



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica
Superior d'Enginyeria
Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica
Universitat Politècnica de València

Análisis del diseño de visualización interactiva de información

Trabajo Fin de Máster

Máster Universitario en Gestión de la Información

Autor: Mauro Attardi

Tutor: José Antonio Ontalba Ruipérez

Cotutor: Jorge Ignacio Serrano Cobos

Septiembre de 2016

Resumen

En el siguiente trabajo, titulado “Análisis del diseño de visualización interactiva de información”, se estudiará la visualización de información como disciplina, destacando las diferencias entre infografía, la visualización de datos estática y la visualización de datos dinámica o interactiva, para después analizar los distintos tipos de gráficos que se pueden utilizar en el apartado concreto de la Visualización Interactiva de Información.

Ya que el objetivo del trabajo es identificar la tipología de visualización de datos más adecuada al análisis de redes sociales multinivel, se compararán además las características de los programas para el análisis de redes sociales y finalmente se desarrollarán varios ejemplos de visualización gráfica interactiva de las redes sociales multinivel, en el contexto del mundo universitario español.

Palabras clave: Visualización de información, grafos, storytelling, red social multinivel, infografía

Abstract

The following MSc thesis, titled "Analysis of the design of interactive visualization of information", studies the visualization of information as a discipline, highlighting the differences between infographic, static data visualization and dynamic or interactive data visualization, in order to analyze the different types of graphics which can be used in the interactive visualization of information.

Since the main objective of this MSc thesis is to identify the most adequate typology of data visualization for the analysis of multilevel social networks, the characteristics of programs for social network analysis will be compared and finally some examples of interactive graphic visualization of multilevel social networks will be developed, within the framework of the Spanish university community.

Keywords: Information visualization, graph, storytelling, multilevel social network, infographic

Tabla de contenidos

| | | |
|-------|--|----|
| 1 | Introducción | 12 |
| 1.1 | Objetivo general | 12 |
| 1.2 | Objetivos Específicos..... | 12 |
| 2 | Metodología | 13 |
| 3 | La visualización de información como disciplina..... | 15 |
| 3.1 | El crecimiento de la información, introducción al Big Data | 15 |
| 3.1.1 | ¿Qué es Big Data?..... | 15 |
| 3.1.2 | ¿Para qué sirve el Big Data?..... | 16 |
| 3.1.3 | ¿Hacia dónde nos lleva?..... | 16 |
| 3.2 | Visualización de (Big) Data..... | 17 |
| 3.2.1 | Que es la Visualización de Datos..... | 17 |
| 3.2.2 | En qué consiste | 18 |
| 3.3 | Diferentes maneras de expresar los datos | 20 |
| 3.3.1 | Infografía | 21 |
| 3.3.2 | Visualización de datos estática | 26 |
| 3.3.3 | Visualización de datos dinámica / interactiva | 29 |
| 3.4 | Historia moderna de la visualización de información | 31 |
| 3.4.1 | John Snow y la mejor infografía de la historia | 31 |
| 3.4.2 | Comienzos del 1900, el método gráfico a sus albores | 32 |
| 3.4.3 | La abstracción de la realidad..... | 33 |
| 3.4.4 | Avances tecnológicos: la informática | 33 |
| 3.5 | Los gurús de la visualización de la información | 35 |
| 3.5.1 | Edward Tufte..... | 35 |
| 3.5.2 | Stephen Few | 38 |
| 3.5.3 | Tufte, el artista y Few, el ingeniero | 38 |
| 3.6 | Empresas más prometedoras en la Visualización de Datos..... | 39 |
| 3.7 | Los Principales tipos de Visualización de Datos actuales | 41 |
| 3.7.1 | Infografía interactiva | 41 |
| 3.7.2 | Infografía creada analizando tendencias..... | 42 |
| 3.7.3 | Infografías animadas | 43 |
| 3.7.4 | Infografía con imágenes reales | 43 |
| 3.7.5 | Infografías que utilizan metáforas | 44 |
| 3.7.6 | Infografías que relacionan datos | 45 |



| | | |
|-------|---|-----|
| 3.7.7 | Infografía con perspectiva variable | 46 |
| 3.7.8 | Infografía que explican un proceso (o cuentan una historia) | 47 |
| 3.7.9 | Infografías que proporcionan accesos a los “Raw data” | 47 |
| 4 | Los tipos de gráficos | 49 |
| 4.1 | Los gráficos | 49 |
| 4.2 | Los tipos de datos..... | 50 |
| 4.2.1 | Datos Cualitativos | 50 |
| 4.2.2 | Datos Cuantitativos | 50 |
| 4.3 | Los principales tipos de gráficos..... | 51 |
| 4.3.1 | Gráfico de barras | 51 |
| 4.3.2 | Gráfico de líneas | 56 |
| 4.3.3 | Relación entre gráficos de líneas y gráficos de barras..... | 56 |
| 4.3.4 | Gráfico de sectores..... | 57 |
| 4.3.5 | Gráfico de dispersión..... | 58 |
| 4.3.6 | Cartograma | 60 |
| 5 | Las redes sociales y el análisis de redde sociales | 61 |
| 5.1 | Que es una Red Social..... | 62 |
| 5.1.1 | Elementos de una red social | 63 |
| 5.2 | Análisis de redes sociales..... | 68 |
| 5.2.1 | Métricas | 69 |
| 5.2.2 | Aplicación en la vida real..... | 71 |
| 5.3 | Redes Sociales Multinivel..... | 79 |
| 5.4 | Software para la Análisis de Social Network..... | 80 |
| 5.4.1 | Funcionalidades principales | 81 |
| 5.4.2 | Programas populares..... | 84 |
| 5.4.3 | Comparativas | 86 |
| 6 | Aplicación de visualización de datos..... | 103 |
| 6.1 | Ejemplo 1: Tree Map con Timeline | 104 |
| 6.1.1 | El gráfico Tree Map de la librería D3plus | 104 |
| 6.1.2 | Datos a disposición | 105 |
| 6.1.3 | Código del gráfico “Tree Map” en la librería D3plus | 106 |
| 6.1.4 | Mapping con el software Ebimap | 108 |
| 6.1.5 | Visualización del gráfico | 111 |
| 6.1.6 | Errores de la librería | 113 |
| 6.2 | Ejemplo 2: Red de enlaces entre 9 de las Universidades Españolas más importantes..... | 114 |



| | | |
|-------|---|-----|
| 6.2.1 | Descubrir el tamaño de las páginas web..... | 115 |
| 6.2.2 | Calcular la cantidad de enlaces desde una universidad hacia otra..... | 116 |
| 6.2.3 | Código del gráfico “Network” en la librería D3plus..... | 120 |
| 6.2.4 | Gráficos resultantes..... | 125 |
| 6.3 | Ejemplo 3: Creación de una simple red multinivel. Red de enlaces entre nueve de las Universidades Españolas más importantes y entre sus departamentos de matemáticas..... | 131 |
| 6.3.1 | Creación de los nodos de los departamentos de matemáticas..... | 133 |
| 6.3.2 | Creación de las aristas del gráfico “network”..... | 135 |
| 6.3.3 | Gráficos resultantes..... | 139 |
| 7 | Conclusiones..... | 145 |
| 8 | Referencias..... | 147 |

Tabla de ilustraciones

| | |
|---|----|
| ILUSTRACIÓN 1 CRECIMIENTO DE LOS DATA CALCULADO EN ZETTABYTE DEL 2005 AL 2019 (HIPOTESIS), 30 DE AGOSTO DE 2016, DE HTTP://WWW.NOJITTER.COM/POST/240170228/THE-NETWORK-IMPACT-OF-BIG-DATA | 16 |
| ILUSTRACIÓN 2 BÚSQUEDA DE LA PALABRA "DATA VISUALIZATION" EN GOOGLE TRENDS, 30 DE AGOSTO DE 2016, DE HTTPS://WWW.GOOGLE.COM/TRENDS/EXPLORE?DATE=ALL&Q=DATA%20VISUALIZATION | 17 |
| ILUSTRACIÓN 3 EJEMPLO DE INFOGRAFÍA, 30 DE AGOSTO DE 2016, DE HTTP://FLOWINGDATA.COM/2010/12/09/PICTURING-SOCIAL-ORDER/ | 19 |
| ILUSTRACIÓN 4 BÚSQUEDA DE LA PALABRA "INFOGRAFÍA" EN GOOGLE TRENDS, 30 DE AGOSTO DE 2016, DE HTTPS://WWW.GOOGLE.COM/TRENDS/EXPLORE?DATE=ALL&Q=INFOGRAPHIC | 21 |
| ILUSTRACIÓN 5 EJEMPLO DE INFOGRAFÍA ESTADÍSTICA, 30 DE AGOSTO DE 2016, DE HTTP://EN.INMOREAU.COM/WP-CONTENT/UPLOADS/2011/05/INTERNET-STATISTICS-TN.JPG | 22 |
| ILUSTRACIÓN 6 EJEMPLO DE INFOGRAFÍA TEMPORAL SOBRE LA HISTORIA DE LA NUTELLA, 30 DE AGOSTO DE 2016, DE HTTPS://S-MEDIA-CACHE-AK0.PINIMG.COM/564X/12/D3/37/12D337E0FB21A1BFE18BF35E21ECEFAD.JPG ... | 23 |
| ILUSTRACIÓN 7 EJEMPLO DE INFOGRAFÍA CON LA RECETA DE LA PIZZA, 30 DE AGOSTO DE 2016, DE HTTP://ENCYCLOPPIZZERIA.COM/WP-CONTENT/UPLOADS/2014/09/E3C0342A287D7E7AFB8A9499D505905E.JPG | 24 |
| ILUSTRACIÓN 8 EJEMPLO DE INFOGRAFÍA BASADA EN LA LOCALIZACIÓN QUE REPRESENTA LOS PAÍS MÁS CORRUPTOS DE EUROPA EN 2015, 30 DE AGOSTO DE 2016, DE HTTP://WWW.VOXEUROP.EU/FILES/IMAGES/BRIEF/CORRUPTION.JPG | 25 |
| ILUSTRACIÓN 9 INFOGRAFÍA BIOGRÁFICA DE MINA MAZZINI, UNA DE LAS CANTANTES ITALIANAS MÁS CONOCIDA, 30 DE AGOSTO DE 2016, DE HTTP://WWW.ESTACIONDISEÑO.ES/PORTFOLIO/INFOGRAFIA-BIOGRÁFICA/# | 25 |
| ILUSTRACIÓN 10 EJEMPLOS DE LOS GRÁFICOS MÁS CONOCIDOS, 30 DE AGOSTO DE 2016, DE HTTP://2.BP.BLOGSPOT.COM/-0PNOZSBHYPA/VcJ3iLP6lKI/AAAAAAAAAD8/y4LwZwDgW_8/s1600/FUSIONCHARTGRAPH.JPG | 26 |
| ILUSTRACIÓN 11 PARTE DE UNA INFOGRAFÍA QUE DA 5 RAZONES SEGÚN LAS CUALES ES MEJOR SER FREELANCE, 30 DE AGOSTO DE 2016, DE HTTPS://S-MEDIA-CACHE-AK0.PINIMG.COM/564X/F4/9D/69/F49D691D7D45D0B8CCD2A0628A7BE587.JPG | 27 |
| ILUSTRACIÓN 12 POPULARIDAD DE FACEBOOK EN TODO EL MUNDO SEGÚN SU USO, 30 DE AGOSTO DE 2016, DE HTTP://WWW.KEYWEBMETRICS.COM/2013/07/BIG-DATA-VISUALIZATIONS/ | 29 |
| ILUSTRACIÓN 13 EJEMPLO DE UNA DASHBOARD, 30 DE AGOSTO DE 2016, DE HTTP://PLOTTING-SUCCESS.SOFTWAREADVICE.COM/FILES/2013/08/SISENSE-INTERACTIVE-DASHBOARD.JPG | 30 |
| ILUSTRACIÓN 14 MAPA DONDE SNOW INDICÓ LOS FALLECIDO EN LA EPIDEMIA DE CÓLERA DE 1854, 30 DE AGOSTO DE 2016, DE HTTP://WWW.YOROKOBU.ES/WP-CONTENT/UPLOADS/COLERA.GIF | 31 |
| ILUSTRACIÓN 15 ZOOM DEL MAPA DONDE SNOW INDICÓ LOS FALLECIDOS EN LA EPIDEMIA DE CÓLERA DE 1854, 30 DE AGOSTO DE 2016, DE HTTP://BLOGS.ELPAIS.COM/.A/6A00D8341BFB1653EF01538DD8656F970B-PI | 32 |
| ILUSTRACIÓN 16 MAPA DEL PLANO DEL METRO DE LONDRES DE LOS AÑOS 30, 30 DE AGOSTO DE 2016, DE HTTPS://ELPASTELDESAGAN.FILES.WORDPRESS.COM/2015/01/1933-UNDERGROUND-MAP.JPG | 33 |
| ILUSTRACIÓN 17 EJEMPLO DE GRÁFICO SPARKLINES DE TUFTE, 30 DE AGOSTO DE 2016, DE HTTPS://EN.WIKIPEDIA.ORG/WIKI/SPARKLINE | 36 |
| ILUSTRACIÓN 18 PARTE DEL INFORME QUE LOS INGENIEROS DE LA THIOKOL ENVIRON A LA NASA, 30 DE AGOSTO DE 2016, DE HTTP://COLLEGE.CENGAGE.COM/HISTORY/US/RESOURCES/SHARED/PRIMARY/SOURCE/IMAGES/CHALLENGER2.GIF | 37 |
| ILUSTRACIÓN 19 GRÁFICO ELABORADO POR TUFTE DONDE SE REPRESENTA DE MANERA MÁS CLARA LA PROBABILIDAD DE RUPTURA DE LOS JUNTOS DEPENDIENDO DE LAS TEMPERATURAS, 30 DE AGOSTO DE 2016, DE HTTP://WWW.WEBSIMPLICITY.CO.ZA/I/E/EDWARD-TUFTE-CHALLENGER-SCATTERPLOT-CHART-3.JPG | 37 |
| ILUSTRACIÓN 20 LISTA DE LAS 20 EMPRESAS MÁS PROMETEDORAS SEGÚN LA REVISTA CIO REVIEW, 30 DE AGOSTO DE 2016, DE HTTP://AVANXO.COM/WP-CONTENT/UPLOADS/2014/06/CIO-REVIEW-2014-MOST-PROMISING-CLOUD-COMPANIES-AVANXO.PDF | 40 |
| ILUSTRACIÓN 21 EJEMPLO DE INFOGRAFÍA INTERACTIVA, 30 DE AGOSTO DE 2016, DE HTTPS://PODIO.COM/SITE/CREATIVE-ROUTINES | 42 |
| ILUSTRACIÓN 22 EJEMPLO DE INFOGRAFÍA CREADA ANALIZANDO TENDENCIAS, "THE YEAR IN NEWS", 30 DE AGOSTO DE 2016, DE HTTP://ECHELONINSIGHTS.COM/WP-CONTENT/UPLOADS/2014/12/THEYEARINNEWS20141.PNG | 42 |
| ILUSTRACIÓN 23 EJEMPLOS DE INFOGRAFÍA ANIMADA, 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWRESEARCH.ORG/FILES/2014/04/847889448.GIF | 43 |
| ILUSTRACIÓN 24 EJEMPLOS DE INFOGRAFÍA CON IMÁGENES REALES, 30 DE AGOSTO DE 2016, DE HTTP://SNIP.LY/VF80#HTTP://MARION-LUTTENBERGER.SQUARESPACE.COM/#/INFOGRAPHICS/ | 44 |



| | |
|--|----|
| ILUSTRACIÓN 25 EJEMPLO DE INFOGRAFÍA QUE UTILIZA METÁFORAS, 30 DE AGOSTO DE 2016, DE HTTPS://SITES.GOOGLE.COM/A/MTLSD.NET/INFOGRAPHICS-SIMPLE/SERVICES/METAPHOR-BASED-INFOGRAPHICS ... | 45 |
| ILUSTRACIÓN 26 EJEMPLO DE INFOGRAFÍA QUE RELACIONA DATOS, 30 DE AGOSTO DE 2016, DE HTTP://WWW.CONCERTHOTELS.COM/100-YEARS-OF-ROCK/ | 46 |
| ILUSTRACIÓN 27 EJEMPLO DE INFOGRAFÍA CON PERSPECTIVA VARIABLE, 30 DE AGOSTO DE 2016, DE HTTP://HEREISTODAY.COM/ | 46 |
| ILUSTRACIÓN 28 EJEMPLO DE INFOGRAFÍA QUE EXPLICA UN PROCESO, 30 DE AGOSTO DE 2016, DE HTTP://WWW.BIZBRAIN.ORG/COFFEE/ | 47 |
| ILUSTRACIÓN 29 EJEMPLO DE INFOGRAFÍAS QUE PROPORCIONAN ACCESOS A LOS “ROW DATA”, 30 DE AGOSTO DE 2016, DE HTTP://BLOG.VISME.CO/EXAMPLES-DATA-VISUALIZATIONS/ | 48 |
| ILUSTRACIÓN 30 EJEMPLO DE UN GRÁFICO DE BARRAS..... | 51 |
| ILUSTRACIÓN 31 EJEMPLO DE UN GRÁFICO DE BARRAS DONDE SE QUIERE FOCALIZAR SOBRE LA EVOLUCIÓN EN EL TIEMPO | 51 |
| ILUSTRACIÓN 32 EJEMPLO DE GRÁFICO DE BARRAS HORIZONTAL..... | 52 |
| ILUSTRACIÓN 33 EJEMPLO DE GRÁFICO DE BARRAS AGRUPADO..... | 52 |
| ILUSTRACIÓN 34 EJEMPLO DE UN GRÁFICO DE BARRAS APILADO..... | 53 |
| ILUSTRACIÓN 35 EJEMPLO DE UN HISTOGRAMA..... | 53 |
| ILUSTRACIÓN 36 EJEMPLO DE HISTOGRAMA BI-DIRECCIONAL..... | 54 |
| ILUSTRACIÓN 37 EJEMPLO DE PIRÁMIDE DE POBLACIÓN | 54 |
| ILUSTRACIÓN 38 EJEMPLO DE PIRÁMIDE DE POBLACIÓN PROGRESIVA | 55 |
| ILUSTRACIÓN 39 EJEMPLO DE PIRÁMIDE DE POBLACIÓN REGRESIVA..... | 55 |
| ILUSTRACIÓN 40 EJEMPLO DE PIRÁMIDE DE POBLACIÓN ESTANCADA..... | 56 |
| ILUSTRACIÓN 41 EJEMPLO DE GRÁFICO DE LINEAS | 56 |
| ILUSTRACIÓN 42 EJEMPLO DE POLÍGONO DE FRECUENCIA | 57 |
| ILUSTRACIÓN 43 EJEMPLO DE GRÁFICO DE SECTORES..... | 58 |
| ILUSTRACIÓN 44 EJEMPLO DE GRÁFICO DE SECTORES DONDE SE VISUALIZA EL PORCENTAJE DE SUPERFICIE DE LOS PRINCIPALES PAÍSES DE LA UE | 58 |
| ILUSTRACIÓN 45 EJEMPLO DE GRÁFICO DE DISPERSIÓN CON CORRELACIÓN NULA..... | 59 |
| ILUSTRACIÓN 46 EJEMPLO DE GRÁFICO DE DISPERSIÓN CON CORRELACIÓN LINEAL..... | 59 |
| ILUSTRACIÓN 47 EJEMPLO DE RELACIÓN LINEAL POSITIVA | 59 |
| ILUSTRACIÓN 48 EJEMPLO DE CARTOGRAMA | 60 |
| ILUSTRACIÓN 49 EJEMPLO DE GRAFO DE UNA RED SOCIAL..... | 62 |
| ILUSTRACIÓN 50 EJEMPLO DE UNA RED MODO 1 | 63 |
| ILUSTRACIÓN 51 EJEMPLO DE UNA RED MODO 2 | 64 |
| ILUSTRACIÓN 52 TABLA DE DEFINICIONES PLEX-MODES..... | 65 |
| ILUSTRACIÓN 53 EJEMPLO DE RED CON 2 TIPOS DE NATURALEZA DE LOS ACTORES | 66 |
| ILUSTRACIÓN 54 EJEMPLO DE RED MULTIDIMENSIONAL PARCIAL Y NO COMPLETA..... | 67 |
| ILUSTRACIÓN 55 JACOB LEVY MORENO, 30 DE AGOSTO DE 2016, DE HTTP://FILES.LECTURAS- MIC2.WEBNODE.ES/SYSTEM_PREVIEW_DETAIL_200000062-982499921F/MORENO.PNG | 68 |
| ILUSTRACIÓN 56 GRÁFICO EJEMPLO DE INTERMEDIACIÓN..... | 71 |
| ILUSTRACIÓN 57 GRAFO CON LA MEDIA DE LOS PASES DE LOS DOS FINALISTAS, ESPAÑA Y HOLANDA, 30 DE AGOSTO DE 2016, DE HTTPS://C2.STATICFLICKR.COM/4/3719/13952898494_48C9EA8046_B.JPG | 72 |
| ILUSTRACIÓN 58 MULTITUD POLARIZADA (POLARIZED CROWD), 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWINTERNET.ORG/2014/02/20/MAPPING-TWITTER-TOPIC-NETWORKS-FROM-POLARIZED-CROWDS- TO-COMMUNITY-CLUSTERS/ | 73 |
| ILUSTRACIÓN 59 MULTITUD APRETADA (TIGHT CROWD), 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWINTERNET.ORG/2014/02/20/MAPPING-TWITTER-TOPIC-NETWORKS-FROM-POLARIZED-CROWDS- TO-COMMUNITY-CLUSTERS/ | 74 |
| ILUSTRACIÓN 60 AGRUPACIONES DE MARCAS (BRAND CLUSTERS), 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWINTERNET.ORG/2014/02/20/MAPPING-TWITTER-TOPIC-NETWORKS-FROM-POLARIZED-CROWDS- TO-COMMUNITY-CLUSTERS/ | 75 |
| ILUSTRACIÓN 61 CLUSTER DE COMUNIDADES (COMMUNITY CLUSTERS), 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWINTERNET.ORG/2014/02/20/MAPPING-TWITTER-TOPIC-NETWORKS-FROM-POLARIZED-CROWDS- TO-COMMUNITY-CLUSTERS/ | 76 |
| ILUSTRACIÓN 62 RED DE DIFUSIÓN (BROADCAST NETWORKS), 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWINTERNET.ORG/2014/02/20/MAPPING-TWITTER-TOPIC-NETWORKS-FROM-POLARIZED-CROWDS- TO-COMMUNITY-CLUSTERS/ | 77 |



| | |
|--|-----|
| ILUSTRACIÓN 63 RED DE SOPORTE (SUPPORT NETWORK), 30 DE AGOSTO DE 2016, DE HTTP://WWW.PEWINTERNET.ORG/2014/02/20/MAPPING-TWITTER-TOPIC-NETWORKS-FROM-POLARIZED-CROWDS-TO-COMMUNITY-CLUSTERS/ | 78 |
| ILUSTRACIÓN 64 REPRESENTACIÓN DE UNA RED MULTINIVEL, EN ESTE CASO DE DOS NIVELES | 79 |
| ILUSTRACIÓN 65 EJEMPLO DE LISTA DE ADYACENCIA Y MATRIZ DE ADYACENCIA, 01 SEPTIEMBRE DE 2016, HTTP://WIC.LITISLAB.FR/2010/SLIDES/COMBE_WIVE10_SLIDES.PDF | 81 |
| ILUSTRACIÓN 66 ORIENTACIÓN DE LOS ENLACES ENTRE NODOS | 82 |
| ILUSTRACIÓN 67 PESO EN LAS RELACIONES | 82 |
| ILUSTRACIÓN 68 RED CON DIFERENTES TIPOS DE NODOS | 82 |
| ILUSTRACIÓN 69 EJEMPLO DE UN GRAFO CENTRADO EN UN SOLO ACTOR | 83 |
| ILUSTRACIÓN 70 CAPTURA DE PANTALLA DE LA INTERFAZ GRÁFICA DEL PROGRAMA UCINET | 84 |
| ILUSTRACIÓN 71 CAPTURA DE PANTALLA DEL PROGRAMA PROPIETARIO POLINODE | 85 |
| ILUSTRACIÓN 72 CAPTURA DE PANTALLA DEL SOFTWARE NETMINER | 86 |
| ILUSTRACIÓN 73 CAPTURAS DE PANTALLA DE PAJEK. | 88 |
| ILUSTRACIÓN 74 CAPTURA DE PANTALLA DE GEPHI. | 89 |
| ILUSTRACIÓN 75 EJEMPLO DE MATRIZ DE ADYACENCIA..... | 90 |
| ILUSTRACIÓN 76 GRÁFICO REALIZADO CON LA LIBRERÍA IGRAPH | 90 |
| ILUSTRACIÓN 77 GRÁFICO REALIZADO CON LA LIBRERÍA STATNET | 91 |
| ILUSTRACIÓN 78 VISUALIZACIÓN PATHS TO THE WHITE HOUSE, 30 DE AGOSTO DE 2016, DE HTTP://ELECTIONS.NYTIMES.COM/2012/RESULTS/PRESIDENT/SCENARIOS | 93 |
| ILUSTRACIÓN 79 ELECCIÓN DEL CANDIDATO QUE SE PIENSA QUE GANE EN LA VISUALIZACIÓN PATHS TO THE WHITE HOUSE, 30 DE AGOSTO DE 2016, DE HTTP://ELECTIONS.NYTIMES.COM/2012/RESULTS/PRESIDENT/SCENARIOS | 93 |
| ILUSTRACIÓN 80 RESULTADO DE LA ELECCIÓN DEL CANDIDATO QUE SE PIENSA QUE GANE EN LA VISUALIZACIÓN PATHS TO THE WHITE HOUSE, 30 DE AGOSTO DE 2016, DE HTTP://ELECTIONS.NYTIMES.COM/2012/RESULTS/PRESIDENT/SCENARIOS | 94 |
| ILUSTRACIÓN 81 EJEMPLO RESULTANTE DE ELEGIR QUE LOS DEMOCRÁTICOS GANARAN EN VIRGINIA Y LOS REPUBLICANOS EN FLORIDA Y NORTH CAROLINA EN LA VISUALIZACIÓN PATHS TO THE WHITE HOUSE, 30 DE AGOSTO DE 2016, DE HTTP://ELECTIONS.NYTIMES.COM/2012/RESULTS/PRESIDENT/SCENARIOS | 95 |
| ILUSTRACIÓN 82 VISUALIZACIÓN DEL TAMAÑO DE LA INDUSTRIA MANUFACTURERA CHINA, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/04/08/BUSINESS/GLOBAL/ASIA-MAP.HTML?_R=0 | 96 |
| ILUSTRACIÓN 83 SELECCIÓN DE UN PAÍS EN LA VISUALIZACIÓN DEL TAMAÑO DE LA INDUSTRIA MANUFACTURERA CHINA, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/04/08/BUSINESS/GLOBAL/ASIA-MAP.HTML?_R=0 | 97 |
| ILUSTRACIÓN 84 VISUALIZACIÓN DE LAS CONEXIONES ENTRE LOS NOMINADOS A LOS OSCAR DE 2013, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/02/20/MOVIES/AMONG-THE-OSCAR-CONTENDERS-A-HOST-OF-CONNECTIONS.HTML) | 98 |
| ILUSTRACIÓN 85 SELECCIÓN DE UNA PERSONA EN LA VISUALIZACIÓN DE LAS CONEXIONES ENTRE LOS NOMINADOS A LOS OSCAR DE 2013, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/02/20/MOVIES/AMONG-THE-OSCAR-CONTENDERS-A-HOST-OF-CONNECTIONS.HTML | 98 |
| ILUSTRACIÓN 86 PASANDO EL RATÓN EN UNA PELÍCULA EN LA VISUALIZACIÓN DE LAS CONEXIONES ENTRE LOS NOMINADOS A LOS OSCAR DE 2013, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/02/20/MOVIES/AMONG-THE-OSCAR-CONTENDERS-A-HOST-OF-CONNECTIONS.HTML) | 99 |
| ILUSTRACIÓN 87 VISUALIZACIÓN DE LAS EMPRESAS DEL 500 S&P, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/05/25/SUNDAY-REVIEW/CORPORATE-TAXES.HTML | 100 |
| ILUSTRACIÓN 88 LEYENDA DEL TAMAÑO DE LOS CÍRCULOS EN LA VISUALIZACIÓN SOBRE LAS EMPRESAS DEL 500 S&P | 100 |
| ILUSTRACIÓN 89 LEYENDA DE LOS COLORES EN LA VISUALIZACIÓN SOBRE LAS EMPRESAS DEL 500 S&P, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/05/25/SUNDAY-REVIEW/CORPORATE-TAXES.HTML | 100 |
| ILUSTRACIÓN 90 ZOOM SOBRE UNA EMPRESA EN LA VISUALIZACIÓN SOBRE LAS EMPRESAS DEL 500 S&P, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/05/25/SUNDAY-REVIEW/CORPORATE-TAXES.HTML | 101 |
| ILUSTRACIÓN 91 VISUALIZACIÓN POR SECTORES EN LA VISUALIZACIÓN SOBRE LAS EMPRESAS DEL 500 S&P, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/05/25/SUNDAY-REVIEW/CORPORATE-TAXES.HTML | 101 |
| ILUSTRACIÓN 92 EJEMPLO DE BÚSQUDA EN LA VISUALIZACIÓN SOBRE LAS EMPRESAS DEL 500 S&P, 30 DE AGOSTO DE 2016, DE HTTP://WWW.NYTIMES.COM/INTERACTIVE/2013/05/25/SUNDAY-REVIEW/CORPORATE-TAXES.HTML | 102 |
| ILUSTRACIÓN 93 CAPTURA DE PANTALLA DEL APARTADO “EJEMPLOS” DE LA PÁGINA WEB D3PLUS.ORG, 30 DE AGOSTO DE 2016, DE HTTP://D3PLUS.ORG/EXAMPLES/ | 103 |



| | |
|---|-----|
| ILUSTRACIÓN 94 EJEMPLO DE TREE MAP CON LA LIBRERÍA D3PLUS..... | 104 |
| ILUSTRACIÓN 95 GESTIÓN DE LAS FECHAS EN UN TREE MAP | 105 |
| ILUSTRACIÓN 96 FICHERO CON LOS DATOS EN EL PROGRAMA MICROSOFT EXCEL. | 105 |
| ILUSTRACIÓN 97 EJEMPLO DE MAPEO CON EL SOFTWARE EBIMAP | 108 |
| ILUSTRACIÓN 98 FICHERO CON LOS DATOS EN EL PROGRAMA MICROSOFT EXCEL. | 109 |
| ILUSTRACIÓN 99 FICHERO CON LOS DATOS ABIERTOS CON UN EDITOR DE TEXTO (NOTEPAD++) | 109 |
| ILUSTRACIÓN 100 INTERFAZ DE ORIGEN DEL MAPA EN EBIMAP | 109 |
| ILUSTRACIÓN 101 FICHERO DE ORIGEN CSV ABIERTO CON LA INTERFAZ DEL MAPA EN EL SOFTWARE EBIMAP..... | 110 |
| ILUSTRACIÓN 102 GRÁFICO TREE MAP CON LOS DATOS DEL 2011 | 112 |
| ILUSTRACIÓN 103 GRÁFICO TREE MAP CON LOS DATOS DEL 2010 | 112 |
| ILUSTRACIÓN 104 GRÁFICO TREE MAP CON LOS DATOS DEL 2009 | 113 |
| ILUSTRACIÓN 105 ERROR EN EL GRÁFICO TREE MAP | 113 |
| ILUSTRACIÓN 106 BÚSQUEDA CON EL COMANDO "SITE:" DE GOOGLE EN EL DOMINIO UPM.ES | 115 |
| ILUSTRACIÓN 107 RESULTADO DE LAS QUERIES CON EL COMANDO DE GOOGLE "SITE:" | 115 |
| ILUSTRACIÓN 108 CREAMOS UN INFORME EN MAJESTIC SEO | 117 |
| ILUSTRACIÓN 109 SELECCIÓN DE DESCARGA DE INFORME AVANZADO | 118 |
| ILUSTRACIÓN 110 FICHERO CSV GENERADO POR MAJESTIC SEO CON LOS BACKLINKS DE WWW.UPV.ES | 118 |
| ILUSTRACIÓN 111 BACKLINKS DEL DOMINIO UB.EDU QUE VIENEN DE LAS 8 UNIVERSIDADES | 119 |
| ILUSTRACIÓN 112 RESULTADO DE LAS QUERIES CON EL COMANDO DE GOOGLE "SITE:" ABIERTO CON MICROSOFT EXCEL... 120 | 120 |
| ILUSTRACIÓN 113 RESULTADO DE LAS QUERIES CON EL COMANDO DE GOOGLE "SITE:" ABIERTO CON EL NOTEPAD++ | 121 |
| ILUSTRACIÓN 114 MODELO DE LO QUE TIENE QUE REALIZAR EL MAPA CREADO PARA FORMATEAR LAS ARISTAS | 122 |
| ILUSTRACIÓN 115 INTERFAZ DE ORIGEN DEL MAPA USADO EN EL GRÁFICO NETWORK | 123 |
| ILUSTRACIÓN 116 ARISTAS EN EL FORMATO REQUERIDO POR LA LIBRERÍA D3PLUS EN EL GRÁFICO "NETWORK" | 123 |
| ILUSTRACIÓN 117 GRÁFICO DE TIPO NETWORK GENERADO CON LA LIBRERÍA D3PLUS QUE REPRESENTA LOS ENLACES ENTRE LAS PÁGINAS WEB DE LAS 10 UNIVERSIDAD ESPAÑOLA MÁS IMPORTANTES, SIN CONSIDERAR SUBDOMINIOS (EJ. WWW.UPV.ES/*)..... | 125 |
| ILUSTRACIÓN 118 PASANDO EL RATÓN EN UN NODO, SE EVIDENCIAN LAS ARISTAS QUE LO INVOLUCRAN | 126 |
| ILUSTRACIÓN 119 COMPORTAMIENTO DEL GRÁFICO "NETWORK" CUANDO SE HACE CLICK EN UN NODO | 127 |
| ILUSTRACIÓN 120 GRÁFICO CON LA OPCIÓN "LABEL" DE LAS ARISTAS ACTIVADA. | 128 |
| ILUSTRACIÓN 121 EJEMPLOS DE CÓMO ACTÚA LA LIBRERÍA D3PLUS PARA POSICIONAR DINÁMICAMENTE LOS NODOS DE UN GRÁFICO DE TIPO "NETWORK" | 130 |
| ILUSTRACIÓN 122 ZOOM HACIA DOS NODOS DE DOS NIVELES DIFERENTES | 132 |
| ILUSTRACIÓN 123 EJECUCIÓN DE LA QUERY SITE:MAT.UB.EDU/ | 133 |
| ILUSTRACIÓN 124 URL DE CADA DEPARTAMENTO DE MATEMÁTICAS Y EL RESULTADO DE LA QUERY "SITE:" | 133 |
| ILUSTRACIÓN 125 BÚSQUEDA EN EL FICHERO CON LOS BACKLINKS DE LA PÁGINA WEB MAT.UB.EDU/ DE TODOS LOS ENLACES PROVENIENTES DE MATEMATICAS.UCM.ES | 135 |
| ILUSTRACIÓN 126 CAPTURA DE PANTALLA DEL EXCEL CON LOS VALORES DE LAS ARISTAS DE LA SUBRED DE WEBSITES DE DEPARTAMENTOS DE MATEMÁTICAS | 137 |
| ILUSTRACIÓN 127 EXPLICACIÓN DE LA FUNCIÓN DEL MAPA PARA INSERTAR LAS ARISTAS EN EL ARRAY..... | 137 |
| ILUSTRACIÓN 128 ARISTAS DEL ARRAY CONNECTIONS[] | 138 |
| ILUSTRACIÓN 129 GRÁFICOS DE TIPO NETWORK DE LOS ENLACES ENTRE LAS 9 UNIVERSIDADES Y SUS DEPARTAMENTOS DE MATEMÁTICAS..... | 139 |
| ILUSTRACIÓN 130 RESPUESTA DEL GRÁFICO DE LOS DEPARTAMENTOS DE MATEMÁTICAS CUANDO SE PASA EL RATÓN ENCIMA DE UN NODO, EN ESTE CASO MAT.UB.EDU | 140 |
| ILUSTRACIÓN 131 RESPUESTA DEL GRÁFICO DE LOS DEPARTAMENTOS DE MATEMÁTICAS CUANDO SE HACE CLICK ENCIMA DE UN NODO, EN ESTE CASO MAT.UB.EDU | 141 |
| ILUSTRACIÓN 132 LEYENDA DE LOS COLORES DE LAS REDES..... | 141 |
| ILUSTRACIÓN 133 ORIENTACIÓN DE LOS AXIS OBSERVADA POR LA LIBRERÍA D3PLUS | 142 |
| ILUSTRACIÓN 134 DECLARACIÓN DEL ARRAY POSITIONS[] | 142 |
| ILUSTRACIÓN 135 GRÁFICO CON LAS DOS REDES SUPERPUESTAS | 143 |
| ILUSTRACIÓN 136 COMPARACIÓN ENTRE EL NODO UPV.ES Y SU DEPARTAMENTO DE MATEMÁTICAS, EL NODO UPV.ES/ENTIDADES/DMAA..... | 144 |



1 Introducción

1.1 Objetivo general

El objetivo general de este trabajo es identificar la tipología de visualización de datos más adecuada al análisis de redes sociales multinivel.

1.2 Objetivos Específicos

Los objetivos específicos de este trabajo son los siguientes:

- Describir la evolución de la visualización de la información.
- Comparar las características de los programas para el análisis de redes sociales.
- Desarrollar varios ejemplos de visualización gráfica interactiva de las redes sociales multinivel, en el contexto del mundo universitario español.

2 Metodología

En este apartado se describirá la metodología utilizada para realizar este trabajo.

Para empezar, se van a establecer las bases teóricas de este trabajo, iniciando describiendo el fenómeno del Big data y sus implicaciones en la vida actual, introduciendo la visualización de datos, desde sus albores hasta la actualidad.

En la siguiente fase, se analiza a fondo la visualización de información, empezando por la visualización estática (con sus antecedentes históricos) y dinámica y llegando a analizar las infografías, los tipos que se utilizan y sus funciones.

Después de eso, se pasa a la fase más “matemática” de la cuestión, es decir se analizan los tipos de datos y los gráficos que se suelen utilizar para representar la información.

Siguiendo con la representación matemático-estadística de los datos, se pasa a analizar en detalle las redes sociales, que son y sus elementos, enfocándose en las redes sociales multinivel.

A continuación, quedándose en el tema de las redes sociales, se van a analizar los programas más populares para realizar el análisis de las redes sociales, comparándolos.

Todo este trabajo teórico de análisis y comparación es necesario ya que el fin principal de este trabajo es la identificación de la tipología de visualización de datos más adecuada al análisis de redes sociales multinivel.

En la parte final del trabajo, la parte práctica, se va aplicar los conocimientos adquiridos durante los apartados anteriores, diseñando y programando unas visualizaciones interactivas.

Los datos que se van a utilizar representan varias características de las páginas web de las universidades españolas más importantes.

Se van a utilizar los datos de las universidades españolas más importantes para estar seguros que, ya que su tamaño es muy grande, tengan la cantidad suficiente de información para poderlas analizar.

En el primer ejemplo, el más básico, se desarrollará un tree map con timeline para visualizar el número de páginas web de escuelas que tiene cada universidad españolas durante tres años. Se realizará este ejemplo más que nada para coger practicidad con la librería.

El segundo y tercer ejemplo, se va a diseñar y programar una visualización interactiva de un grafo representante una red social multinivel y se adaptará a un caso para poder validarlo.

En detalle, la red social estará compuesta por las páginas web de nueve de las más importantes universidades españolas y las aristas representan el número de enlaces entre ellas.



Para generar otro nivel en la red, se va también a calcular los enlaces entre los departamentos de matemática de cada una de estas universidades, de esta manera se podrá crear una visualización de una red social multinivel.

3 La visualización de información como disciplina

3.1 El crecimiento de la información, introducción al Big Data

La sociedad actual está viviendo una época con gran volumen de información, y eso es bastante difícil de asimilar.

Para entender mejor este concepto, es interesante ver que desde los años ochenta la capacidad tecnológica de cada ser humano a nivel mundial de almacenar datos se dobla cada cuarenta meses. Esto significa que cada tres años salen nuevos dispositivos tecnológicos que doblan la capacidad de almacenamiento de otro dispositivo igual con tres años de antigüedad. [1]

Todo este incremento desproporcional de la información se puede resumir con el concepto de Big Data. Cada vez generamos más datos en todo el mundo, hasta tal punto de que la tecnología actual es totalmente incapaz de procesarlos, clasificarlos, interpretarlos y utilizarlos.

3.1.1 ¿Qué es Big Data?

Un conjunto masivo de datos en una determinada época se puede definir como Big Data. No es una tecnología o una forma de procesar los datos, más bien es la causa de que se inventen, constantemente, sistemas capaces de gestionarlos.

No es exactamente una novedad de estos últimos tiempos, de hecho el Big Data ha existido siempre.

A finales del siglo XIX, por ejemplo, el censo de los Estados Unidos costó más de ocho años en realizarse. Diez años más tarde, si no se hubiesen desarrollado avances en la metodología, se estima que el nuevo censo no habría acabado antes de que tuviera que realizarse el siguiente, en 1900.

El incremento exponencial de las capacidades tecnológicas a lo largo del siglo XX y lo que llevamos del XXI ha favorecido que cada vez haya un mayor volumen de datos para procesar.

Según el Quinto Global Cloud Index [2], se prevé que el tráfico y el almacenaje de datos en la nube se cuadruplicarán en el mundo hasta 2019 y pasarán de los 2,1 zettabytes (ZB) a los 8,6 (cada uno equivalente a unos 13,8 billones de horas de música). [2]

Ese volumen de datos significa un crecimiento correspondiente de 350% en el tráfico de la red, que puede llevarse a través de redes LAN privadas (por cable e inalámbrica) y WAN, Internet y las redes celulares.



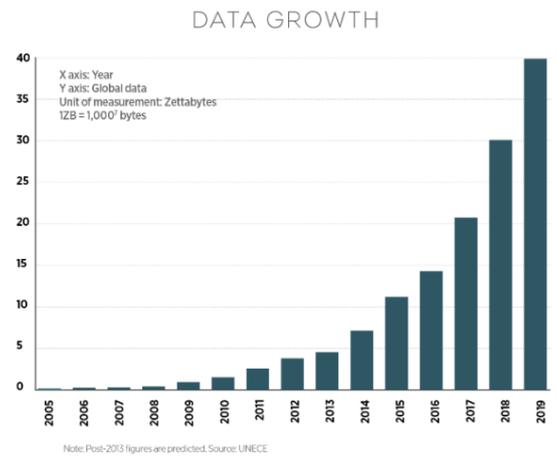


Ilustración 1 Crecimiento de los data calculado en zettabyte del 2005 al 2019 (hipotesis), 30 de agosto de 2016, de <http://www.nojitter.com/post/240170228/the-network-impact-of-big-data>

Como se puede notar de la gráfica, este fenómeno se ve en aumento de manera exponencial. Es importante subrayar que no solo se trata de datos intercambiados por vía electrónica (emails, mensajes etc.) sino también por ejemplo cuando se rellena una hoja de cálculo de Excel, se le hace una foto con el móvil a vuestro perro o simplemente se genera contenido en las redes sociales. Todo esto es información masiva, que acaba creando y alimentando el Big Data.

3.1.2 ¿Para qué sirve el Big Data?

Es imposible encontrar una utilidad a esta gran masa de datos si no está desarrollada la tecnología adecuada.

Afortunadamente en los últimos tiempo se ha credo y desarrollado muchos medios tecnológico que nos permiten almacenar y sobre todo explotar los datos.

Por ejemplo, uno de los sectores donde el Big Data está desarrollando es el del marketing. Las empresas que utilizan el software adecuado para procesar el Big Data hacen uso de él para conocer mejor a sus clientes, con el fin de elaborar ofertas y productos a medida. Sus preferencias, sus gustos, su comportamiento en la red, sus movimientos... Es una información valiosísima.

Otro sector donde el Big Data se está afirmando es el de la medicina.

Según un artículo de la revista Forbes, el Big Data será de muy gran ayuda para prevenir muchas enfermedades y entender la manera más eficiente para curarlas. [3]

3.1.3 ¿Hacia dónde nos lleva?

Hacer una estimación sobre el futuro es prácticamente imposible, aunque está claro que no parece que la generación de datos vaya a reducirse en los próximos años. Seguramente los softwares capaces de darles utilidad también se desarrollarán, como han hecho siempre, y cada vez habrá más información y se expandirán lo escenarios de uso.

3.2 Visualización de (Big) Data

En el precedente apartado se hizo una descripción del Big Data y una introducción a los desafíos que representa. Para seguir analizando este fenómeno no cabe duda que al disponer de cantidades descomunales de información a través de Big Data, uno de los grandes retos que se presenta es cómo representarla.

Por esta razón, el *Main Topic* de este trabajo es la Visualización de los datos.

Buscando la “Data Visualization” en Google Trends se puede ver que es un término muy demandado al día de hoy.

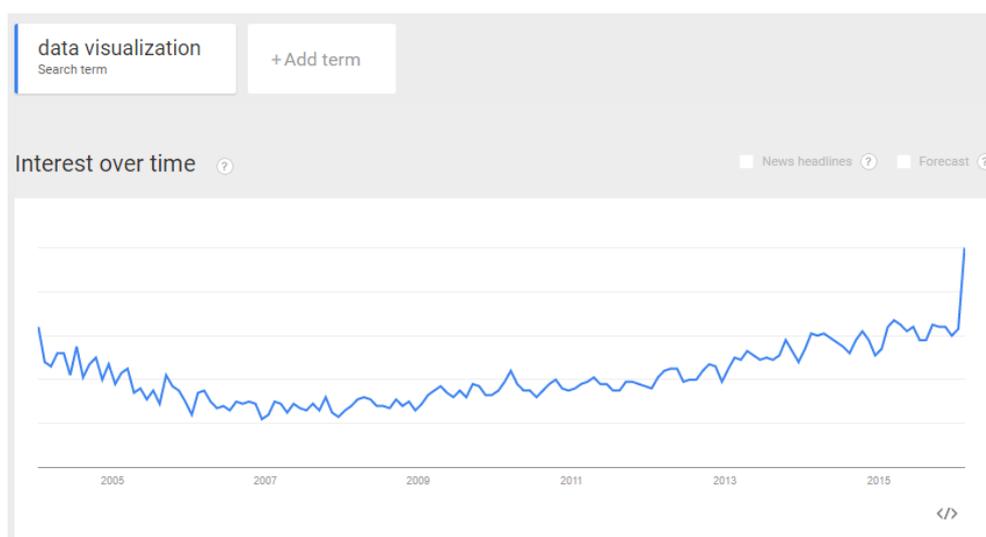


Ilustración 2 Búsqueda de la palabra "Data Visualization" en Google Trends, 30 de agosto de 2016, de <https://www.google.com/trends/explore?date=all&q=data%20visualization>

3.2.1 Que es la Visualización de Datos

La Visualización de Datos se refiere a las técnicas que se usan para comunicar datos o información de cualquier tipo al codificarlos como objetos visuales (por ejemplo puntos, líneas, barras etc.) contenidos en gráficos. Por esta razón, el principal objetivo de la Visualización de Datos es la comunicación de la información de forma clara y eficiente a través de gráficos.

A día de hoy, ya que la mayoría de las empresas disponen de una enorme cantidad de datos por extrapolar, lo que más se busca es la manera de transmitirlos a los clientes. Para transmitir la información de forma efectiva, el diseño, la parte artística y la funcionalidad deben ir de la mano, permitiendo la comunicación de los aspectos claves en la forma más intuitiva y que provoque la curiosidad en el usuario.

La visualización de datos es una de las herramientas indispensables para la representación de la información, de datos y del propio conocimiento en general. Eso consiste en transformar la información en imágenes prácticas y fáciles de entender con el propósito de hacer llegar el mensaje de manera rápida, interesante y comprensible al usuario.

Con la difusión de Internet y la gran cantidad de datos almacenados en la nube – Big Data - la gestión y la visualización de la información se convierten en segmentos protagonistas del marketing de contenidos para incrementar el nivel de empresas y marcas. Las infografías tienen un increíble potencial viral, permitiendo transmitir un mensaje a los clientes (de diferentes tipologías) en poco tiempo.

Al día de hoy las infografías aparecen en todos los medios: periódicos, revistas, medios digitales, webs, blogs. Y además aparece en todos los soportes: periódicos y revistas, televisión, ordenadores, tabletas y Smartphone.

3.2.2 En qué consiste

Hemos visto en los párrafos anteriores que se disponen de multitudes de informaciones relacionadas con cualquier campo de nuestras vidas.

Una de las cuestiones más recurrente en nuestra época es ¿cómo se pueden comunicar a la población de una forma comprensible y sobre todo atractiva la información recopilada (los datos) relacionados, por ejemplo, con los mercados financieros, los resultados de las elecciones políticas o el transporte?

Está claro que estos datos afectan a la vida diaria de la gente, pero son muchísimas tipología de datos y de cada tipología se pueden extrapolar muchas estadísticas y sobre todo muchísimas manera de visualizarlas.

El arte para visualizar los datos no es una labor fácil.

Desde que la visualización de datos empezó a nacer, son muchas las disciplinas que contribuyen a enriquecer la información primaria, es decir, combinar de manera armoniosa ciencia y arte.

Entre estas disciplinas se podría ver que un científico de datos colabora con un diseñador, un publicista con un matemático o un comunicador.

La clave del éxito en este campo es la mezcla de diferentes perfiles que contribuyen a unir los matices artísticos y científicos, ofreciendo una miríada de posibilidades a la hora de moldear los datos.

La visualización de datos tiene dos objetivos generales: reducir y revelar los datos.

Hay que reducir los gráficos ya que se dispone de demasiada cantidad de información. Por esta razón es necesario elaborar gráfico de resumen adecuado, por ejemplo gráfica mensual o anual sobre un aspecto.

Se necesita también revelar los datos, es decir, inventarse gráficos que muestren algo que nunca se ha visto antes.

Hoy en día el mercado necesita el arte visual para implementarlo en ambos tipos de datos. Por esta razón muchas empresas están buscando nuevos métodos para mostrar los datos y se inspiran en el mundo artístico para conseguirlo.

Existe una creciente predisposición a contratar a un profesional creativo con formación en Big Data cuya labor consiste en analizar y visualizar datos de una forma accesible y reveladora. [4]



En definitiva, la parte artística de la visualización es quizás la más importante, ya que hay que comunicar pero sin olvidarse de diferenciarse y llamar la atención de los clientes.



Ilustración 3 Ejemplo de infografía, 30 de agosto de 2016, de <http://flowingdata.com/2010/12/09/picturing-social-order/>

Este nuevo campo no solo se aplica al sector privado.

Por ejemplo, en el sector público un Artista de los Datos puede trabajar con grandes cantidades de datos para recalcar las tendencias y los patrones significativos de la población, la demografía y la salud pública.

Para tener éxito en esta profesión, la persona que se ocupa de esta tarea debe tener una formación en ciencias, matemáticas y Data Visualization.

Sin embargo, según mi opinión, debe también ser creativo, tener fuertes habilidades comunicativas y de marketing y estar familiarizado con una variedad de software de visualización.

3.3 Diferentes maneras de expresar los datos

Como ya se repitió varias veces, para realizar una visualización de la información es muy poco probable que a colaborar sea un solo individuo. Es un proyecto laborioso, que una creatividad, estadística, búsqueda, documentación, gráfica y muchas más especialidades.

Esto se puede analizar más a fondo en el artículo “The 8 hats of data visualisation design” de Andy Kirk, donde se estudian algunos roles profesionales implicados en la visualización de la información. [5]

Según Andy Kirk, se necesitan ocho figuras profesionales, que son las siguientes:

1. El Líder, que es una persona que está buscando una solución a un problema determinado, por simple curiosidad o por oportunidad. Normalmente esta figura es un explorador, que no se conforma con lo estándar y le gusta indagar.
2. El científico de datos es la persona que se ocupa de la adquisición, manipulación y preparación de los datos.
3. El periodista es el narrador, la persona que implanta el enfoque narrativo a lo que se quiere visualizar.
4. El experto de programación, que es el que se ocupa de encontrar la manera más rápida y eficiente de recolectar datos, manipularlos y catalogarlos.
5. El diseñador, que es el que ocupa de desarrollar la parte más artística y visual, dándole una estética atractiva pero sencilla.
6. El científico cognitivo, que es el que se ocupa de asegurarse que el resultado sea eficaz y que sea intuitivo. Tiene que tener en cuenta desde las teorías de los colores hasta la interacción hombre-ordenador.
7. El comunicador es quien se ocupa de enderezar las cualidades puramente artísticas del diseñador hacia el cliente, hacia una parte más comercial. Se tiene que asegurar que diferentes tipos de personas entiendan el mensaje.
8. El Project Manager, es decir el gerente, se tiene que ocupar de que el proyecto no salga demasiado de las líneas, tiene que coordinar el equipo y asegurarse que se respeten los tiempos.

La visualización de información tiene, entre multitudes de fines, la misión de contar un hecho, simplificar, comparar, descubrir, etc., con el objetivo de transformar los datos en información, la información en conocimiento.

La visualización de la información favorece nuestra capacidad de asimilar informaciones con el fin de amplificar el conocimiento.

En los próximos párrafos nos vamos a centrar en tres aspectos de la visualización de la información: La infografía, la visualización de datos estática y la visualización de datos dinámica.



3.3.1 Infografía

Analizando la palabra “Infografía” en Google Trends se puede fácilmente ver cómo está subiendo su popularidad.

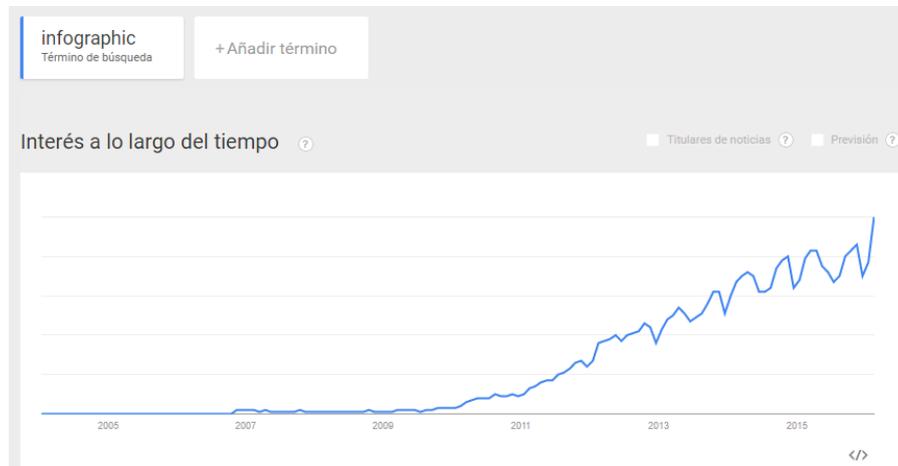


Ilustración 4 Búsqueda de la palabra "Infografía" en Google Trends, 30 de agosto de 2016, de <https://www.google.com/trends/explore?date=all&q=infographic>

A menudo cuando se piensa en Infografía se asocia a los últimos tiempos, a las redes sociales y a internet. La verdad es que todo empezó alrededor de mediados del siglo XIX, en Londres, donde durante una gran epidemia del cólera un médico llamado John Snow elaboró lo que se suele llamar como “el mejor infográfico de la historia”. [6]

Pero de eso se hablará más adelante, en el capítulo 3.4.1.

La infografía es un concepto que es muy fácil de hacerse viral. Esto está demostrado por la gran cantidad de infografía que se pueden ver en las grandes redes sociales como Facebook, Twitter o últimamente Pinterest.

Según la Wikipedia es una representación visual de información “en la que intervienen diversos tipos de gráficos y signos no lingüísticos y lingüísticos (pictogramas, ideogramas y logogramas) formando descripciones, secuencias narrativas o argumentativas e interpretaciones, presentadas de manera gráfica normalmente figurativa, que pueden o no coincidir con grafismos abstractos y/o sonidos”. [7].

Una infografía puede comunicar al usuario mucha información en muy poco tiempo y sobre todo de manera clara y amena. Esto se verifica porque detrás de una infografía, que la mayoría de los internautas podrán ver como un simple dibujo o conjunto de grafos, hace falta mucho estudio pero sobre todo mezclar información y conocimiento de ramas diferentes.

Llegados a este punto, se necesita plantearse un par de preguntas: ¿Merece la pena todo este esfuerzo para realizar una infografía? Y abstrayendo aún más, ¿Por qué usar infografías?

Podemos resumir las motivaciones en la siguiente lista, aunque en cada campo se pueden obtener motivaciones diferentes:

- Para transmitir mensajes
- Para mostrar muchísima información de forma compacta y fácil de entender.
- Para describir causas y efectos y saber por qué suceden determinados acontecimientos.
- Porque nos enseñan la evolución temporal de algo según determinados parámetros.
- Para mostrar mucha información utilizando recursos gráficos para no saturar una imagen.

3.3.1.1 Tipos de Infografías

Existen muchísimos tipos de infografías, y cada una de ellas se dirige a un público objetivo, con el fin de comunicarle un mensaje a través de un contexto que el conozca.

Lo más utilizados son los siguientes:

- La infografía estadística

El fin de esta infografía es comunicar información compleja y densa de manera rápida y clara.

Puede contener gráficas, tablas, diagramas que pueden resumir mucha información estadística. Normalmente en esta infografía se utilizan gráficas típicamente estadística que el público ya ha visto alguna vez, como por ejemplo gráficas de tortas, barras, anillos, puntos de dispersión etc.

Este tipo de infografía es muy utilizado por los diarios en la actualidad e incluso noticieros.

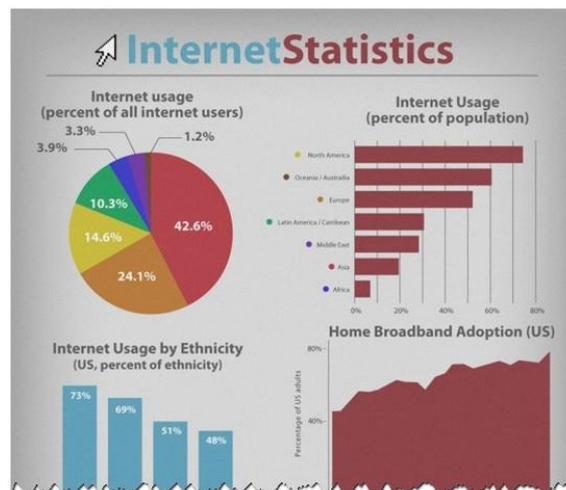


Ilustración 5 Ejemplo de infografía estadística, 30 de agosto de 2016, de <http://en.inmoreau.com/wp-content/uploads/2011/05/Internet-Statistics-tn.jpg>

- La infografía temporal o de línea de tiempo

Utilizada para representar la información y eventos que han ocurrido en una etapa en una franja de tiempo.

Es una representación de una secuencia cronológica a través de una línea de tiempo que ayuda a quién observa a realizar asociaciones temporales de manera muy simple e inmediata.

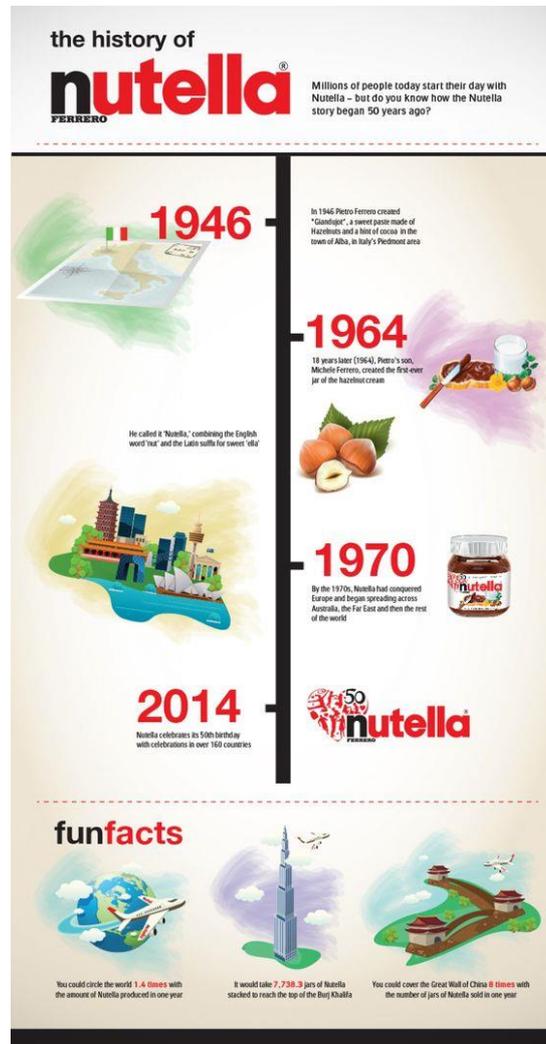


Ilustración 6 Ejemplo de infografía temporal sobre la historia de la Nutella, 30 de agosto de 2016, de <https://s-media-cache-ak0.pinimg.com/564x/12/D3/37/12D337e0fb21a1bfe18bf35e21ecefad.jpg>

- Infografías de Procesos de Producción

Este tipo de infografías son usualmente utilizadas para enseñar un proceso de elaboración de un producto. Últimamente se puede ver muchos en las revistas y página web de cocina, ya que son muy impactantes y muy útiles para mostrar recetas utilizando gráficos. Por ejemplo, buscando en la red social Pinterest [8] las palabras “#Cook #Infography” se puede ver como salen miles de imágenes con todo tipo de recetas.

Usar imágenes que se relacionan con los estados del producto puede ser un gran ejemplo del entendimiento del proceso de elaboración.

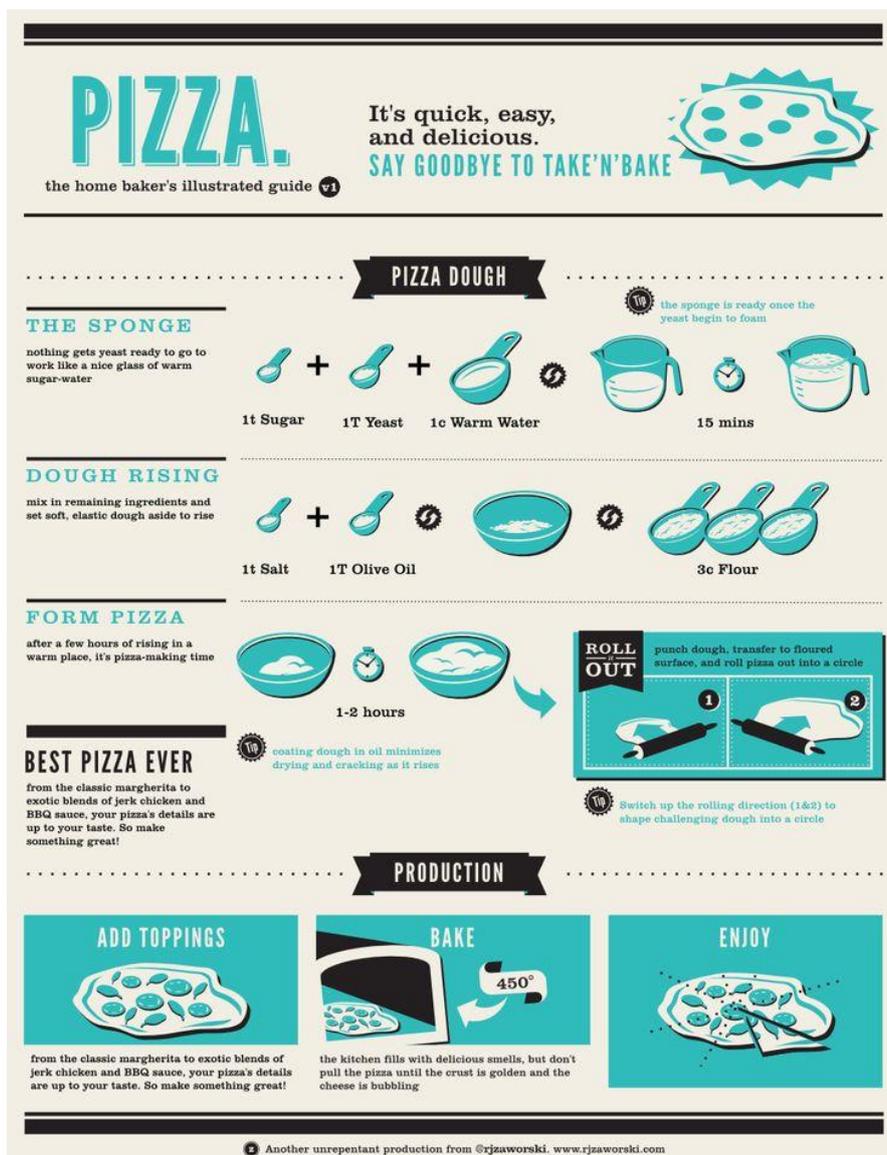


Ilustración 7 Ejemplo de infografía con la receta de la Pizza, 30 de agosto de 2016, de <http://encyclopizza.com/wp-content/uploads/2014/09/e3c0342a287d7e7afb8a9499d505905e.jpg>

- Infografías Geográficas basada en la Localización

Esta infografía se puede encontrar en cualquier lugar, desde mapas en las escuelas hasta complejos gráficos astronómicos, ya que puede ser entendida por muchos públicos.

Los mapas de ciudad y países pueden ser considerados un ejemplo de esto.

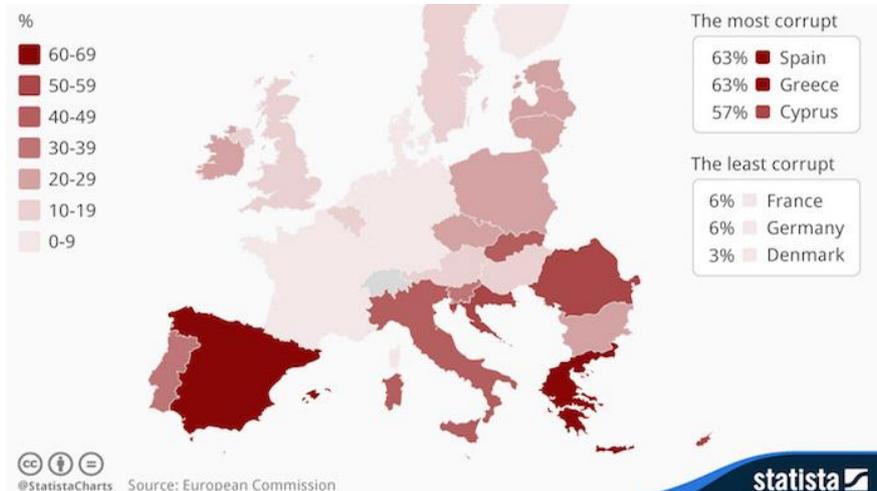


Ilustración 8 Ejemplo de infografía basada en la localización que representa los país más corruptos de Europa en 2015, 30 de agosto de 2016, de <http://www.voxeurop.eu/files/images/brief/corruption.jpg>

- Infografía Biográfica

Este tipo de infografía sirve para detallar la vida o las obras de un personaje conocido. En esta se detallan los principales acontecimientos y logros de un personaje.



Ilustración 9 Infografía biográfica de Mina Mazzini, una de las cantantes italianas más conocida, 30 de agosto de 2016, de <http://www.estaciondiseno.es/portfolio/infografia-biografica/#>

3.3.2 Visualización de datos estática

David Mccandless asegura en su célebre libro que “La información es bella” [9], sin embargo en realidad los datos en si son muy poco artísticos y tenemos que trabajar con ellos para darles una forma.

Cuando nos encontramos con datos en el mundo real, necesitan muchísimo tiempo para transformarlos, limpiarlos y adquirirlos. Es verdad que hoy en día existen varias herramientas automáticas que realizan esto proceso, pero en la mayoría se necesita también bastante tiempo de trabajo manual.

Cuando se piensa en la visualización de datos, la mayoría de la gente la asocia a los típicos gráficos de barra o de tarta. Estos gráficos sí que pueden dar una idea de la información que se quiere representar, pero del punto de vista artístico y comercial ni son bonitos ni captan la atención del cliente a primera vista.

Por ejemplo, comparados con los gráficos de barras, los gráficos de burbujas representan más puntos de datos en menos espacio, los gráficos de anillos indican claramente las relaciones entre elementos y conjuntos y los diagramas de árbol muestran de manera muy clara las categorías jerárquicas.

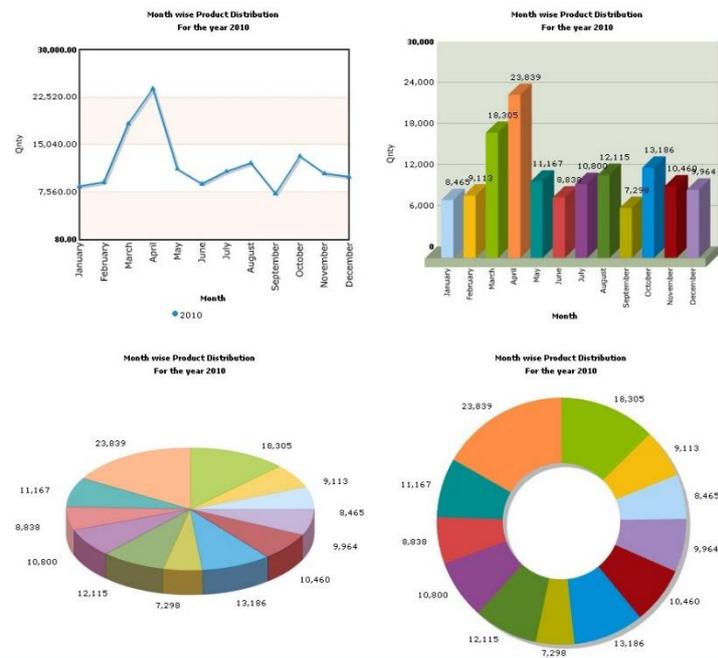


Ilustración 10 Ejemplos de los gráficos más conocidos, 30 de agosto de 2016, de http://2.bp.blogspot.com/-OpnOZSbhYPA/VCI3iLP6IKI/AAAAAAAAAD8/y4LwZwDGw_8/s1600/FusionChartGraph.jpg

Hay muchas corrientes de pensamiento sobre las gráficas. Mucha gente sigue pensando que ninguno de los gráficos que acabamos de ver es más efectivo que un simple gráfico de barras para una comparación entre valores.

Además, el hecho que dispongamos de herramientas que nos permiten fácilmente crear formas de visualización visualmente llamativas, no quiere decir que sean la más efectiva.

Es muy fácil caer en pensar que la parte estética es más importante que la práctica, que el “Eye Candy” es más importante de la claridad y la eficaz.



Sin embargo, en un mundo como el actual donde el usuario medio dispone de mucha información y le dedica muy poco tiempo, hace falta destacar, encontrar la manera de generar gráfica clara pero al mismo tiempo atractivas, que impresionen al público pero que transmitan el mensaje.

Estamos hablando de visualización de datos estática. Es decir, un gráfico que recopila un estudio hecho disponiendo de una cantidad de datos. Este gráfico es ideal para representar datos almacenados, y que no va a subir cambios ni van a actualizarse.

En el párrafo anterior nos hemos enfocado sobre la infografía, y es verdad que los dos conceptos, infografía y visualización estática, se parecen bastante. Sin embargo cabe subrayar las pequeñas diferencias:

Según Brian Shaler [10], visualizador y evangelista de datos, “una Infografía es la mera representación de datos o de un concepto a través sobre todo de objetos o símbolos. A menudo se enseñan sus partes con la mano, aunque el mecanismo de dibujo utilizado no tiene impacto en la categorización. Por el contrario, en la visualización de datos las reglas concretan cómo las variaciones y relaciones dentro de un conjunto de datos se convierten en características y objetos visuales. La Visualización de datos permite que los datos sean los protagonistas, y a menudo están generados y catalogados por ordenadores”.

Otra diferencia importante es la interacción con el usuario.

Una infografía es muy artística y posee muchas informaciones y objetos, pero sigue siendo plana y parada.



Ilustración 11 Parte de una infografía que da 5 razones según las cuales es mejor ser freelance, 30 de agosto de 2016, de <https://s-media-cache-ak0.pinimg.com/564x/f4/9d/69/f49d691d7d45d0b8ccd2a0628a7be587.jpg>

Por el contrario, una visualización de datos puede contener varias funciones y varias herramientas para poder “jugar” con los datos, por ejemplo controladores que permite aplicar filtro a los datos y focalizarse en lo que más importa. Esta última parte la analizaremos a fondo en el siguiente párrafo.

Las infografías se han vuelto virales, son un material 'atractivo', llaman la atención y son para una audiencia masiva. Sin embargo, una visualización de datos incorpora más elementos estadísticos e informáticos, más datos en sí y se pueden sacar muchas más conclusiones.

Además, para extrapolar los datos, se pueden utilizar múltiples fuentes de información, como pueden ser por ejemplo sensores, Smartphones, GPS o monitores médicos.

En una visualización de datos estática se suelen utilizar más números, se manejan más tipos de datos y la parte artística no es la protagonista, pues se intenta que el actor principal sea la información en sí.

En las infografías se utilizan muchas imágenes creadas a través de los datos, se visualizan menos datos en si centrándose directamente en las conclusiones. Se da mucha importancia a la parte artística y creativa.



3.3.3 Visualización de datos dinámica / interactiva

Como se ha analizado en los párrafos anteriores, vivimos en una época donde circula una cantidad exorbitante de información, y la única manera de poder traer beneficio de ella es saberla enseñar correctamente al público. En detalle, se ha visto que es una infografía y de que se trata cuando se realiza una visualización de datos estática.

Una de las principales fuentes de información de hoy en día son las redes sociales. Todo el mundo genera información desde varios dispositivos, como ordenadores, Smartphone, Smart tv y últimamente smartwatch.

Es interesante observar que cuando se habla de información ya no se entiende solo texto y foto generados por el usuario, también se incluye posición del mismo, y estadística calculadas por el dispositivo.



Ilustración 12 Popularidad de Facebook en todo el mundo según su uso, 30 de agosto de 2016, de <http://www.keywebmetrics.com/2013/07/big-data-visualizations/>

En la imagen se puede ver el uso de Facebook en todo el mundo. La intensidad de la luz indica la popularidad de Facebook.

Como se puede notar de esta imagen, el mensaje y los datos que se quieren transmitir son muy claros e impactante, el usuario es fascinado de la imagen, mucho más que si se hubiese informado de esta estadística mediante tablas de datos.

Por esta razón, ya se necesita utilizar otra manera de representar la información, ya que esa puede cambiar en intervalos de tiempos muy breve (a veces minutos o segundos) y una imagen estática publicada hoy en una página web, en menos de una semana puede resultar obsoleta.

La alternativa para visualizar este tipo de datos es la visualización de datos dinámica o interactiva.

Normalmente se utilizan pantallas con varios indicadores, para poder “navegar” entre los datos y centrarse en lo que interesa, aplicándole filtros.

Estas pantallas se llaman dashboards, y últimamente se pueden encontrar en muchas páginas web, periódicos digitales etc.

Visitors - Map Distribution



U.S Visitors and Income By State (source)

| State | Stock Size | Satisfaction Index | Visitors | Income (Sum) | \$ / visitor |
|----------------|------------|--------------------|-----------|----------------|--------------|
| New York | | A+ | 3,313,508 | \$2,071,008.60 | \$0.625 |
| California | | A+ | 2,008,366 | \$1,824,459.42 | \$0.908 |
| Pennsylvania | | A+ | 503,698 | \$424,459.38 | \$0.842 |
| Vermont | | A | 318,332 | \$324,459.02 | \$1.019 |
| West Virginia | | A | 218,560 | \$202,327.00 | \$0.925 |
| Massachusetts | | B+ | 202,271 | \$58,035.27 | \$0.286 |
| Oklahoma | | B+ | 102,503 | \$82,539.38 | \$0.805 |
| Maine | | B | 98,689 | \$63,935.80 | \$0.647 |
| Alabama | | B- | 54,201 | \$62,502.02 | \$1.153 |
| North Carolina | | B- | 46,185 | \$53,297.32 | \$0.647 |



Income Per Visitor



Total Income



Ilustración 13 Ejemplo de una dashboards, 30 de agosto de 2016, de <http://plotting-success.softwareadvice.com/files/2013/08/SiSense-interactive-dashboard.jpg>

Es el extremo de la visualización de datos. Tal vez se utilizan datos muy grandes, que es imposible visualizar (y se entendidos) sin poderle aplicarle filtro.

Otras veces, y desde mi punto de vista es la parte más futurista, los datos se generan en tiempo real.

3.4 Historia moderna de la visualización de información

Para poder entender mejor la compleja realidad de la visualización de la información, es necesario conocer la historia de la misma. En nuestro caso vamos a centrarnos sobre todo en las últimas décadas, analizando cómo han evolucionado las manifestaciones gráficas en relación con el desarrollo tecnológico.

3.4.1 John Snow y la mejor infografía de la historia

John Snow era un médico británico que vivió en Londres en mediados del siglo XIX. Durante esta época Londres era una ciudad caótica, sucia y desordenada. Además, eran muy frecuentes las epidemias. [6]

Una de las infecciones que fue más terrible era el cólera. La mayoría de los médicos pensaba que el medio de propagación del virus eran las vías respiratorias, es decir que el virus estaba en el aire.

Sin embargo, había otra corriente que afirmaba que el cólera entraba por el sistema digestivo, es decir que el agua de Londres estaba contaminada, y se moría por beber agua y comer alimentos, no por respirar aire infectada.

Uno de los médicos más destacados de esta corriente era John Snow.

En septiembre de 1854 un agresivo brote de cólera causó la muerte de casi mil personas en una pequeña parte del barrio de Soho, cerca de donde vivía Snow. John Snow entonces convenció las autoridades a cerrar la bomba del agua en esta zona, y el brote cesó de inmediato.

Para demostrar esta tesis, Snow compró un mapa de Londres, que se puede ver en la siguiente imagen, y registró todos los fallecidos basándose en información de hospitales cercanos.

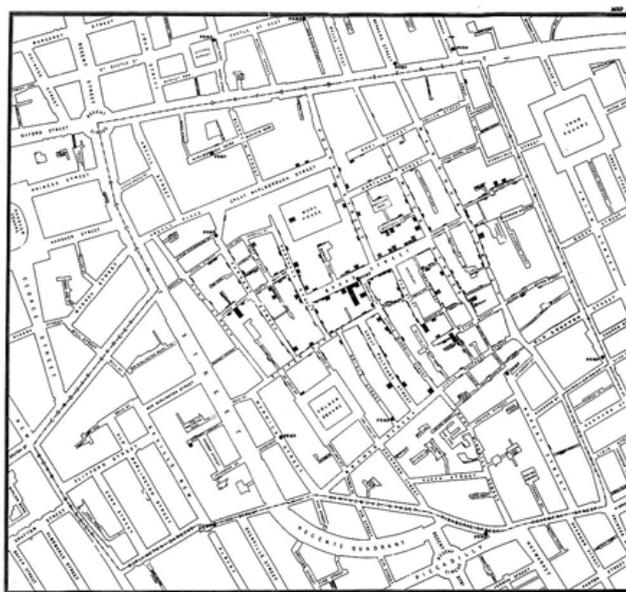


Ilustración 14 Mapa donde Snow indicó los fallecidos en la epidemia de cólera de 1854, 30 de agosto de 2016, de <http://www.yorokobu.es/wp-content/uploads/colera.gif>

En el gráfico Snow trazó las muertes con líneas finas, y se puede ver que la mayoría de ellas eran cerca de los pozos de agua contaminados (PUMP), como se puede ver en la siguiente imagen.

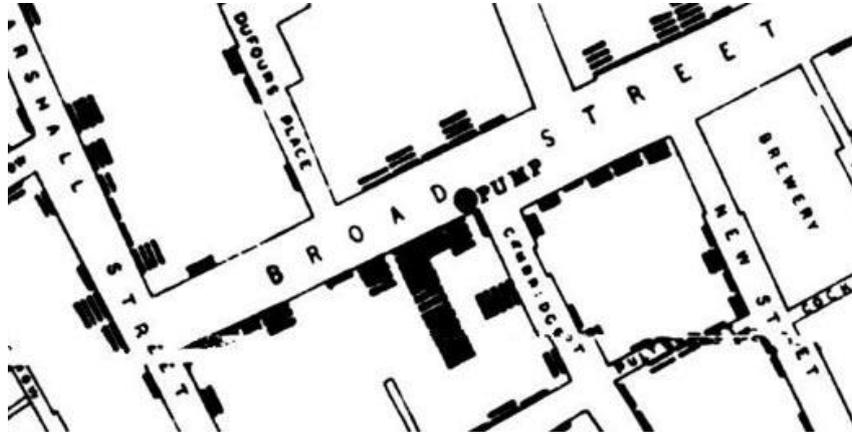


Ilustración 15 Zoom del mapa donde Snow indicó los fallecidos en la epidemia de cólera de 1854, 30 de agosto de 2016, de <http://blogs.elpais.com/.a/6a00d8341bfb1653ef01538dd8656f970b-pi>

3.4.2 Comienzos del 1900, el método gráfico a sus albores

Es obligatorio subrayar que en las décadas entre los años 1860 y 1890 se consolidó el método gráfico, ya que todos los artistas capaces de crear representaciones visuales a partir de la información empezaron a experimentar con esta nueva técnica. Por esta razón este periodo de tiempo podría denominarse “Los años dorados del método gráfico”.

El método gráfico fue oficialmente reconocido por las agencias gubernamentales y aceptado en las publicaciones como un elemento más de información.

Posteriormente a esta época, en las primeras décadas del siglo XX, hubo una época de grandes cambios en la vida social, política, cultural y económica de nuestra sociedad. Uno de los más importantes fue la primera guerra mundial.

Estos cambios alteraron varios aspectos de nuestras vidas, y en particular en la representación visual de datos hubo un cambio radical. El inicial entusiasmo por la representación visual se transformó en una producción masiva de todo tipo de gráficas. La estética se volvió mucho más importante y se llegó a descuidar el contenido.

Esta etapa pudo definirse como un retroceso en la evolución de la visualización de datos, ya que se perdieron muchos valores y técnicas aprendidas en las épocas anteriores.

3.4.3 La abstracción de la realidad

En los años 30, gracias a un simple mapa, el plano del metro de Londres, hubo un cambio clave en la historia de la visualización de datos.

Hasta entonces, todos los mapas, a pesar de las diferentes tipología gráficas, aún conservaban las mismas características, por ejemplo la de representar en detalles los recorridos de los medios de transporte. Era muy importante representar si un recorrido tenía varias curvas o iba en una colina o en plano.

Henry C. Beck, el creador de este revolucionario mapa, introdujo una nueva forma de representación de la realidad: la esquematización de un mapa.



Ilustración 16 Mapa del plano del metro de Londres de los años 30, 30 de agosto de 2016, de <https://elpasteldesagan.files.wordpress.com/2015/01/1933-underground-map.jpg>

Gracias a él, la representación gráfica de las líneas de transporte comenzó a favorecer la comprensión y la usabilidad, dejando de un lado las características geográficas. Los rasgos geográficos desaparecen, y se nota un nivel mayor de abstracción de la realidad. Después de este evento, el concepto de abstracción empezó a aplicarse a distintos ámbitos.

3.4.4 Avances tecnológicos: la informática

Durante el siglo XX el desarrollo de las nuevas tecnologías y la abstracción - que favoreció el progreso de nuevas ideas - beneficiaron el desarrollo de los gráficos multivariados, ya que no solo se representaba en 2D sino también se podía generar gráfica en 3D.

Gracias a estas técnicas fue posible crear nuevos tipos de gráfica y analizar la información desde más puntos de vista. Cuanto más avanza la informática, y más se van desarrollando gráfica a alta resolución y nuevas formas de gráficas.



También fue de ayuda la publicación de John W. Turkey “El futuro de las análisis de información”, que reconocía la análisis de la información como una rama de la estadística, diferente a la estadística matemática.

Esta publicación, junto con otras más, logró que la gráfica de datos volvió a ser respetada.

3.5 Los gurús de la visualización de la información

3.5.1 Edward Tufte

Edward Rolf Tufte [11] (1942, Kansas City, Misuri, EE. UU.), es un estadístico y un artista, profesor emérito en Ciencias Políticas, estadística e informática en la Universidad de Yale. Además, dictó cursos sobre evidencia estadística y diseño de información y de interfaces.

Él escribió, diseñó, y auto publicó cuatro libros sobre la visualización de datos, que son considerados unos clásicos.

Es defensor del minimalismo referente a la representación gráfica de datos y de la eliminación de todo tipo de objeto que entorpezca su comprensión. Su estilo premia la información sobre el ornamento.

Su notoriedad en el ambiente de la visualización de datos es muy alta, ya que muchas citas clave de este ambiente son suyas. Por ejemplo, "El mundo es complejo, dinámico, multidimensional, el papel es estático y plano. ¿Cómo vamos a representar la rica experiencia visual del mundo en la mera planicie", es una de las citas más famosas de Edward Tufte [12].

Otra cita muy conocida es: "Power Corrupts. PowerPoint Corrupts Absolutely" [13] (el poder corrompe, PowerPoint corrompe absolutamente), es un crítico a la presentaciones creadas con el software Microsoft Powerpoint en los ámbitos técnicos, científicos y empresariales ya que considera que es ineficiente para transmitir información.

Es interesante saber que creado términos como "chartjunk" [14], que se refiere aquellos elementos de una gráfica que no son esenciales y dificultan su lectura, y el "ratio tinta-información", que mide la cantidad de tinta que se aconseja utilizar para representar efectivamente un conjunto de datos. Más en detalle, Edward Tufte define el ratio en su libro "The Visual Display of Quantitative Information [15]" de esta manera:

"Datos-Tinta: La mayor parte de la tinta en un gráfico debería representar la información de los datos, la tinta cambia como cambian los datos. Datos-Tinta es el núcleo no borrable de un gráfico, la tinta no redundante colocada en respuesta a la variación en los números representados. Entonces:"

$$\text{La razón Datos-Tinta} = \frac{\text{Datos-Tinta}}{\text{Total de tinta usada en imprimir el gráfico}}$$

Edward Tufte señala que: "debemos de maximizar la razón Datos-Tinta, dentro de lo razonable. Igualmente afirma que debe borrarse la tinta de no-datos dentro de lo razonable."

Tufte también creó el concepto de Sparklines [16], que son pequeños gráficos para mostrar tendencias y variaciones de variables. Estos pueden ser utilizados por ejemplo para representar acciones de bolsa, temperaturas, precipitaciones, etc.



Los sparklines tienen la característica de ocupar muy poco espacio e ir acompañados de texto.

| category | chart | jan | feb | mar | apr | mai | jun | jul | aug | sep | oct | nov | dec |
|----------|---|------|------|------|------|------|------|------|------|------|------|------|------|
| revenue |  | 1119 | 1139 | 1152 | 1150 | 1146 | 1236 | 1314 | 1336 | 1312 | 1267 | 1226 | 1294 |
| licenses |  | 463 | 472 | 483 | 488 | 490 | 534 | 557 | 563 | 552 | 542 | 526 | 482 |
| service |  | 463 | 472 | 483 | 488 | 490 | 534 | 557 | 563 | 552 | 542 | 526 | 631 |
| reseller |  | 193 | 195 | 186 | 174 | 166 | 168 | 200 | 210 | 208 | 183 | 174 | 180 |

Ilustración 17 Ejemplo de gráfico sparklines de Tufte, 30 de agosto de 2016, de <https://en.wikipedia.org/wiki/Sparkline>

3.5.1.1 El ejemplo del accidente del transbordador espacial Challenger

En este apartado se contará un ejemplo donde se puede ver como la visualización de información de vez en cuando puede tener una importancia fundamental.

El 28 de enero de 1986 el transbordador espacial Challenger, durante la misión espacial de la NASA número STS-51-L, se desintegró 73 segundos después del lanzamiento, provocando la muerte de los siete miembros de la tripulación. La nave se desintegró sobre el océano Atlántico, cerca de la costa del centro de Florida (Estados Unidos) [17].

Esto ha sido calificado como el accidente más grave en la historia de los lanzamientos espaciales.

La causa de la desintegración del shuttle fue que una junta tórica del cohete acelerador sólido (SRB) derecho fallara durante el despegue. Este fallo de la junta tórica causó la apertura de una brecha, permitiendo que se desperdiese hacia el exterior el gas caliente presurizado contenido en el interior del motor del cohete sólido y contactara con la estructura adyacente de conexión con el SRB y con el tanque externo de combustible.

Esto provocó la separación de la conexión posterior del SRB derecho y causó el fallo estructural del depósito externo. La nave espacial fue destruida rápidamente por las fuerzas aerodinámicas.

La causa principal de la ruptura de la junta tórica fue que la noche anterior hubo temperaturas muy bajas. La empresa constructora de la junta, la Thiokol, había avisado a los ingenieros de la NASA que con bajas temperatura podía aumentar el riesgo de fallos. El día anterior por la tarde en toda prisa los ingenieros de la Thiokol redactaron un informe con estas informaciones, donde salía que en varios test se produjeron fallos en las juntas cuando había bajas temperaturas.

Sin embargo, los datos fueron representados en este informe de manera muy desordenada y no fácilmente interpretable, y los ingenieros de la NASA decidieron seguir adelante con el proyecto.

Este fue un ejemplo de como la visualización de la información tiene un papel muy importante en la toma de decisiones.

Por esto, el diseñador de información Edward Tufte utilizó el accidente del Challenger como un ejemplo de los problemas que puede causar la falta de claridad a la hora de representar información [18].



De hecho, según Tufte, si los ingenieros de Morton Thiokol hubiesen presentado más claramente los datos sobre la relación entre las temperaturas frías y los fallos de las juntas de los cohetes aceleradores sólidos, quizá habrían logrado convencer a los administradores de la NASA de posponer el lanzamiento.

Uno de los gráficos que le anduvieron es el siguiente:

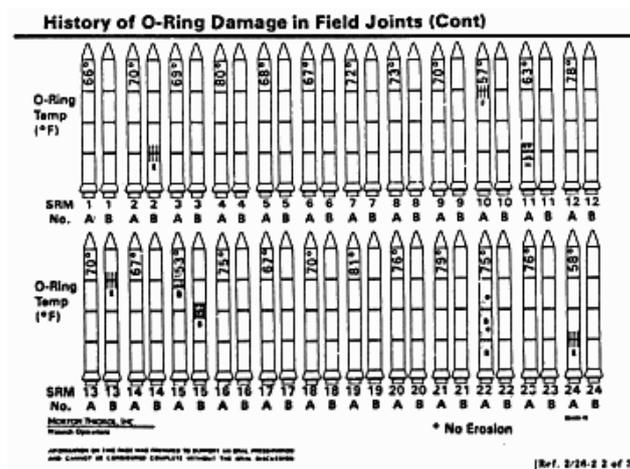


Ilustración 18 Parte del informe que los ingenieros de la Thiokol enviaron a la NASA, 30 de agosto de 2016, de <http://college.cengage.com/history/us/resources/shared/primary/source/images/challenger2.gif>

En la imagen se pueden ver dibujos de los cohetes aceleradores sólidos, en orden cronológico, con escritas las temperaturas y los daños en las juntas.

Tufte argumentó que este gráfico no representa de manera clara e inmediata la información y que cuando los ingenieros y directores terminaron analizando el gráfico, no sabían más de lo que tenían antes.

Entonces Tufte realizó un gráfico con la misma información, pero representada de manera mucho más clara, ya que se puede entender fácilmente.

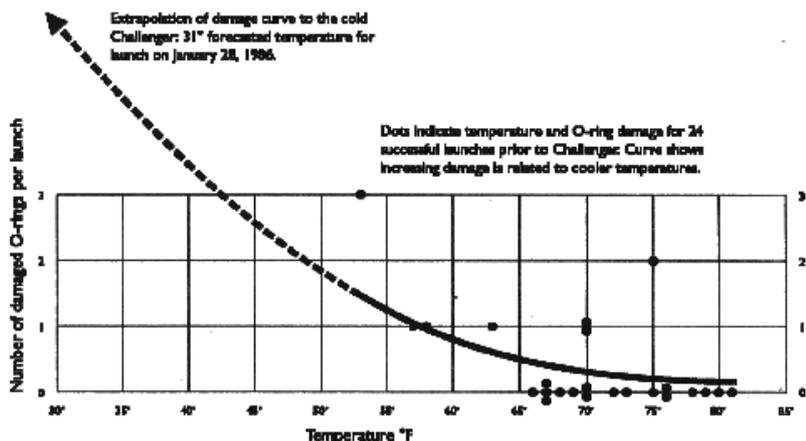


Ilustración 19 Gráfico elaborado por Tufte donde se representa de manera más clara la probabilidad de ruptura de los juntos dependiendo de las temperaturas, 30 de agosto de 2016, de <http://www.websimplicity.co.za/i/e/Edward-Tufte-Challenger-scatterplot-chart-3.jpg>

En el eje horizontal se pueden ver las temperaturas y en el vertical el número de rupturas en las juntas tóricas. Se puede rápidamente notar que cuando más baja es la temperatura, más daños en las juntas ha habido durante los experimentos.

Como conclusión, Tufte argumentó que la mala presentación de información podría haber afectado las decisiones de la NASA durante el último vuelo del Columbia y podría ser una de las responsables de la muerte de los siete miembros de la tripulación.

3.5.2 Stephen Few

Otro gurú de la visualización de datos es Stephen Few [19].

Tiene más de 30 años de experiencia como innovador, consultor y educador en los campos de business intelligence e visualización de la información, es un líder en el aspecto de la visualización de datos que más se acerca a la comunicación con sentido.

Su obra más conocida es “Show me the numbers”, que se considera como una guía clave sobre los gráficos comerciales. [20]

3.5.3 Tufte, el artista y Few, el ingeniero

Las obras y los pensamientos de estos dos grandes de la visualización de datos pueden parecer a su vez muy parecidos u opuestos.

Ellos dos normalmente se citan casi siempre juntos cuando se habla visualización de datos, como si fuesen una individualidad. Los dos tienen una visión similar sobre los principios básicos de la visualización de datos y parecen estar en la misma onda.

Sin embargo, mirando más a fondo, se puede notar una gran diferencia entre los dos.



3.6 Empresas más prometedoras en la Visualización de Datos

En la época actual, los datos estructurados y no estructurados están teniendo un crecimiento exponencial. Es la era del Big Data, estamos sumergidos de datos de todos los tipos y la mayoría de las veces no sabemos cómo manejarlos y sacar conclusiones.

Es muy difícil analizarlos de manera eficiente, identificar solo la información relevante y aplicar filtros para eliminar la información sobrante. Además, si se quiere enseñar esta información a los clientes, se tiene que trabajar también en la visualización de datos, ya que estamos hablando de muchísima información que se tiene que resumir y dibujarla con un diseño entrañable.

Por internet se encuentran herramienta que prometen ocuparse de estas tareas, pero la mayoría de estas necesitan mucho estudio y mucho trabajo, y casi siempre no se llega a un resultado profesional.

Por esta razón es muy interesante conocer las empresas que se dedican únicamente a esto, de esta manera se puede observar cuales son los países del mundo más evolucionados en este aspecto.

A continuación, una lista de las 20 empresas más prometedoras según la revista CIO Review, una revista especializada en tecnologías de la información. [21]

Se procederá solo mencionarlas, diciendo su sede principal y a que se dedican en particular, y se enfocará la atención sobre la única empresa española.

| Company | Management | Description |
|---|--|---|
| AnyChart Park City, UT anychart.com | Anton Baranchuk Founder & CEO | Provides cross-platform data visualization solutions |
| Boost Labs, LLC Gaithersburg, MD boostlabs.com | Ali Allage CEO | Provides innovative full stack visualization services for private, public, and nonprofit industry clients |
| Cadeon Inc. Alberta, Canada cadeon.com | Phil Unger CEO | Provides solution consulting and system integration in visualization, data virtualization, content analytics, agile data warehousing, and enterprise information architecture |
| Crunch Data, Inc. Thousand Oaks, CA crunch-data.com | Douglas Textor CEO | Provides solutions for data visualization to ensure BI readiness |
| datapine GmbH Berlin, Germany datapine.com | Martin Blumenau & Jakob Rehermann Managing Partners | Develops highly scalable SaaS tool that revolutionizes the way of performing database analysis by making even complex functionalities available to non-technical persons |
| Diamond Visionics Vestal, NY dvcsim.com | David Gdovin Co-Founder & President | Provides cutting-edge image-generation software and visualization tools for training and simulation, mission rehearsal, and homeland security |
| Enthought, Inc. Austin, TX enthought.com | William Cowan CEO | Provides easy installation of the core scientific analytic and scientific Python packages, creating a robust platform to explore, develop, and visualize |



| | | |
|--|---|--|
| ERNESTO OLIVARES VISUAL INFORMATION Cantabria, Spain ernestoolivares.es | Ernesto Olivares CEO | Customizing infographics tool to provide the best visual solutions. |
| FTI Technology Washington, DC ftitechnology.com | JR Jenkins Sr. Director, Product Marketing | A global provider of e-discovery and information governance software, services and consulting |
| Genialis Houston, TX genialis.com | Nejc Škoberne CEO | Builds end-to-end bioinformatics software solutions by combining Biomedical data management, Customizable data analysis pipelines and Interactive visualizations |
| i.Predictus LLC Parsippany, NJ ipredictus.com | Monica C. Smith CEO | A provider of automated data ingestion and visualization solutions to predict media performance and drive ROI |
| ICIMO LLC Apex, NC icimo.com | Bryce Gartner CEO | Transforms businesses into data-driven organizations by building profitable relationships between users and their data |
| Interactive Network Technologies, Inc Houston, TX int.com | Olivier Lhemann President | Specializes in the development of graphics toolkits and graphics intensive applications |
| Interstates Control Systems, Inc Sioux Center, IA interstates.com | Jack Woelber, President Stephen Dekker, Division Manager of Business Intelligence | Interstates helps to collect transactional data through historical records and other data collection tools and transforms them into ready-to-use visual forms |
| iVEDiX Pittsford, NY ivedix.com | Rajesh Kutty President & CEO | Offering a social media integrated application to solve complex data challenges, visually. |

| | | |
|---|-------------------------------------|--|
| Pellucid Analytics Boulder, CO pellucid.com | Adrian Crockett Co-founder & CEO | Provides a content platform blended with technology and industry-leading expertise, design, and advisory services |
| Seabourne, Inc. Washington, DC seabourneinc.com | Mike Reich Founder & CEO | Provides data integration, consolidation, and visualization tools |
| Seed Scientific Brooklyn, NY seedscientific.com | Adam Bly CEO | Leverages big data to craft solutions, fuel innovation, and improve decision-making for clients around the world in the commercial, public, and social sectors |
| Vizonomy LLC Washington, DC vizonomy.com | Ricardo Saavedra Founder | A computer software company delivering product management, geospatial data analysis/ visualization and web application development services |
| Zebra BI Ljubljana, Slovenia zebra.bi | Andrej Lapajne CEO | Provides data visualization solutions as tools for Excel which simplifies a company's requirement for reporting and managing data |

Ilustración 20 lista de las 20 empresas más prometedoras según la revista CIO Review, 30 de agosto de 2016, de <http://avanxo.com/wp-content/uploads/2014/06/CIO-Review-2014-Most-Promising-CloudC-Companies-Avanxo.pdf>

Entre las 20, está presente también la empresa Ernesto Olivares Visual Information [22], con sede en Cantabria, en España.

Ernesto Olivares Visual Information es una empresa de diseño de infografías, que se ocupa de transmitir mensajes de forma clara, simple y atractiva.

Su misión, en los últimos 15 años, ha sido ayudar a las empresas y proporcionarles las herramientas para que puedan visualizar grandes cantidades de datos de manera clara y sencilla.



3.7 Los Principales tipos de Visualización de Datos actuales

Cada día se producen 2.5 trillones de bytes de datos, esto equivale al 90% de la información creada en el mundo en los últimos dos años. [23]

Mucha gente se pregunta de dónde puede salir esta multitud de datos. Bueno, prácticamente de todas partes; desde sensores, redes sociales, imágenes digitales y videos. Disponemos de muchos más datos de los que nos imaginamos.

Por esta razón, la tarea fundamental es organizarlos para que tengan un sentido y se pueda generar algo con ellos. Y ahora es cuando la visualización de datos entra en escena.

Algunos refieren el periodo que estamos viviendo y el que nos espera en un futuro próximo como la “Revolución industrial de los datos” [24], por esto se tendrá que conseguir que las visualizaciones de datos sean siempre más eficientes, innovadoras y creativas. Y claro, siempre muy fácil de entender.

Por eso, se ha decidido dedicar un apartado a las posibilidades de visualización de datos actuales, para hacer el punto de la situación de donde estamos situados e intentar a hipotetizar lo que nos espera en los próximos años.

Se han recolectado 9 tipos de visualización de datos que según mi opinión representan de forma más o menos clara las posibilidades que se tienen hoy en día. No se van a analizar tipos de datos diferentes, sino que es posible también en la misma infografía utilizar varios de ellos.

3.7.1 Infografía interactiva

Un tipo de infografía que está teniendo un éxito enorme en los últimos años es la infografía interactiva. En esta infografía se combinan grandes cantidades de datos en una sola página, ya que el usuario puede navegar entre ellos y, la mayoría de las veces, decidir en cual focalizarse.

Se suele utilizar el color para distinguir fácilmente las tendencias y el espectador en un solo vistazo entiende el mensaje que se quiere dar y es atraído ya que puede navegar en la misma infografía.

Un ejemplo de este tipo de infografía es “La rutina diaria de famosos creativos” [25]



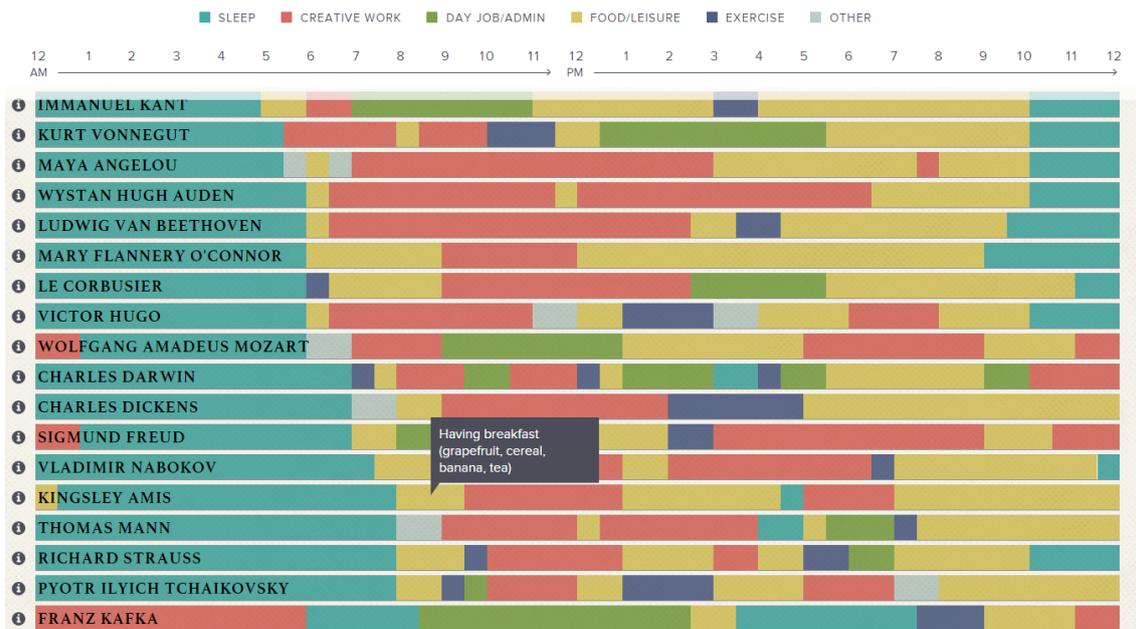


Ilustración 21 Ejemplo de infografía Interactiva, 30 de agosto de 2016, de <https://podio.com/site/creative-routines>

En esta infografía se puede ver la rutina diaria de varios famosos creativos. Las tareas se representan con colores y pasando encima sale un pequeño cuadro donde se pueden ver más detalles.

Combina varios colores, muchos datos en una sola página y sin duda entretiene mucho al usuario.

3.7.2 Infografía creada analizando tendencias

Las infografías creadas por ejemplo con comentarios de redes sociales (datos generados en Twitter por ejemplo) son muy de moda en la época actual.

Por ejemplo, la siguiente infografía (se visualizará solo una parte) enseña los argumentos principales de los americanos en 2014. Ha sido creada analizando 184.500.000 menciones de Twitter, y se ha conseguido ordenar, manejar y representar estos datos en una única página.

WEEKLY SHARE OF NEWS CONVERSATION BY STORY

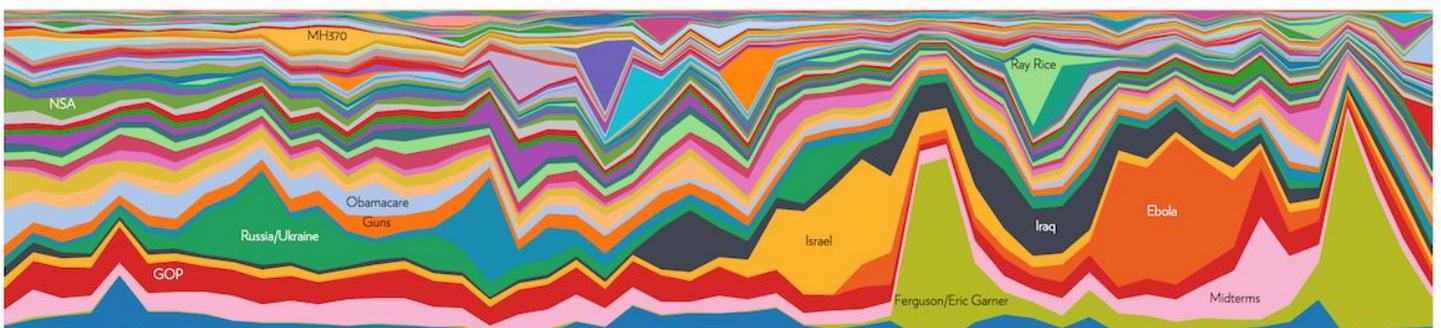


Ilustración 22 Ejemplo de infografía creada analizando tendencias, "The year in news", 30 de agosto de 2016, de <http://echeloninsights.com/wp-content/uploads/2014/12/theyearinnews20141.png>

3.7.3 Infografías animadas

Con la popularidad que están teniendo las GIFs últimamente, se han puesto de moda las infografías animadas. Generalmente no son muy grandes ya que gracias al movimiento pueden representar una cantidad impresionante de datos.

En la siguiente infografía se puede ver el porcentaje de americanos divididos por grupo de edad y sexo desde el 1950 hasta lo que se prevé ser en el 2060.

Siendo una GIF, se tomarán varias capturas de la infografía.

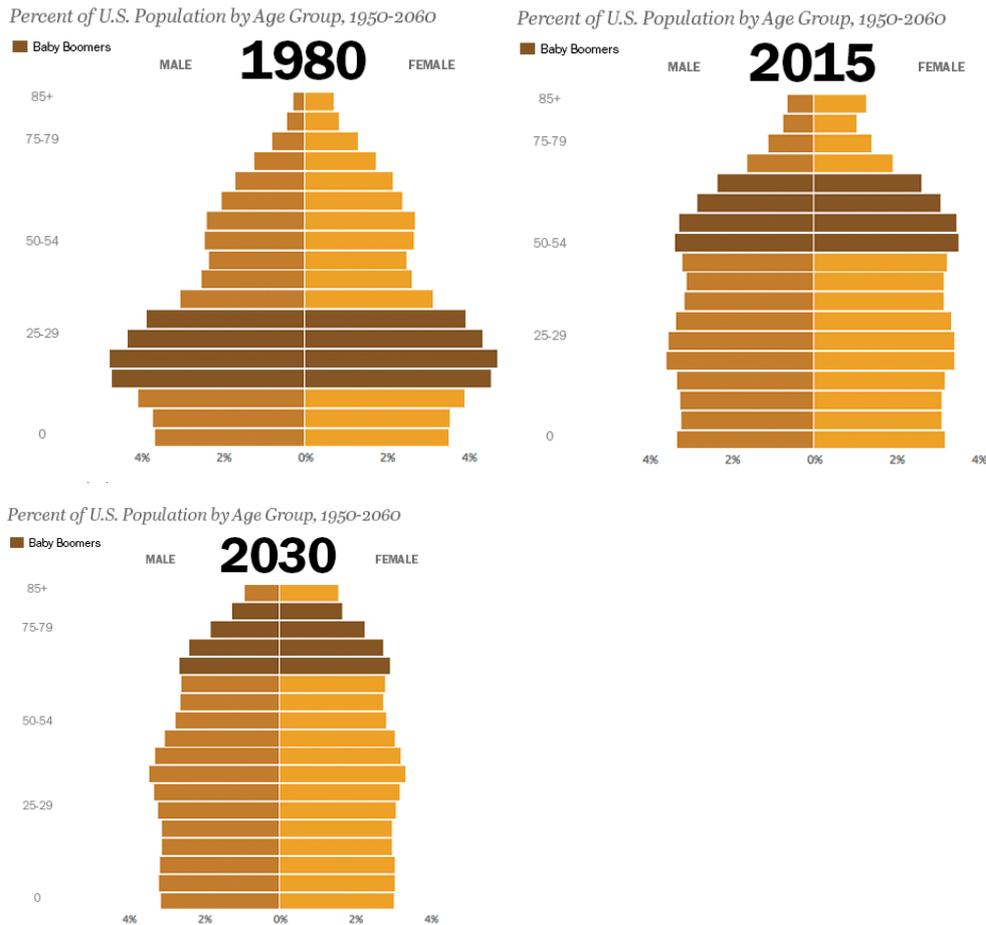


Ilustración 23 Ejemplos de infografía animada, 30 de agosto de 2016, de <http://www.pewresearch.org/files/2014/04/847889448.gif>

3.7.4 Infografía con imágenes reales

Esta infografía tiene como particularidad la utilización de imágenes reales per representar los datos, es decir en lugar de las gráficas comunes, imágenes reales que realicen su función.

A continuación se enseñarán varios ejemplos de gráficas creadas con imágenes reales. Esta infografía se está poniendo de moda sobre todo en los últimos tiempos ya que tiene gran impacto sobre todo en las redes sociales.





Ilustración 24 Ejemplos de infografía con imágenes reales, 30 de agosto de 2016, de <http://snip.ly/vF80#http://marion-luttenberger.squarespace.com/#/infographics/>

3.7.5 Infografías que utilizan metáforas

En esta categoría se analizan las infografías que tienen la particularidad de visualizar información sobre un argumento utilizando la parte gráfica de otro. Ese tipo de infografía es muy impactante. Por ejemplo, en el siguiente ejemplo, se representa un confronto entre dos marcas de cámaras de foto, Canon vs Nikon, y es representado como si fuese una lucha de boxeo.



Ilustración 25 Ejemplo de infografía que utiliza metáforas, 30 de agosto de 2016, de <https://sites.google.com/a/mtlsd.net/infographics-simple/services/metaphor-based-infographics>

3.7.6 Infografías que relacionan datos

En esta sección se van a analizar las infografías especialmente creadas para, además de visualizar mucha cantidad de datos, enfocarse sobre las relaciones que están entre ellos. En el siguiente ejemplo se pueden ver los principales géneros musicales de los últimos 100 años y, haciendo clic sobre uno, se pueden ver sus influencias.

Además, es posible escuchar un ejemplo de cada género.



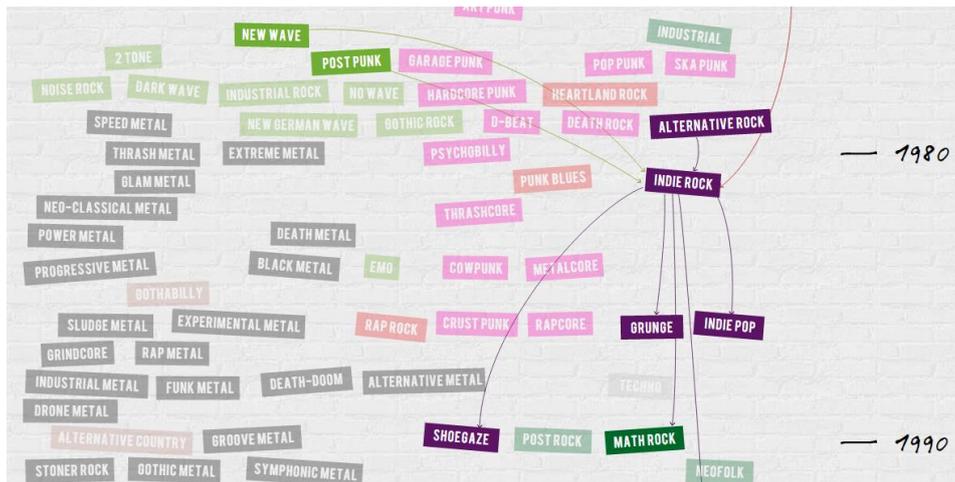


Ilustración 26 Ejemplo de infografía que relaciona datos, 30 de agosto de 2016, de <http://www.concerthotels.com/100-years-of-rock/>

3.7.7 Infografía con perspectiva variable

Este tipo de infografía normalmente representa mucha cantidad de datos que se tienen que representar de una forma definida, como por ejemplo líneas temporales muy largas o distancias.

En el ejemplo que sigue se sitúa el día actual en una línea del tiempo, viendo lo pequeño que es un día si es visto en un contexto muy amplio.

Ya que es una infografía animada, se van a representar unos fotogramas.

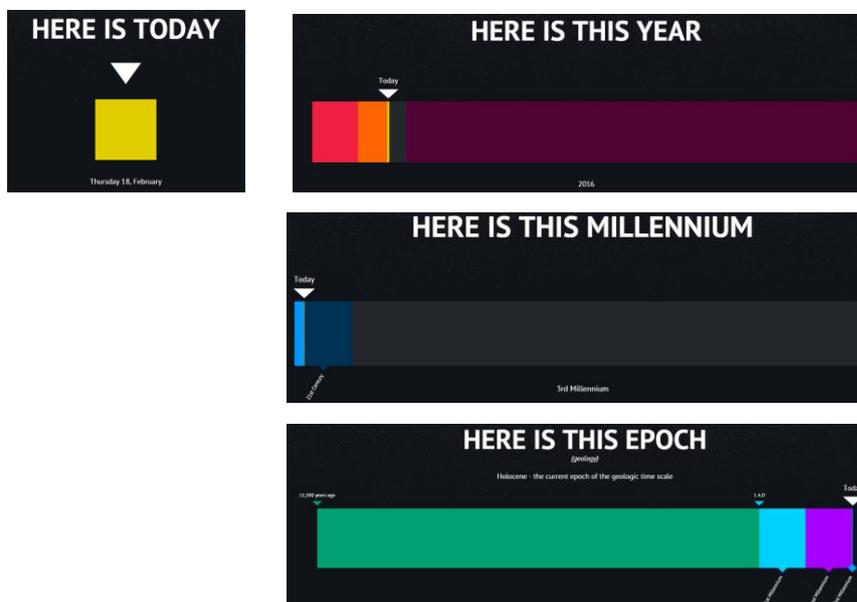


Ilustración 27 Ejemplo de infografía con perspectiva variable, 30 de agosto de 2016, de <http://hereistoday.com/>

3.7.8 Infografía que explican un proceso (o cuentan una historia)

Se trata de infografía, la mayoría de las veces animadas y con muchos elementos gráficos, que tienen su funcionalidad en explicar un proceso, una historia, un cambio de una comida etc. Es decir, el protagonista de esta infografía evoluciona durante el tiempo.

En la siguiente infografía animada se puede ver la vida de un grano de café, desde la plantación a la taza.

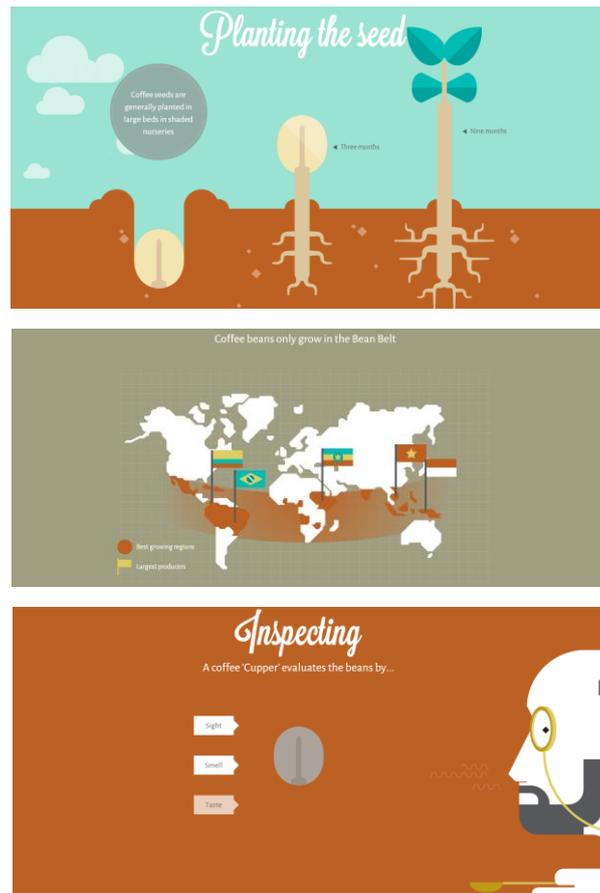


Ilustración 28 Ejemplo de infografía que explica un proceso, 30 de agosto de 2016, de <http://www.bizbrain.org/coffee/>

3.7.9 Infografías que proporcionan accesos a los “Raw data”

Este tipo de infografía permite al usuario de tener acceso también a los datos brutos originales. Además de esto, esta es una infografía normal, con sus colores, diseños y tal vez animaciones, así pues es el espectador el que elige si quedarse en la capa superficial o entrar más adentro y explorar los Raw data.

En el siguiente ejemplo se visualiza en un mapa de burbujas las mayores violaciones de datos del mundo. A primera vista se ven varias burbujas, cada una con el nombre de la compañía y el tamaño de la violación de datos. Cada burbuja es proporcional al tamaño.

Si se hace clic en una, la burbuja cambia de aspecto y se explica cómo fue el robo de datos. Además, se tiene la posibilidad de aplicar filtros (por ejemplo por categoría de empresa).

Hasta ahora la infografía parece ser un conjunto de las anteriores. Sin embargo, en la parte inferior de ella, se tiene la posibilidad de visualizar los Raw Data. En este caso son un Excel compartido públicamente a través de Google Drive.

Como se hizo con ejemplos anteriores, también ahora se visualizarán unas capturas de pantalla de la infografía ya que es interactiva y animada.

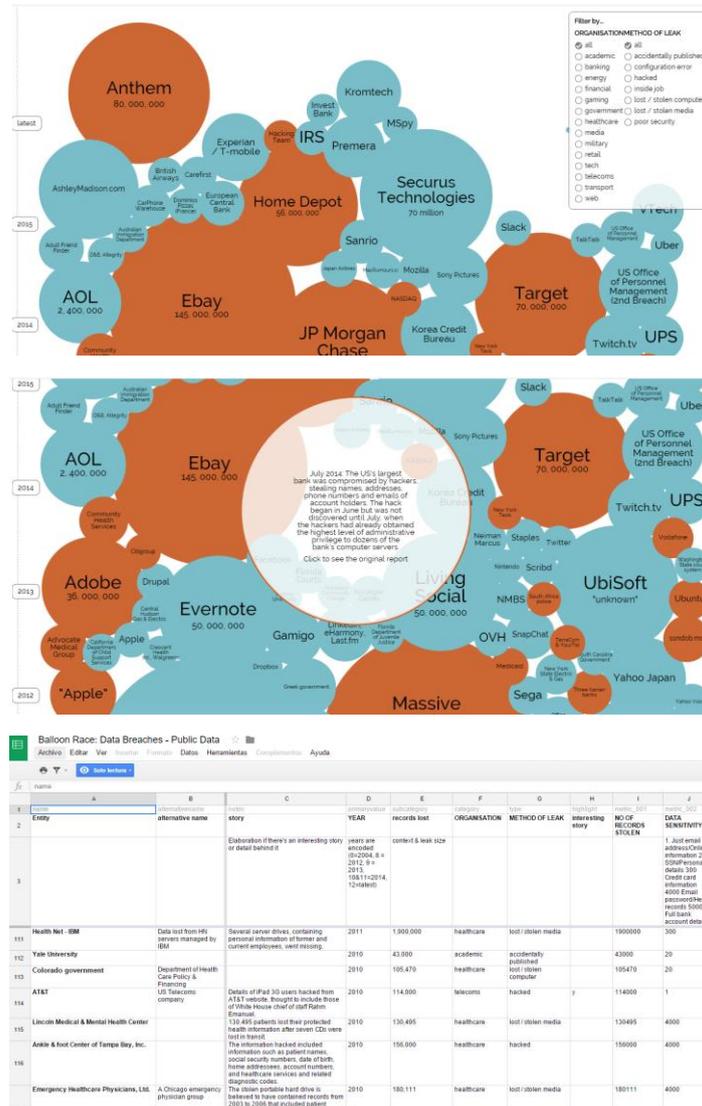


Ilustración 29 Ejemplo de Infografías que proporcionan accesos a los "Raw data", 30 de agosto de 2016, de <http://blog.visme.co/examples-data-visualizations/>



4 Los tipos de gráficos

4.1 Los gráficos

En este apartado se analizarán los principales tipos de gráficos estadísticos.

Esta sección se ha desarrollado como una introducción al siguiente apartado, donde se analizarán más en detalle las Redes Sociales y las maneras de visualizarlas.

Un gráfico estadístico es una representación visual de una serie de datos estadísticos.

En estadística se denominan “gráficos” las imágenes que, combinando la utilización de colores, puntos, líneas, símbolos, números, texto y un sistema de coordenadas, permiten representar información. [26]

Además de poderse utilizar como sustituto de las tablas, los gráficos constituyen por sí mismos una eficaz herramienta para el análisis de los datos, ya que en ocasiones son el medio más efectivo no sólo para describir y representar la información, sino también para analizarla directamente.

Sin embargo en este apartado solo vamos a catalogar los principales gráficos utilizados como vehículo de presentación de datos, sin afrontar su otra mansión como herramienta de análisis.

Un gráfico, para ser un buen vehículo de representación de datos, tiene que tener las siguientes propiedades:

- Atrae la atención del lector
- Representa la información de forma sencilla, clara y precisa
- no induce al error
- Facilita la comparación de datos y destaca tendencias y diferencias
- Muestra el mensaje, tema o asunto del texto al que acompaña

4.2 Los tipos de datos

Para adentrarnos en el análisis de los gráficos, tenemos que definir los tipos de datos que existen:

4.2.1 Datos Cualitativos

Los datos cualitativos son datos que se refieren a cualidades o modalidades que no pueden expresarse numéricamente.

Estos datos pueden ser:

- Ordinales: los datos cualitativos ordinales siguen un orden o secuencia, por ejemplo el abecedario, los meses del año etc.
- Categóricos: los datos cualitativos categóricos no siguen ningún orden, por ejemplo el estado civil de las personas: solteros, casados, viudos, divorciados y separados etc.

4.2.2 Datos Cuantitativos

Los datos cualitativos son datos que se refieren a cantidades o valores numéricos.

Estos datos pueden ser:

- Discretos: los datos cuantitativos discretos son datos que toman valores enteros (0, 1, 2, 3...). Por ejemplo, el número de hijos, el número de alumnos de una clase etc.
- Continuos: los datos cuantitativos discretos son datos que pueden tomar cualquier valor dentro de un intervalo, por ejemplo la estatura o el peso de las personas etc.



4.3 Los principales tipos de gráficos

En esta sección se pretenden analizar los principales tipos de gráfico. Ya que existen muchísimos tipos de gráficos, se han elegidos los gráficos mencionados en el apartado “Explica” del INE, Instituto Nacional de Estadística, en la sección “Tipos de gráficos”. [27]

4.3.1 Gráfico de barras

Un gráfico de barras es una representación gráfica en un eje cartesiano de las frecuencias de una variable cualitativa o discreta.

En uno de los ejes se posicionan las distintas categorías de la variable cualitativa o discreta y en el otro el valor o frecuencia de cada categoría en una determinada escala.

En el siguiente ejemplo se puede ver la producción de cereales en España en el año 2007.

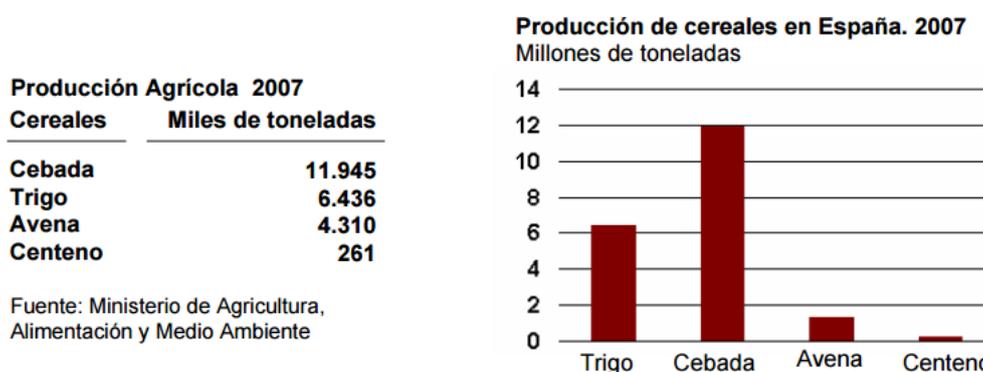


Ilustración 30 Ejemplo de un gráfico de barras

En el eje horizontal se ha posicionado el tipo de cereal mientras en el eje vertical la producción en millones de toneladas.

En el ejemplo anterior se ha utilizado el gráfico de barras para comparar magnitudes de varias categorías, sin embargo este tipo de gráfico se puede utilizar también para ver la evolución en el tiempo de una magnitud concreta.

En el siguiente ejemplo, se representa la producción de cereales en España desde el año 2000 hasta el 2008.

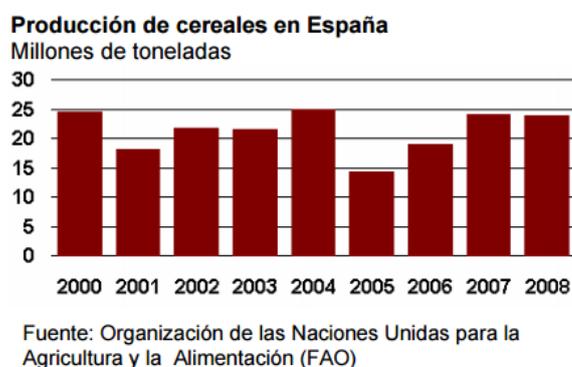


Ilustración 31 Ejemplo de un gráfico de barras donde se quiere focalizar sobre la evolución en el tiempo

En el ejemplo anterior se ha elegido representar la información a través de un gráfico de barras con orientación vertical, sin embargo existe la posibilidad de utilizar un gráfico de barras con orientación horizontal, donde las categorías se sitúan en el eje vertical y las barras crecen horizontalmente.

Los gráficos de barras horizontales suelen usarse cuando hay muchas categorías o sus nombres son demasiado largos.

En el siguiente ejemplo se representa el porcentaje habitantes usuarios de internet del año 2007 por países, según la fuente Unión Internacional de Telecomunicaciones.

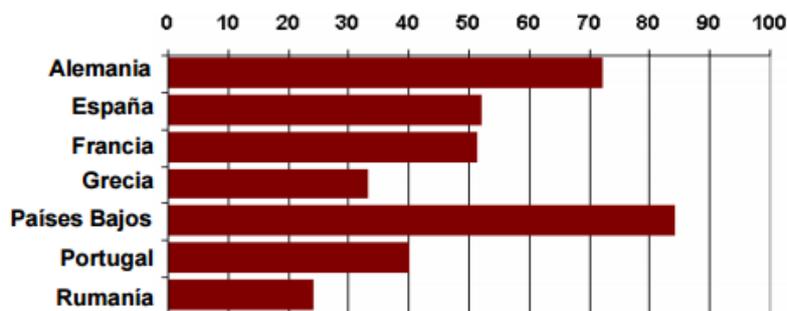


Ilustración 32 Ejemplo de gráfico de barras horizontal

En los ejemplos anteriores se han visualizado gráficos que representaban solo una única serie de datos. Sin embargo, a través de los gráficos de barras, se pueden representar varias series de datos en el mismo gráfico.

4.3.1.1 Gráfico de barra agrupado

Este gráfico contiene varias series de datos y cada una se representa por un tipo de barra de un mismo color o textura.

En el siguiente ejemplo se representa el número de alumnos (en miles) que terminó Bachillerato por su oposición académica.

La fuente de los datos es el Ministerio de Educación y se refiere al curso 2006/07

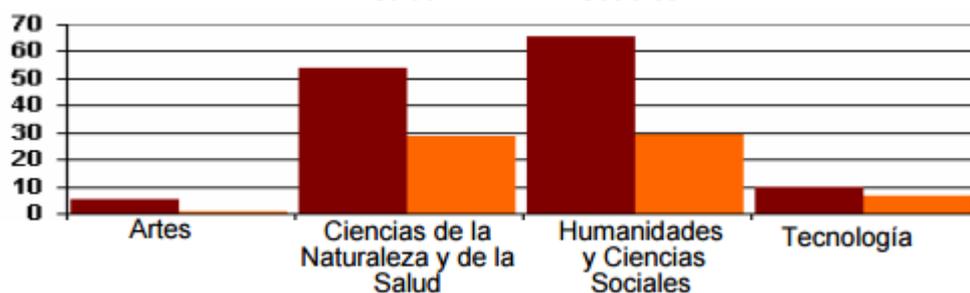


Ilustración 33 Ejemplo de gráfico de barras agrupado

4.3.1.2 Gráfico de barras apilado

Este gráfico contiene varias series de datos, la barra se divide en segmentos de diferentes colores o texturas y cada uno de ellos representa una serie. A continuación se muestra la representación del mismo conjunto de datos del ejemplo anterior utilizando un gráfico de barras apilado

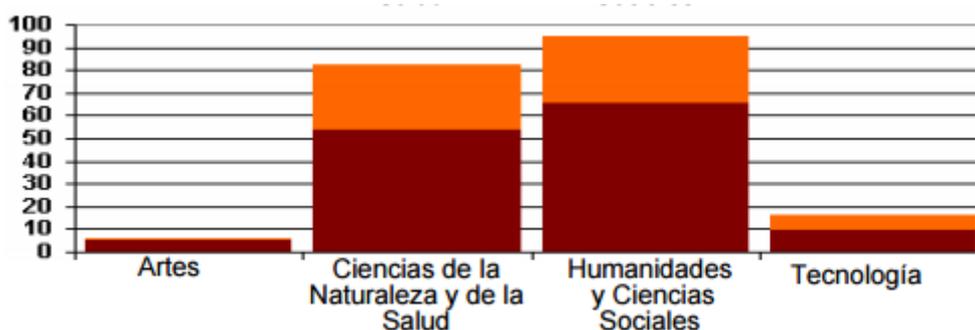


Ilustración 34 Ejemplo de un gráfico de barras apilado

4.3.1.3 Histograma

El Histograma se utiliza para representar las frecuencias de una variable cuantitativa continua.

En uno de los ejes se posicionan las clases de la variable continua (los intervalos o las marcas de clase, es decir los puntos medios de cada intervalo) y en el otro eje las frecuencias. No se prevé la utilización de separación entre las barras.

En el siguiente ejemplo se representa el número de hogares según ingresos, en 2008, según una encuesta de Presupuestos Familiares del INE, el Instituto Nacional de Estadística. Para representar los ingresos se han elegido intervalos.

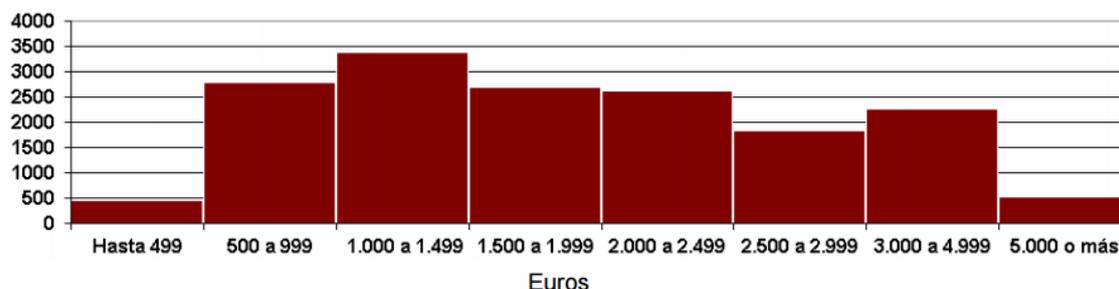


Ilustración 35 Ejemplo de un histograma

4.3.1.4 Histograma Bi-direccional

El gráfico Bi-direccional es un gráfico parecido al gráfico de barras horizontal visto anteriormente, solo que contiene dos series de datos cuyas barras de frecuencias crecen en sentidos opuestos.

En el siguiente ejemplo se representa el consumo de tabaco según sexo y grupos de edad en el año 2006. EN el gráfico se representa el porcentaje de fumadores diarios y la fuente de los datos es la Encuesta Nacional de Salud del año 2006 del INE, el Instituto Nacional de Estadística.

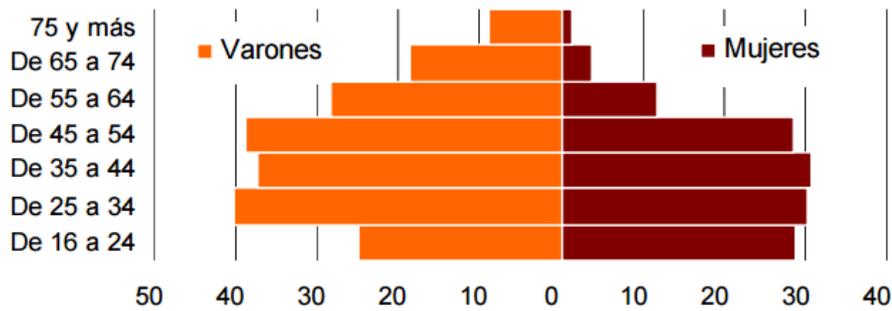


Ilustración 36 Ejemplo de histograma bi-direccional

4.3.1.5 Pirámide de población

Una pirámide de población es un caso particular de un histograma Bi-direccional que se focaliza en enseñar la estructura demográfica de una población, normalmente por sexo y edad, en un momento determinado del tiempo.

La utilidad de estos gráficos es que nos dan una visión de la juventud, madurez o vejez de una población y, por tanto, su grado de desarrollo.

En el siguiente ejemplo se representa porcentaje de la población española según sexo y grupos de edad en el año 2009. La fuente de los datos es la Explotación estadística del Padrón del año 2009 del INE, el Instituto Nacional de Estadística.

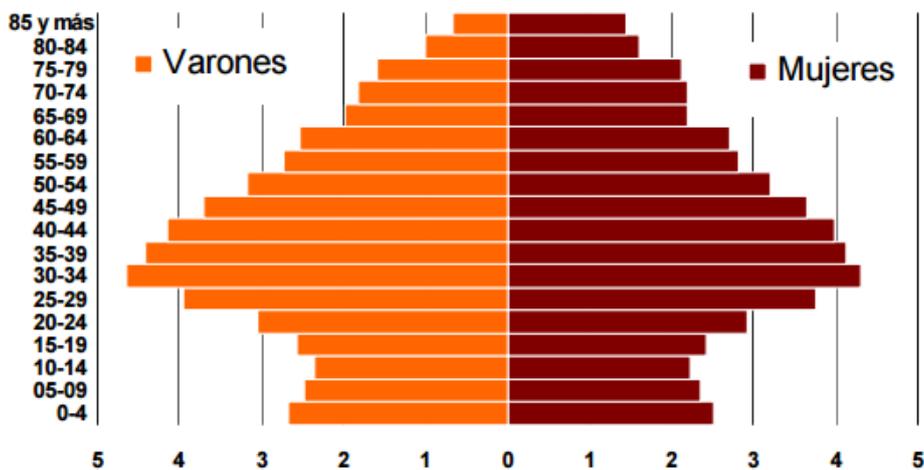


Ilustración 37 Ejemplo de pirámide de población

Del gráfico se puede observar que en el eje vertical se posicionan las clases de edad y en el horizontal los porcentajes de población.

En una de las direcciones se colocan las barras que representan la distribución por edad de los varones y en la otra la distribución por edad de las mujeres.

Es posible encontrar también varias series de datos en un mismo gráfico, como por ejemplo, la población en distintos años o de distintos países.

En el siguiente ejemplo se puede ver la pirámide de población española según sexo y grupos de edad en el año 2009 y una estimación de lo que será en el año 2018.

La fuente de los datos es la Explotación estadística del Padrón, estimaciones y proyecciones del año 2009 del INE, el Instituto Nacional de Estadística.

Como ya dicho anteriormente, a través de este gráfico se puede entender el estado de desarrollo de un país.

Según sus resultados, existen varios tipos de pirámides de población:

- Pirámide de población progresiva

Se denomina Pirámide de población progresiva cuando en el gráfico resultante se puede notar que existe un alto porcentaje de población joven que va desapareciendo según avanzan las edades.

Este resultado es típico de países subdesarrollados cuya esperanza de vida es baja y la tasa de natalidad alta.

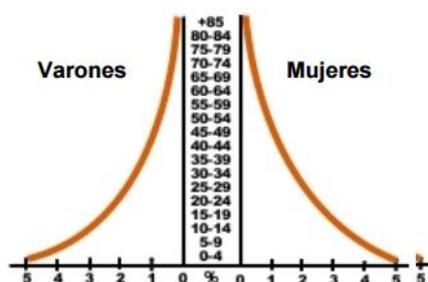


Ilustración 38 Ejemplo de Pirámide de población progresiva

- Pirámide de población regresiva

Se denomina Pirámide de población regresiva cuando en el gráfico resultante se puede notar que en la base existe menos población que en el medio y la población envejecida es considerable.

Este resultado es típico de países desarrollados cuya natalidad está descendiendo y la esperanza de vida es alta.

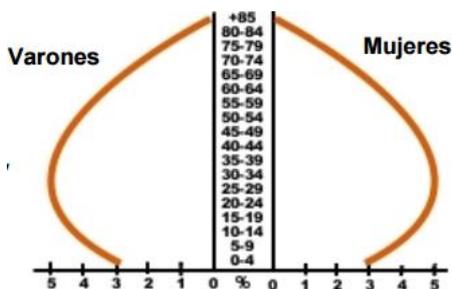


Ilustración 39 Ejemplo de Pirámide de población regresiva

- Pirámide de población estancada

Se denomina Pirámide de población estancada cuando en el gráfico resultante se puede notar que los tramos intermedios de edades tienen la misma población que la base.

Este resultado es típico de países en vías de desarrollo donde se ha controlado la mortalidad y se empieza a controlar la natalidad.

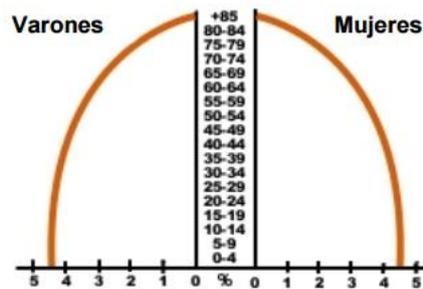


Ilustración 40 Ejemplo de Pirámide de población estancada

4.3.2 Gráfico de líneas

La función principal de un gráfico de líneas es mostrar cómo cambia algo con el correr del tiempo. En este gráfico, un conjunto de puntos es conectado por medio de líneas que, entre todas, logran mostrar la dinámica más o menos regular del comportamiento de algo en relación con otra variable.

Este gráfico se suelen usar para presentar tendencias temporales.

En el eje horizontal se representa la variable que indica las unidades de tiempo y en el eje vertical se introduce la escala de la variable cuya variación en el tiempo queremos ver.

En el siguiente ejemplo se representa el número de parados (en miles de personas) en España desde 2005 a 2009 por nivel de formación alcanzada. Cada línea de color diferente representa un nivel de formación. En el eje horizontal podemos ver el tiempo y en el vertical la cantidad de personas.

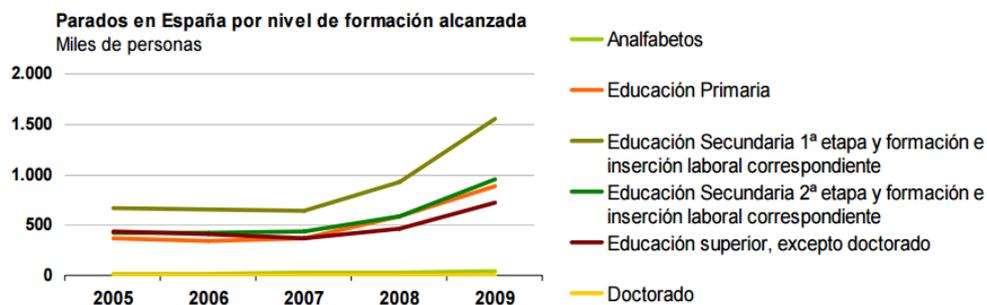


Ilustración 41 Ejemplo de gráfico de líneas

La fuente de estos datos en la Encuesta de Población Activa que el INE realizó en 2009.

4.3.3 Relación entre gráficos de líneas y gráficos de barras

Juntando los dos tipos de gráficos vistos anteriormente, se puede obtener varios tipos de gráficos.

Uno de estos es el polígono de frecuencias. Es decir, un gráfico de barras donde los puntos medios de cada barra son unidos con líneas.

En el siguiente ejemplo se puede observar el polígono de frecuencia creado a partir de los datos de producción de cereales en España desde el año 2000 al 2008. La fuente de estos datos es la Organización de las Naciones Unidas para la Agricultura y alimentación (FAO).

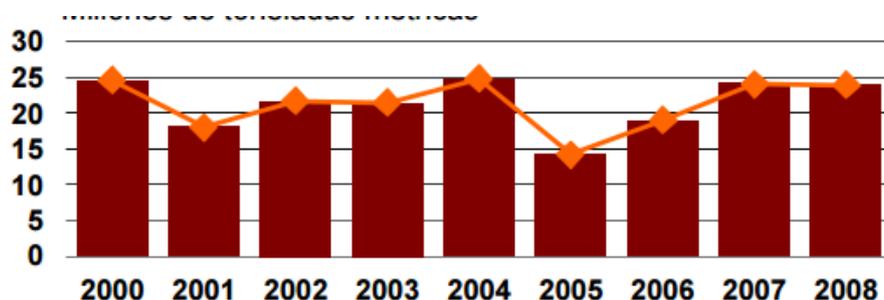


Ilustración 42 Ejemplo de polígono de frecuencia

Otro tipo de gráfico creado juntando un gráfico de barra y de línea es otra interpretación de la pirámide de población, ya que se puede representar también con un gráfico de líneas.

Esto es muy útil cuando se van a utilizar varias series de datos, y se quieren compararlas. Por ejemplo, si se quieren comparar las pirámides de varios países, ver la población extranjera o comparar la población de un mismo país en varios años.

En el siguiente gráfico se ha querido visualizar el porcentaje de la población Española en 2010, dividida por franja de edad, dividida por sexo y por residencia (español o extranjeros). La fuente de los datos es la Explotación estadística del Padrón realizada por el INE en 2010.

4.3.4 Gráfico de sectores

Los gráficos de sectores, también llamados de torta, muestran la distribución de un determinado total en diferentes partes. Es valiosa para los casos en los que se conoce el absoluto, y lo que interesa es conocer la forma en la que eso se repartió en diferentes partes.

Este gráfico es una representación circular de las frecuencias relativas de una variable cualitativa o discreta que permite poder compararla fácilmente.

El círculo representa la totalidad que se quiere observar y cada sector representa la proporción de cada categoría del variable respecto el total. La porción suele expresarse en porcentajes.

En el siguiente ejemplo, el porcentaje de viajeros hospedados en hoteles españoles por categoría del establecimiento, categorizado según número de estrellas. La fuente es la Encuesta de Ocupación en Alojamientos Turísticos, realizada por el INE en 2009.

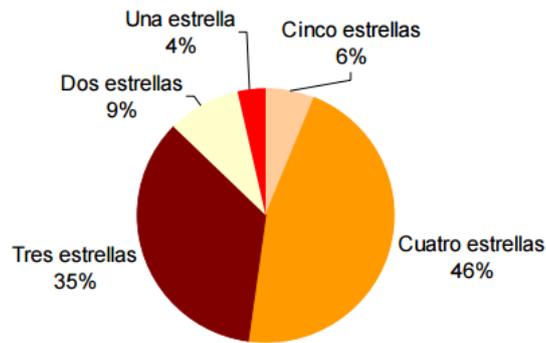


Ilustración 43 Ejemplo de gráfico de sectores

Hay que observar que estos tipos de gráficos son útiles cuando se representan pocas categorías, sin embargo si hay muchas categorías son incomprensibles. En el siguiente ejemplo se visualiza el porcentaje de superficie de los países europeos.

Como se puede notar, la mayoría de datos no se pueden entender.

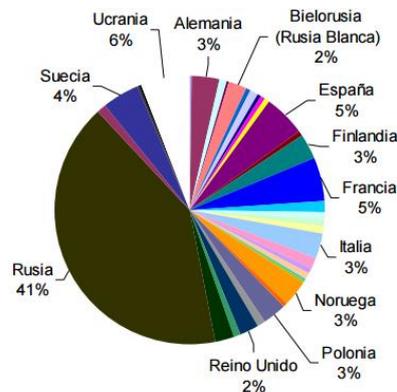


Ilustración 44 Ejemplo de gráfico de sectores donde se visualiza el porcentaje de superficie de los principales países de la UE

4.3.5 Gráfico de dispersión

Un gráfico de dispersión muestra en un eje cartesiano la relación que existe entre dos variables. La particularidad de este gráfico es que informa del grado de correlación entre las dos variables, es decir, nos muestra si el incremento o disminución de los valores de una de las variables, denominada variable independiente y que se suele representar en el eje horizontal, altera de alguna manera los valores de la otra, denominada variable dependiente y que representa generalmente en el eje vertical.

Según la forma de la nube de puntos que se genera, se puede deducir el tipo de correlación:

- **Correlación nula**
Se obtiene correlación nula cuando no hay ningún tipo de relación entre las variables. En este caso, se dice que las variables son independientes. En el siguiente gráfico se puede observar un ejemplo de correlación nula.

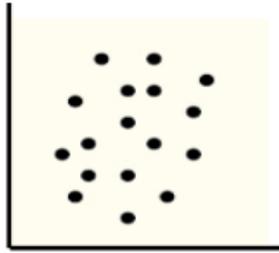


Ilustración 45 Ejemplo de gráfico de dispersión con correlación nula

- Correlación lineal

Se obtiene correlación lineal cuando hay relación entre las variables de forma lineal. Más en detalle, la correlación lineal es negativa cuando al aumentar de los valores de la variable independiente disminuyen los valores de la variable dependiente y, al contrario, la correlación lineal es positiva si al aumentar los valores de la variable independiente aumentan los valores de la variable dependiente.

En los siguientes gráficos se puede observar un ejemplo de correlación lineal negativa y positiva.

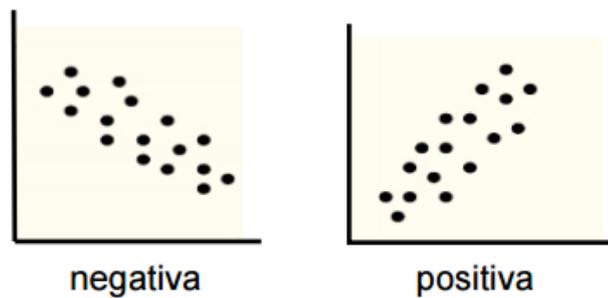


Ilustración 46 Ejemplo de gráfico de dispersión con correlación lineal

Normalmente en el caso de correlación lineal, para facilitar la visión, se suele dibujar la recta de regresión, que se obtiene a través del ajuste lineal.

Esta recta nos sirve para entender la relación de los puntos, sin tener que analizar a fondo el gráfico.

En el siguiente ejemplo se puede observar una relación lineal positiva con la recta de regresión dibujada.

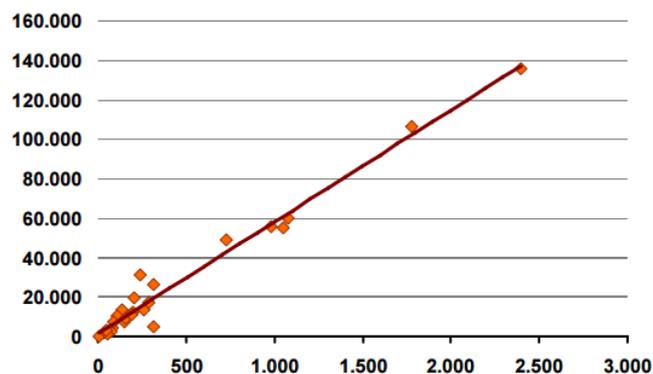


Ilustración 47 Ejemplo de relación lineal positiva

4.3.6 Cartograma

Un gráfico Cartograma es un mapa en el que se representan datos estadísticos. Normalmente se suelen dividir por regiones, y en cada una se representan directamente los números o se colorean las distintas zonas en función del dato que representan.

Los colores utilizados suelen ser explicados en la leyenda.

En el siguiente ejemplo se visualiza la tasa de paro en España en el 2009. La fuente de este gráfico es la encuesta “España en cifras” realizada por el INE en 2009.



Ilustración 48 Ejemplo de cartograma

5 Las redes sociales y el análisis de redes sociales

En este apartado vamos a enfocarnos en el análisis de redes sociales. Hay que detallar que estamos hablando de redes sociales como grupos sociales, no se refiere únicamente a las redes sociales online.

Una red social es compuesta por un conjunto de actores y por las relaciones que hay entre ellos, y existen varios tipos diferentes de red social. A nivel básico, una red social se puede representar mediante un grafo o una matriz. Además de las redes sociales básica, en este apartado se van a analizar también las redes sociales multinivel.

Posteriormente, nos enfocaremos en la teoría de la visualización de las redes sociales, la teoría de grafos, que sirve para resolver problemas que pueden ser modelos mediante un grafo y resueltos mediante algoritmos específicamente desarrollados para un grafo.

El análisis de redes sociales, como ya dijimos anteriormente, no es solo análisis de las redes sociales entendidas como análisis de contenido de Social Media. Es un estudio numérico, algebraico, que representa el conocimiento en formato de grafo. Es una mezcla entre la sociología y las matemáticas (el álgebra de grafos) en el que hay actores o entidades que interactúan y que representan estas acciones mediante un grafo.

Para concluir este apartado, vamos a investigar sobre los softwares para el análisis de las redes sociales, analizando los programas más populares en el mercado y haciendo una comparativa.



5.1 Que es una Red Social

Para introducir el análisis de redes sociales, es obligatorio definir las redes sociales, ya que este término puede crear ambigüedades.

Existen multitudes de definiciones de las redes sociales.

Por ejemplo, Streeter y Gillespie definen una red social “como cualquier conjunto limitado de entidades sociales conectadas” [28] o según Wasserman y Faust una red social es “un conjunto finito de actores y las relaciones que los vinculan” [29].

Además según Knoke y Yang una red social “es una estructura formada por un conjunto de actores, donde parte de ellos están conectados por un una o más relaciones” [30].

Para dar una definición global, podemos decir que una red social es una estructura compuesta por un conjunto finito de actores y que contiene también una serie de relaciones entre ellos, que puede ser representada por uno o varios grafos. Los miembros de los grafos son los nodos, que representan los actores, y las artistas, que representan las relaciones entre ellos.

Para determinar si un actor pertenece a la red, se utiliza el límite. Dicho también “boundarie”, el límite es un criterio gracias al cual establecemos que un conjunto de actores pertenecen a la red y otros no [31].

En la siguiente imagen se puede ver un ejemplo de grafo de una red social, donde hay actores que pertenecen a la red y otros no.

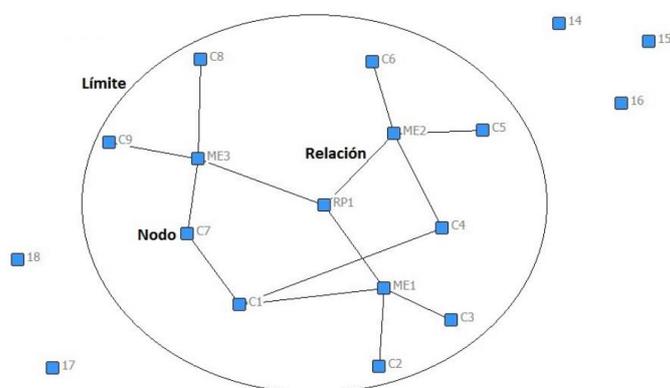


Ilustración 49 Ejemplo de grafo de una red social

En resumen, si tenemos un conjunto de actores, para crear y definir a una red social, antes de todo debemos determinar las relaciones que nos interesa estudiar para poder definir el límite de la red y determinar el conjunto de actores involucrados.

Para que una red social sea relevante, estas relaciones tienen que influir en el comportamiento de los actores.

5.1.1 Elementos de una red social

Como ya se dijo en la introducción, los elementos centrales de las redes sociales son los actores, las relaciones entre ellos y los límites de la red. A continuación se van a analizar más a fondo estos tres elementos.

5.1.1.1 Actores

Los actores de una red se representan en el grafo con los nodos y pueden ser individuos, grupos o instituciones. En mismo grafo pueden existir distintos tipos de actores, y por esta razón puede haberse todo tipo de relaciones, por ejemplo relaciones entre individuos e instituciones.

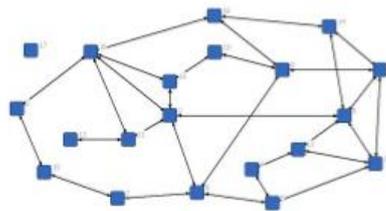
Los grafos con un solo tipo de actores son muy limitados ya que pueden tener solo limitados tipos de relaciones.

Más en detalles, una de las formas de identificar las redes es en base al tipo y la cantidad de los conjuntos de actores involucrados (sets of actors). Existen por eso tres *modes* de redes:

1. Redes Modo 1 (one-mode networks)

En esta tipología aparece un único conjunto de actores y las relaciones que los vinculan. Que aparezca un único conjunto de actores significa que todos los actores tienen que tener la misma naturaleza, por ejemplo todos tienen que ser individuos o todos grupos.

En el siguiente ejemplo se puede ver una red modo 1, donde todos los nodos son del mismo tipo, en este caso estudiantes.



• Students

Ilustración 50 Ejemplo de una red modo 1

2. Redes Modo 2 (two-mode networks)

En esta tipología pueden aparecer:

- Dos conjuntos de actores de naturaleza distinta (por ejemplo individuos y grupos) y las relaciones que los vinculan
- Un conjunto de actores, un conjunto de eventos y las relaciones que los vinculan.

Este segundo tipo de red también se puede llamar red de afiliación (affiliation network).

En este tipo de red y en las superiores (es decir, las con un grado mayor igual que 2), sus componentes no son solos actores, ya que se admiten naturalezas diferentes (actores de distinto tipo, eventos o fenómenos).

Por esta razón se les suelen denominar entidades sociales (social entities).

En este tipo de redes entonces se estudian los vínculos entre conjuntos de entidades sociales finitos en un periodo de tiempo y espacio determinados.

En la siguiente imagen se puede ver un ejemplo de red modo 2, donde los actores son Administradores y Sociedades.

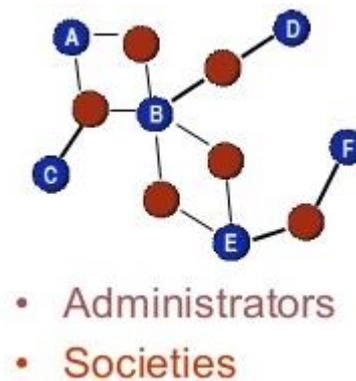


Ilustración 51 Ejemplo de una red modo 2

3. Redes modo N

En esta tipología se estudian 3 o más (N) conjuntos de entidades sociales.

Estas redes pueden ser mucho más complejas de las anteriores, ya que pueden aparecer muchísimos tipos de naturalezas y la cantidad de entidades sociales puede ser muy alta.

Sobre todo, pueden ser mucho más complejas las relaciones entre entidades sociales.

En resumen, dependiendo de la cantidad de tipos distintos de entidades sociales vinculados en una red, se establece el número de modos.

5.1.1.2 Relaciones

Se define relación la conexión entre un par de actores o entidades sociales.

Se pueden categorizar principalmente de dos maneras:

1. Teniendo en cuenta la direccionalidad de la relación.

En este caso, podemos encontrar dos tipos de relación:

- Transitiva (non-directed relation) cuando la relación entre las entidades es recíproca, es decir la relación entre A y B es la misma que entre B y A.
- Directa (directed relation) cuando existe un actor activo y otro pasivo, por ejemplo A tiene relación con B pero eso no implica que B la tenga con A.

2. Teniendo en cuenta la densidad de la relación.
 Este caso se puede realizar solo si la relación puede cuantificarse de tal manera que podamos estudiar las cantidades que pasan en la red entre nodo y nodo (por ejemplo, flujos de dinero en una red de prestamistas).
 La medida resultante de esta cuantificación determina la densidad de la red.

Como anteriormente ya vimos que se pueden distinguir las redes según la cantidad de conjuntos diferenciados de actores involucrados (los modos de la red), también se pueden diversificar según la cantidad de relaciones involucradas.

A continuación se definen los tipos de red diferenciados por relaciones:

- Red Uniplex: Red formada por una única relación entre un conjunto de actores
- Red Multiplex: Red formada por dos o más relaciones vinculando al mismo conjunto de actores.

Ahora que hemos analizado la clasificación de una red según la cantidad de actores y según la cantidad de las relaciones, podemos juntar estas dos metodologías y definir las redes multidimensionales.

En la tabla a continuación se puede observar la clasificación analítica de las redes sociales según la cantidad de actores y la cantidad de las relaciones.

| | Modo 1 | Modo 2 y Modo N |
|-----------|--------------------|--|
| Uniplex | Unimodal-Uniplex | Multimodal-Uniplex |
| Multiplex | Unimodal-Multiplex | Multimodal-Multiplex (Multidimensional) |

Ilustración 52 Tabla de definiciones Plex-Modes

Castels y Monge han completado esta clasificación introduciendo el concepto de red multidimensional [32].

Según Castels y Monge, una red multidimensional es una red multimodal y multiplex.

Se considera red multidimensional porque ya que presenta conjuntos de actores de naturaleza diferente y múltiples relaciones, aparecen múltiples dimensiones de la interacción social.

En el siguiente ejemplo se puede observar una red multidimensional. Hay dos tipos de naturaleza de actores, las persona (cuadrado) y las tecnologías (círculos).

Además, hay diversos tipos de relaciones entre ellos: La flecha rosa significa que “aporta información”, la flecha azul que “recupera información”.

Como ya dijimos anteriormente, una red multidimensional ha de tener 2 o más tipos de naturalezas de actores (en este caso 2) y dos o más tipos de naturalezas de relaciones (en este caso 2).



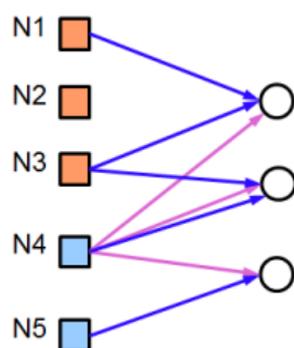


Ilustración 53 Ejemplo de red con 2 tipos de naturaleza de los actores

El concepto de red multidimensional puede ser ampliado diferenciando también entre una red multidimensional parcial y completa.

- Red multidimensional Parcial:
Una red multidimensional es parcial cuando, además de ser multimodal (de modo superior o igual a 2) y multiplex, las relaciones se aplican entre las entidades sociales de distintos conjuntos, pero no sucede que relaciones entre las entidades sociales hacia el interior de un mismo conjunto.
- Red multidimensional Completa:
Una red multidimensional es completa cuando, además de ser multimodal (de modo superior o igual a 2) y multiplex, las relaciones se aplican entre las entidades sociales de distintos conjuntos, y además pueden haber relaciones entre las entidades sociales hacia el interior de un mismo conjunto.

En el siguiente grafo se puede ver un ejemplo de red multidimensional parcial.

Se trata de un grafo utilizado para investigar las relaciones que hay sobre clientes, mediadores y un programa municipal llamado “soluciones habitacionales”.

Existen 3 conjuntos de actores:

- Los clientes, denominados “C” en el grafo, y representados por un círculo. Además, los clientes que adoptan el plan de ayuntamiento son verdes y quien no lo adoptan son amarillos.
- Los mediadores, denominados “M” en el grafo, y representados por un cuadrado.
- EL programa municipal “Soluciones habitacionales” denominado “Plan” y representados por un triángulo.

Existen también 3 tipos de relaciones entre los actores:

- Una relación directa de apoyo político de los clientes hacia los mediadores
- Una relación transitiva de intercambio de información entre los mediadores y el programa de “soluciones habitacionales”
- Una relación directa que representa el otorgamiento de beneficios que el programa da a los clientes

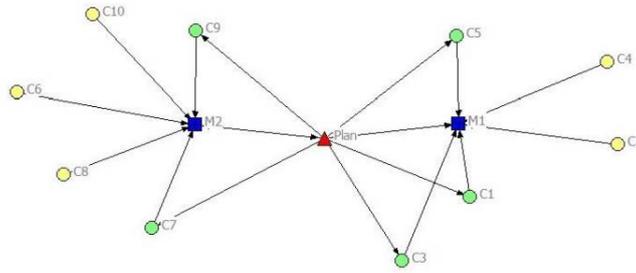


Ilustración 54 Ejemplo de red multidimensional parcial y no completa

Como se puede notar del grafo, estamos en presencia de una red multidimensional parcial y no completa ya que no hay relaciones entre actores de la misma naturaleza.

Por ejemplo, no hay relaciones entre los cliente o los mediadores a su interno, pero si que hay relación entre actores se cada conjunto con actores de los demás conjuntos (Clientes-Mediadores, Mediadores-Plan, Plan-Clientes).

5.1.1.3 Límites de la red

Para determinar la pertenencia de un actor a la red se utiliza el criterio de los límites de la red.

El límite de la red está emparejado con las relaciones que constituyen la misma red, ya que es el tipo de relaciones el que determinan quiénes sujetos pueden participar y quiénes no.

Para demarcar una red y poder realizar un análisis basado en evidencia sobre la misma debemos dar cuenta de los actores determinados en torno a los cuales buscamos inferir una particular red social partiendo de sus vínculos [33].

5.2 Análisis de redes sociales

En el apartado anterior definimos que es una red social, como se puede representar a través de un grafo y focalizamos la atención sobre los elementos del grafo, es decir los actores, las relaciones y los límites.

Aplicando la representación de las redes sociales que vimos anteriormente a la sociedad actual, se puede observar que en los últimos tiempos han crecido enormemente las relaciones entre agentes [34]

Por ejemplo, a la hora de realizar una compra, es mucho más frecuente confiar en la opinión de los usuarios que en la publicidad de la propia marca. Otro ejemplo podría ser la enorme red de proveedores y clientes que se ha formado gracias a la globalización de la economía y la interconexión internacional.

Dicho esto, nos interesa mucho poder representar estas enormes redes a través de un grafo, ya que de esta manera dado un problema, puede ser resuelto mediante algoritmos específicos de la teoría de grafos y la información que podemos obtener es muy relevante.

La visualización, elaboración y análisis de todo tipo de redes y sistemas complejos se puede definir con la palabra sociometría.

El fundador de la Sociometría [35], Jacob Levy Moreno [36], la describió de esta manera:

“La sociometría tiene por objeto el estudio matemático de las propiedades psicológicas de las poblaciones.



Ilustración 55 Jacob Levy Moreno, 30 de agosto de 2016, de http://files.lecturas-mic2.webnode.es/system_preview_detail_200000062-982499921f/Moreno.png

Con este fin utiliza una técnica experimental fundada sobre los métodos cuantitativos y expone los resultados obtenidos por la aplicación de estos métodos. Persigue así una encuesta metódica sobre la evolución y la organización de los grupos y sobre la posición de los individuos en los grupos”.

Una vez que tenemos a disposición un grafo (con sus actores y relaciones) podemos explotar la información contenida en el mismo con el fin de sacar conclusiones. Para hacer este proceso se necesitan utilizar varias métricas.

5.2.1 Métricas

Vamos a analizar las métricas principales que se pueden aplicar a un grafo, clasificadas en tres niveles:

5.2.1.1 Nivel global de un grafo

- **Coeficiente de agrupamiento**
Mide el nivel de agrupamiento de los nodos, para entender cómo de cohesionados o integrados están los agentes/actores.
- **Camino característico**
Mide el grado de separación de los nodos, para determinar cómo de separados o alejados están los nodos, y poder buscar así medidas para recortar más la relación entre agentes/actores.
- **Densidad**
Un grafo puede tener muchas aristas (grafo denso) o muy pocas (grafo disperso). En este sentido, se puede descubrir si hay muchas o poca conexión.
- **Diámetro**
Es la distancia mayor entre cualquier par de nodos. De esta manera, sabemos cómo de “alejados” o “cercaños” están a la hora de comparar varios grafos.
- **Grado medio**
Nos permite descubrir la popularidad de un nodo. En detalle, es el número medio de vecinos (conexiones a otros nodos) que tiene un grado, es decir, nos indicará cuál es la media de conexiones que tiene un nodo.
- **Centralidad**
Nos permite descubrir los actores influyentes. Para eso, hay que ver aquellos nodos que poseen una mayor cantidad de relaciones y de esta manera descubrir los influyentes dentro del grupo. Se llega a saber su “popularidad”, lo que puede dar mucha información para saber la importancia de un nodo dentro de un conjunto.

5.2.1.2 Nivel comunidad (grupos de nodos dentro de un grafo)

- **Comunidades**
Una comunidad es una agrupación de varios nodos por patrones de similitud. Es una herramienta para que el nodo se conozca a sí mismo, para conocer a los demás nodos, para analizar el grupo en concreto y en general a los grupos que viven procesos parecidos.



- Puentes entre comunidades:
Los puentes de comunidades describen las conexiones entre las comunidades y como de comunicables son esas comunidades. Normalmente se utilizan para trazar planes de actuación o de marketing.
- Centros locales vs. Periferia
Dentro de una comunidad es interesante saber cuáles son los nodos que son más centrales o críticos y cuáles no.

5.2.1.3 Nivel nodo (propiedades de los influencers)

- Centralidad
También dicho “eigenvector”, con esta métrica podemos descubrir cuanto de popular es un nodo. Por ejemplo, si el valor de centralidad de un nodo A es igual a 0.64, nos indica que si para cada par de nodos buscamos el camino más corto en el grafo, el 64% de estos caminos pasa por dicho nodo A.
Mide la popularidad de un nodo. De hecho, el algoritmo de Google funcionó durante mucho tiempo así, siendo cada nodo una página web o un recurso de la red.
- Modularidad
La modularidad mide la estructura de las redes. Su objetivo es medir la fuerza de la división de una red en grupos (llamados en este caso módulos). Las redes tienen modularidad alta cuando tienen conexiones fuertes entre los nodos dentro de los módulos, pero escasas conexiones entre nodos de diferentes módulos.
- Intermediación
Es la capacidad de un nodo en ocupar una posición intermedia en las comunicaciones entre el resto de los nodos. Si un nodo es presente en muchas comunicaciones, significa que tiene un gran liderazgo, ya que es punto clave los flujos de comunicación.
En el siguiente ejemplo se puede ver un gráfico con el siguiente grupo de usuarios (A, B, C, D, E, F, G) relacionados entre sí. Como se nota en la imagen, quien tiene más capacidad de intermediación será el nodo B ya que ejerce de “puente” entre el grupo conformado por A y por C. Sin B la comunicación entre esos dos grupos sería imposible.



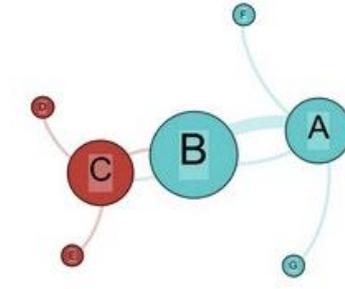


Ilustración 56 Gráfico ejemplo de intermediación

- **Pagerank**
Algoritmo diseñado por los creadores de Google, se basa en dar un valor numérico (denominado ranking) a cada nodo de un grafo. El fin de este valor numérico es medir la conectividad de un nodo.
- **Closeness**
Indica cómo de cerca está un nodo para llegar a contactar con otros. Esto permite conocer la importancia de ese nodo dentro de la red de influencia para eventuales comunicaciones o relaciones con otros nodos.

5.2.2 Aplicación en la vida real

Lo que acabamos de analizar, es decir los grafos con sus componentes y las métricas que se pueden aplicar para obtener resultados, tienen muchas utilidades prácticas.

A continuación se van a exponer unos ejemplos básicos de posibles usos de las métricas de los grafos en diferentes ámbitos:

5.2.2.1.1 Marketing online.

En el marketing online, donde es muy importante la reputación online, se utilizan estas métricas para detectar influencers entre nuestros propios seguidores.

5.2.2.1.2 Deporte

En el deporte, se pueden utilizar estas métricas para elaborar estadísticas sobre los equipos para encontrar puntos débiles y fuertes.

Por ejemplo, dos matemáticos, Javier López Peña y Hugo Touchette de Queen Mary (Universidad de Londres) realizaron un trabajo sobre los equipos del mundial de fútbol de 2010, elaborando una red de pases (passing network) entre los jugadores durante todo el torneo y analizando cómo estas redes se comparan entre los equipos.

Por cada equipo calcularon la centralidad de cada jugador, es decir el nivel de importancia que tiene el jugador para el equipo. En detalle, la centralidad de un jugador se calcula en función de varios parámetros que ya analizamos en el apartado anterior, por cercanía, intermediación y popularidad.

A continuación se puede ver el grafo con la media de los pases de los dos finalistas, España y Holanda.

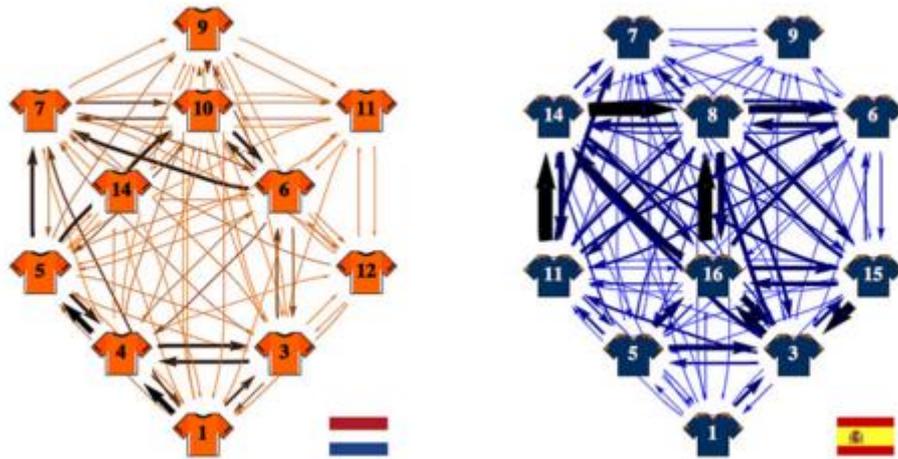


Ilustración 57 grafo con la media de los pases de los dos finalistas, España y Holanda, 30 de agosto de 2016, de https://c2.staticflickr.com/4/3719/13952898494_48c9ea8046_b.jpg

Uno de los datos que más impactó durante la final es que los jugadores de la selección española Xavi y Xabi Alonso fueron la clave en el campeonato de Sudáfrica para que España ganase, en lugar de Iniesta, autor del gol.

5.2.2.1.3 Twitter y redes sociales

En Twitter, se han catalogado seis tipos de conversación que se mantienen entre los usuarios, y para catalogarlas se han utilizados diferentes tipos de grafos.

El centro de investigación Pew, en asociación con la Social Media Research Foundation, ha realizado un informe analizando muchos miles de conversaciones en Twitter, e intentando dibujarlas con grafos. Finalmente se han identificado seis tipos de grafos predominantes.

A continuación se van a enseñar estos seis tipos, explicando la forma en que están formados, el tema principal de discusión y los usuarios que las forman.

- Multitud polarizada (Polarized Crowd).
Las discusiones polarizadas cuentan con dos grupos grandes y densos que tienen poca conexión entre ellos. Los argumentos que se discuten a menudo son temas políticos altamente activos y con muchas diferencias de opiniones. De hecho, por lo general hay poca conversación entre estos grupos a pesar del hecho de que están enfocados en el mismo tema. Es decir, las multitudes polarizadas en Twitter no suelen discutir. Como ya dijimos, se suele formar durante los debates políticos.

En el siguiente grafo se puede ver una multitud polarizada. Ha sido creada recogiendo 688 Tuit entre los días 6 y 8 de Enero de 2013 con el hashtag #My2K. Hay una flecha para cada reacción a los tuits, es decir cada vez que después de un tuit un usuario sigue al autor del tuit, le contesta o lo menciona.

Se ve claramente que hay dos grandes grupos muy densos que hablan cada uno del mismo argumento y casi no se relacionan entre ellos (utilizando las mismas palabras, hashtag o URL).

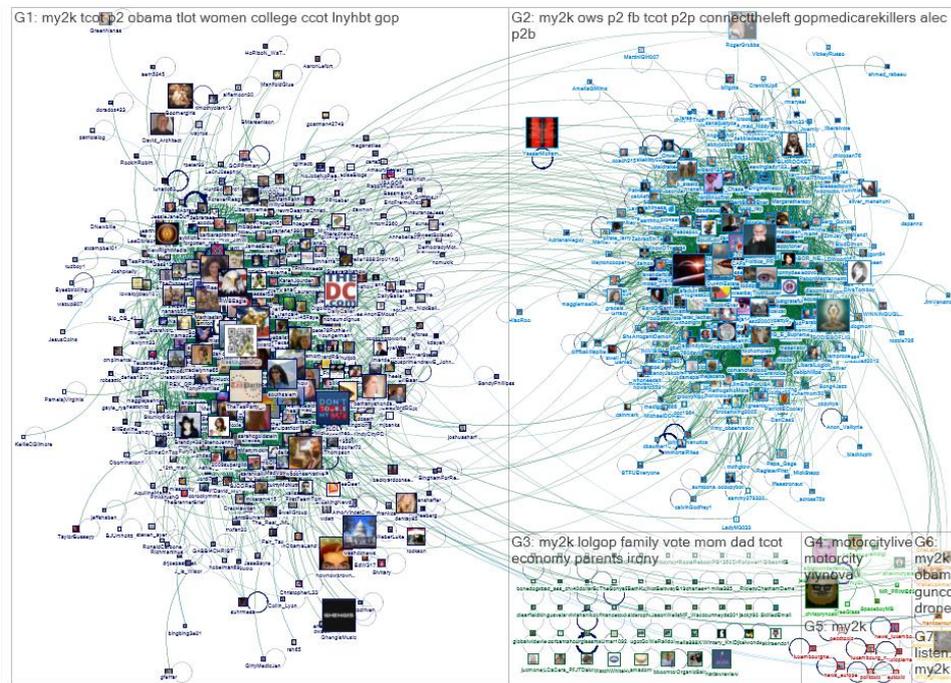


Ilustración 58 Multitud polarizada (Polarized Crowd), 30 de agosto de 2016, de <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>

- Multitud apretada (Tight Crowd). Este tipo de discusión se caracteriza cuando se comunican usuarios altamente interconectados con pocos participantes aislados. Esto pasa sobre todo en conferencias, temas profesionales, grupos de aficionados y otros temas que atraen a las comunidades.

El siguiente grafo ha sido creado analizando los tuits de 268 usuarios entre los días 14 y 18 de Enero de 2013 con el hashtag “#cmgrchat OR #smchat. Estos hashtag se utilizan por gente que utiliza las redes sociales para gestionar sus empresas o negocios (como si fuese una red informal de community managers).

Por ejemplo pueden varias desde las conversaciones sobre temas de salud u organizaciones de rescates de perros.

Se puede ver como dijimos antes que hay muy pocos participantes aislados y que hay muchísimas interconexiones

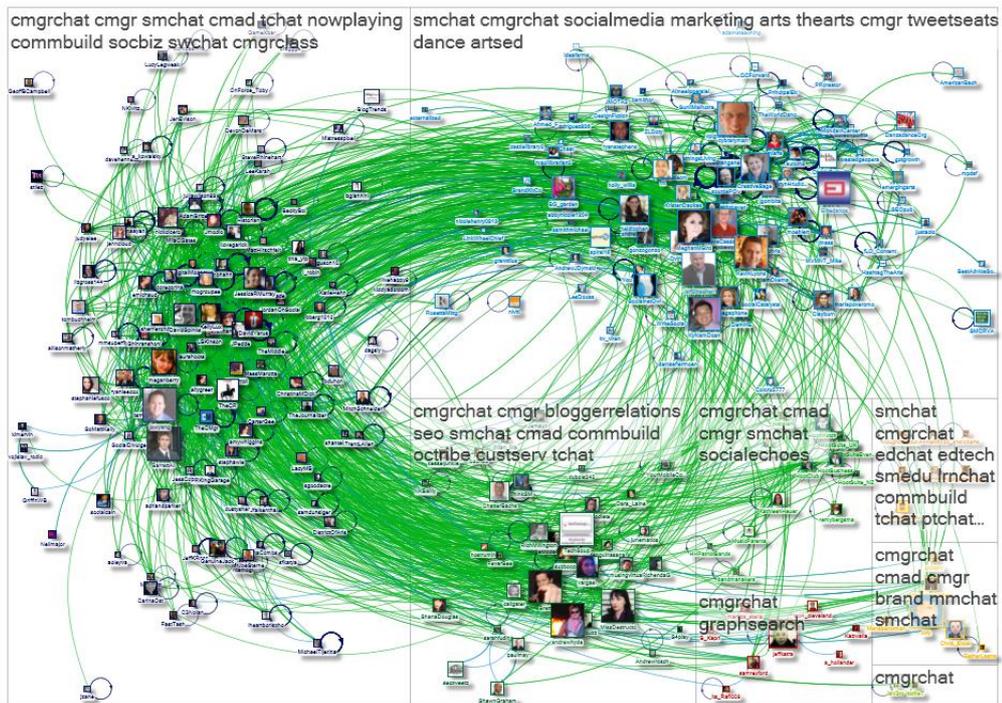


Ilustración 59 Multitud apretada (Tight Crowd), 30 de agosto de 2016, de <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>

- Agrupaciones de marcas (Brand Clusters).
 Este tipo de grafo se materializa cuando se habla de marcas o celebridades. Hay muchas cantidades de usuarios que participan pero no siempre están muy bien conectados entre ellos.
 Este ejemplo ha sido realizado recogiendo 834 los tuits con la palabra o hashtag "Apple" el día 15 mayo de 2013.
 En l parte izquierda de pueden observar los participantes aislados, mientras en la derecha pequeño grupos que se han formado principalmente hablando acerca de un producto de la marca.

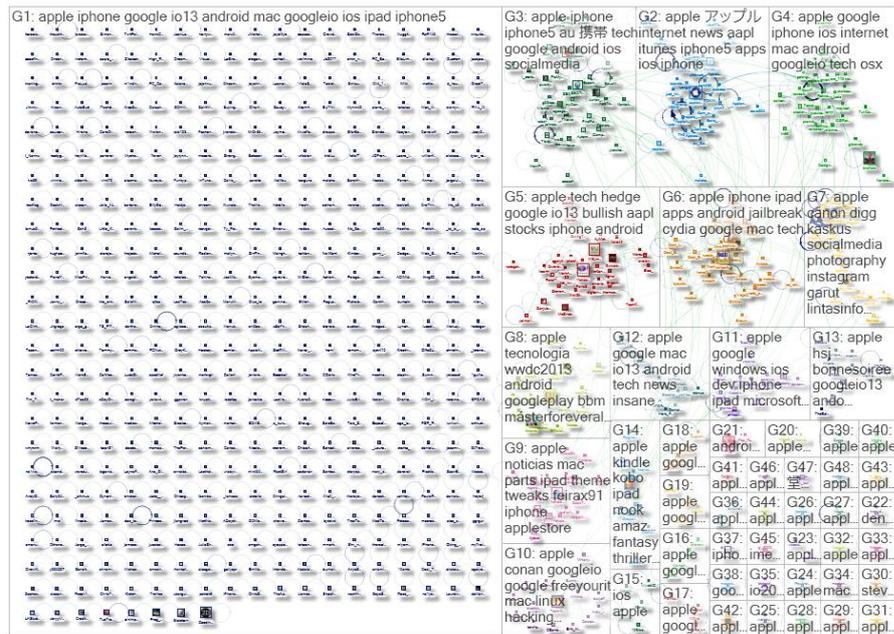


Ilustración 60 Agrupaciones de marcas (Brand Clusters), 30 de agosto de 2016, de <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>

- Cluster de comunidades (Community Clusters).

Algunos temas populares pueden desarrollar múltiples grupos más pequeños, que a menudo se forman alrededor de unos ejes cada uno con su propio público, influyentes, y fuentes de información.

Por ejemplo las noticias mundiales a menudo atraen la cobertura de muchos medios de información, cada uno con su propio grupo de seguidores. Eso crea una colección de grupos y además un buen número de aislamientos

En la siguiente imagen se han recolectado 1260 tuits con la palabra o hashtag Flotus (First Lady of the United States, Michelle Obama).

Se han recogido el día 18 de Enero de 2013 ya que era el cumpleaños de Michelle Obama y se ha aprovechado esta ocasión ya que la administración Obama y muchos activistas han aprovechado este día para crear menciones para generar atención sobre su trabajo.

Esto creó muchos grupos de discusiones pequeños y un buen número de aislamientos, como se puede ver en el lado izquierdo de la siguiente imagen.

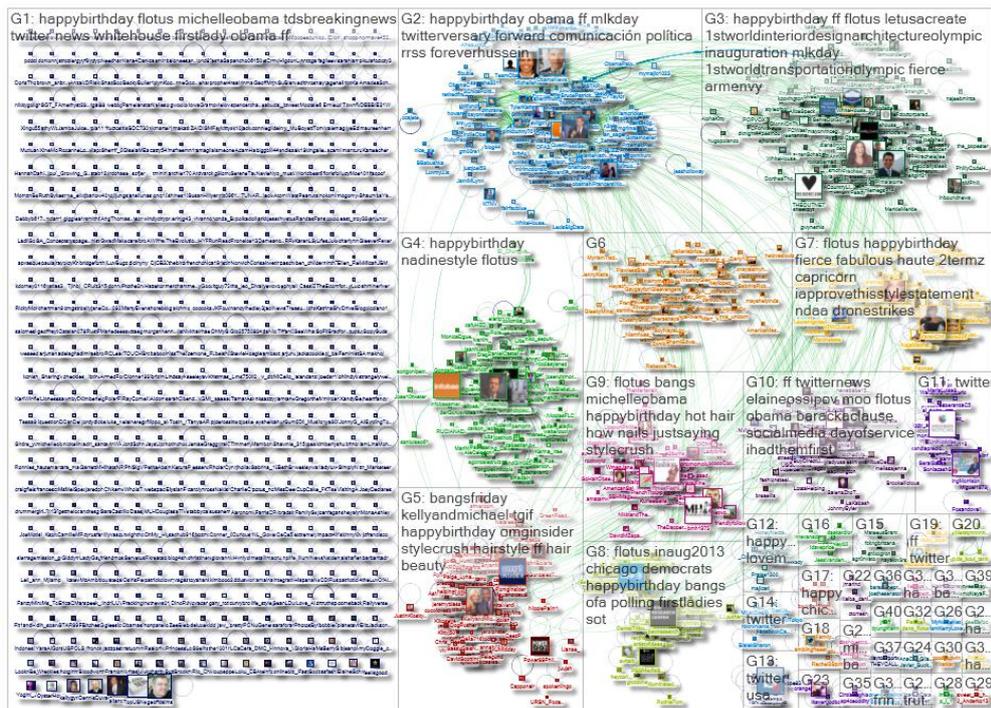


Ilustración 61 Cluster de comunidades (Community Clusters), 30 de agosto de 2016, de <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>

- Red de difusión (Broadcast Networks)

Este tipo de red suele verificarse cuando hay algún nuevo artículo mediático de un medio de prensa y muchas personas lo comparten y lo retuitean. Normalmente hay varias comunidades distintas que interactúan comentando este artículo (y compartiendo su enlace).

La siguiente imagen ha sido elaborada utilizando los tuits de 399 usuarios que contenían la URL a un artículo de Paul Krugman publicado en el New York Times el día 11 de Enero de 2013.

En esta red se puede observar que hay un público de personas que sólo están vinculados a la cuenta de Krugman y comentan directamente (parte izquierda de la imagen).

Al mismo tiempo, en la parte superior e inferior del lado izquierdo de la imagen contiene colecciones densas de personas que podrían ser consideradas partes de diferentes comunidades interesadas en discutir el artículo de Krugman uno con el otro.

En las redes de difusión, hay un uso común de direcciones URL centrales, que al final es el contenido que está impulsando la charla Twitter.



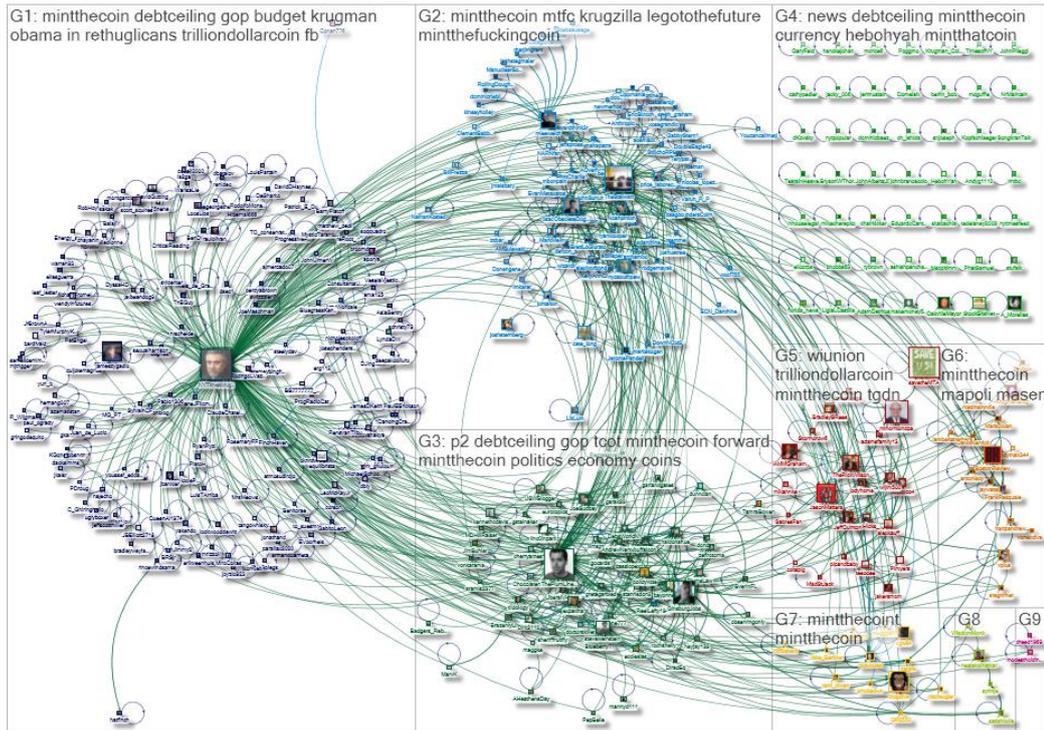


Ilustración 62 Red de difusión (Broadcast Networks), 30 de agosto de 2016, de <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>

- Red de Soporte (Support Network).
Este tipo de grafo se verifica cuando empresas, agencias gubernamentales u organizaciones contestan a las preguntas y las dudas de los clientes. El account de la empresa contesta a varios usuarios que entre si no están conectados, creando círculos exteriores.

En el siguiente ejemplo vemos los tuits de 388 usuarios de Twitter registrados en la semana de 12 a 19 febrero de 2013 contenientes las palabras “DellListens” y “DellCares”.

“DellListens” y “DellCares” son los nombres de dos account de la compañía Dell dedicados al soporte a los usuarios.

De la imagen se puede ver que en la parte izquierda está al centro las cuentas de Dell y tienen muchos rayos, que son los usuarios que están interactuando con ellas. Como dijimos antes, estos usuarios no están conectados entre sí, pues solo se comunican con las cuentas de soporte.

En la parte central del grafo vemos dos grupos. Estos son usuarios interconectados que interactúan uno con el otro y hablan sobre las cuentas de usuario y el soporte en general de Dell.

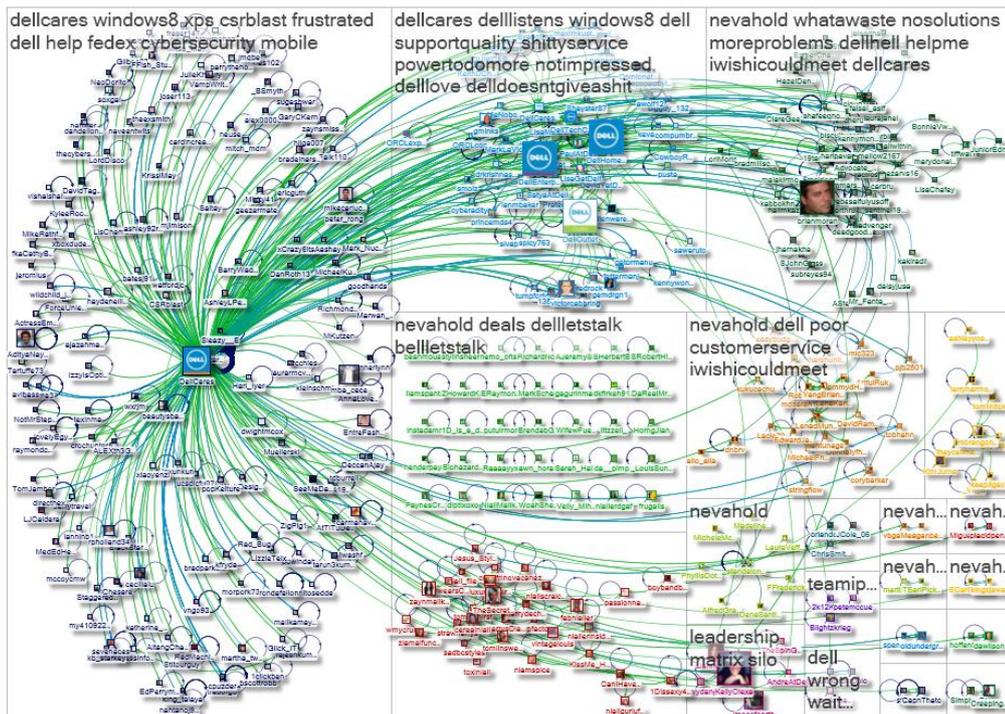


Ilustración 63 Red de Soporte (Support Network), 30 de agosto de 2016, de <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>

5.3 Redes Sociales Multinivel

Hasta ahora hemos hablado de las redes sociales de un solo nivel, compuestas por nodos y aristas. Cada arista era dependiente de las otras, es decir, la existencia de una arista depende de la existencia de las otras aristas que componen la red.

Sin embargo, en un contexto de red multinivel, las aristas de la red son independientes, no solo en el mismo nivel sino también respecto a las aristas de los otros niveles.

En una red social multinivel es posible que exista la relación padre-hijo entre nodos de niveles diferentes.

A continuación se representa una simple red multinivel, en este caso de dos niveles.

En el nivel A, es representada una red con cinco nodos. Esta red podría ser construida a partir de una matriz de adyacencia n por n (en este caso $n = 5$). Lo mismo vale por el nivel B, donde está presente una red cuyos datos provienen de una matriz de adyacencia m por m (en este caso $m = 8$).

Juntando las dos redes con el fin de crear una red multinivel, nos encontramos con una matriz rectangular n por m , donde se representan todas las casuísticas. En esta red podemos ver que las relaciones no solo existen entre nodos que pertenecen al mismo nivel, sin embargo también entre nodos de niveles diferentes.

Por esta razón, a nivel gráfico, se representa también un meso-level X, un nivel intermedio que contiene las relaciones entre las dos redes.

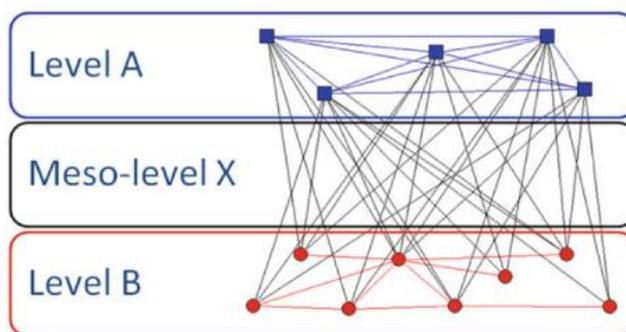


Ilustración 64 Representación de una red multinivel, en este caso de dos niveles

Unos ejemplos de red multinivel podrías ser las relaciones que hay entre investigadores de varios departamentos. Es decir, en un nivel las relaciones entre investigadores del mismo departamento (con su centro, donde aparece quien tiene más contactos etc.) cruzados con la misma red que representa otros departamentos. De esta manera se puede realizar un “zoom” en la red, haciendo que el valor de cada nodo cambie según la red donde se coloca.

Otro ejemplo podría ser la interacción entre empleados de un mismo grupo y la cooperación entre sus unidades organizativas en una institución gubernamental.

5.4 Software para la Análisis de Social Network

En este apartado analizaremos los principales softwares para la Análisis de redes sociales.

Como ya se dijo anteriormente, una red social es una estructura social compuesta por individuos llamados “nodos” que están vinculados por uno o más tipos de interdependencia, como por ejemplo la amistad, los intereses en común etc.

En la vida real, prácticamente cualquier cosa puede ser catalogada como res. Se consideran redes por ejemplo las familias, los equipos de proyecto, las aulas o los equipos. También son redes los sitios web como Twitter o Facebook, gracias a sus usuarios, o incluso el mismo Internet.

Ya que disponemos de muchísimas redes sociales en nuestra vida real, necesitamos un software para poderlas explotar.

Los Software dedicados a la Análisis de redes sociales son programas que facilitan un análisis cualitativo y cuantitativo de las redes sociales mediante la descripción de las características de una red que puede ser a través de datos numéricos o con representación visual.

Las redes se pueden analizar a varios niveles. Por ejemplo, se puedes aplicar las funciones a la red a nivel del individuo (analizando cada nodo), a nivel de arterias (analizando principalmente las relaciones que hay entre los notos) o a nivel global, examinando la red en su totalidad.

Por ejemplo, si analizamos una red a nivel de nodo, podemos estudiar fenómenos como la intermediación y la centralidad de un nodo o atributos individuales de los nodos como la edad, el sexo o los ingresos.

El fin de los softwares de análisis de redes sociales es generar estadísticas analizando estas características a partir de datos de la red. Normalmente los datos se formatean en una lista de adyacencia, o matriz de adyacencia, para poder ser explotados.

En la siguiente imagen podemos ver como un grafo se representa a través de una lista de adyacencia (abajo) o una matriz de adyacencia (en la parte derecha).

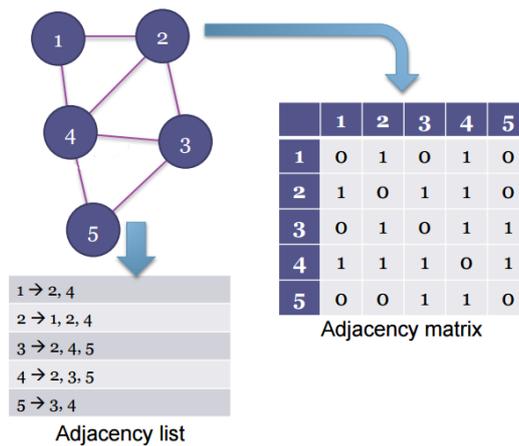


Ilustración 65 Ejemplo de lista de adyacencia y matriz de adyacencia, 01 septiembre de 2016, http://wic.litislab.fr/2010/slides/Combe_WIVE10_slides.pdf

Aunque la mayoría de software de análisis de redes utiliza un formato de datos de texto ASCII, algunos paquetes de software contienen la capacidad de utilizar las bases de datos relacionales para importar y analizar los datos.

Algunos programas de análisis de redes sociales pueden realizar también un análisis predictivo. Se pueden prever cambios en la red analizando fenómenos a nivel de nodo o viceversa, es decir analizando cambios en la totalidad de la red se puede prever el comportamiento de nodos o grupos de nodos.

5.4.1 Funcionalidades principales

En resumen, las funcionalidades que tiene que tener un programa de análisis de redes sociales son las siguientes:

5.4.1.1 Representación

Según Cartwright and Harary la manera más completa y clara de representar las redes es a través de grafos [37].

Como ya dijimos varias veces en este trabajo, los componentes principales de un grafo son los nodos y las aristas. Los elementos de un grafo, para poder realizar correctamente su función, tienen que prever las siguientes casuísticas:

- Orientación de los enlaces entre nodos
Se tiene que poder representar un enlace entre dos nodos de manera indirecta (por ejemplo para indicar la coautoría de dos autores, en la parte izquierda de la imagen) o directa (por ejemplo para representar el flujo de envío de emails, en la parte derecha de la imagen). En la siguiente representación se pueden ver estas dos características.

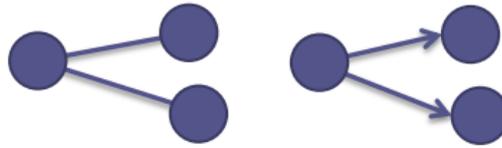


Ilustración 66 Orientación de los enlaces entre nodos

- **Peso en las relaciones**
Se trata de asignar algún valor a las relaciones, para luego poder realizar algún cálculo. En la siguiente imagen se puede ver un simple grafo con unos números que representan el peso que vale la relación entre dos nodos.

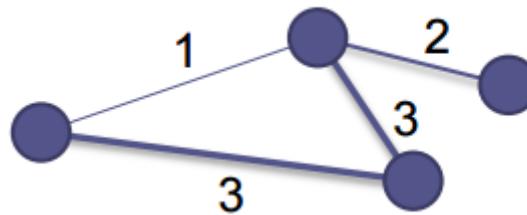


Ilustración 67 Peso en las relaciones

- **Posibilidad de visualizar diferentes tipos de nodos, para poder representar redes modo-2 o superiores**

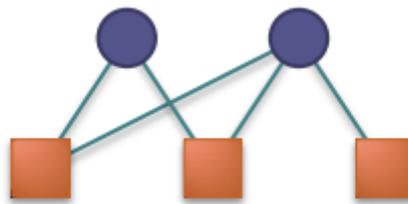


Ilustración 68 Red con diferentes tipos de nodos

- **Representación de la red a través de grafos y de tablas y listas diferentes.**
Se tiene que poder representar un grafo de varias maneras para poder realizar cálculos y manejar los datos. Por ejemplo, se tendría que poder crear una lista de adyacencia o una matriz de adyacencia a partir de un gráfico, y viceversa.

5.4.1.2 Visualización

Uno de los objetivos de un programa de análisis de redes sociales es dar una representación visual del grafo según de diferentes tipologías. Si tenemos un conjunto de datos, se tiene que poder crear grafos diversos según el estudio que queremos realizar, es decir el “protagonista” del grafo podría cambiar.

Por ejemplo, con el mismo dataset podemos focalizar el grafo sobre un actor o enfatizar diversos grupos de tipología de actores. A continuación, un ejemplo de un grafo centrado en un solo actor.

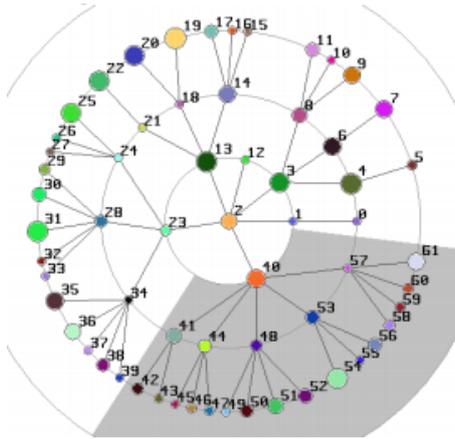


Ilustración 69 Ejemplo de un grafo centrado en un solo actor

5.4.1.3 Uso de indicadores

Es obvio que un programa de análisis de redes sociales debería poder usar indicadores, ya que sin estos no se podría realizar ningún análisis.

Se debe también poder aplicar los indicadores a varios niveles del gráfico, es decir:

- Indicadores globales, a nivel de red (número de nodos, número de relaciones, diámetro etc.)
- Indicadores locales, a nivel de nodos (número de vecinos, número de relaciones en las que participa un nodo etc.)

5.4.1.4 Detección de comunidades

Una de las principales características de las redes sociales es la presencia de comunidades. Es muy común que en un estudio para explotar los datos se intente descubrir las relaciones más comunes entre actores. Por esta razón un programa de análisis de redes sociales debe prever esta posibilidad.

Una comunidad es un grupo de actores que tienen conexiones fuertes entre ellos.

Existen varios algoritmos para detectar las comunidades, los más conocidos son el algoritmo de Newman–Girvan (Newman and Girvan, 2002) y el de Walktrap (Latapy & Pons, 2005).

5.4.2 Programas populares

Normalmente los Software de análisis de red social se dividen en dos categorías: los paquetes basados en interfaces gráficas de usuario (GUI), o los paquetes que utilizan para lenguajes de programación. En general, los paquetes que ofrecen al usuario una interfaz gráfica son mucho más fáciles de aprender, mientras que las herramientas para programar con código se aprenden más lentamente pero son más potentes y extensible.

A continuación se van a listar los programas más conocidos de las dos categorías, para luego focalizarse y realizar un análisis más detallado de unos softwares para cada categoría.

Los programas más conocidos con interfaz gráfico son:

- Ucinet [38]. Programa de pago, Utiliza Pajek y NetDraw para la visualización. Está especializado en el análisis estadístico y matricial. Calcula indicadores y lleva a cabo la agrupación jerárquica
- Pajek [39], un programa para Windows para el análisis y visualización de grandes redes. Es de libre acceso, solo para uso no comercial.
- GUESS [40], un programa dedicado principalmente a la visualización, tiene muchísimos layout diferentes.
- Gephi [41], una plataforma interactiva de visualización y exploración de datos y muxViz (opensource).

A continuación, una captura de pantalla nos enseña la interfaz gráfica del programa Ucinet.

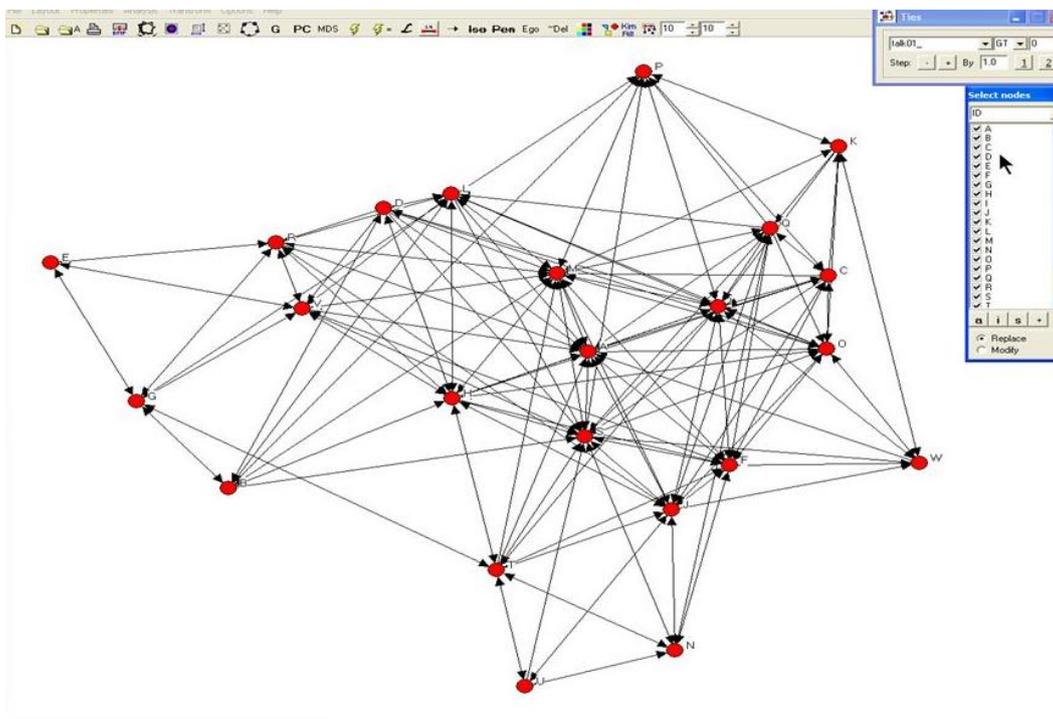


Ilustración 70 captura de pantalla de la interfaz gráfica del programa Ucinet

A nivel empresarial, dirigidos a un público de negocios, los programas de pago con interfaz gráfica para el usuario más conocidos son Orgnet, que proporciona también formación sobre el uso de su software, Polinode, Keyhubs, Keylines, KXEN y Keynetiq.

A continuación, una captura de pantalla del programa propietario Polinode

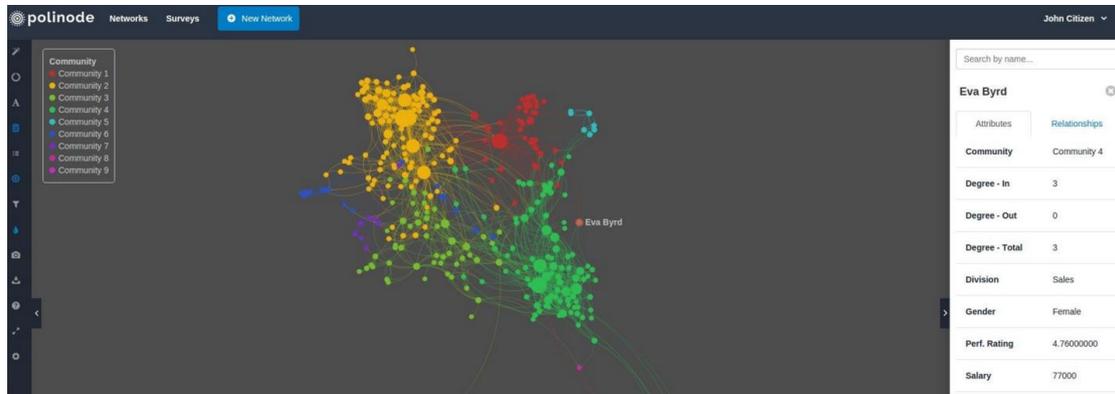


Ilustración 71 captura de pantalla del programa propietario Polinode

Otros programas, como por ejemplo Idir SNA Plus se han desarrollado específicamente para determinadas industrias como las de las telecomunicaciones y el mercado de los juegos online, donde existen conjuntos de datos masivos que necesitan ser analizados.

Como ya dijimos anteriormente, los programas que permiten utilizar lenguajes de programación para explotar los grafos son más difíciles que los que tienen una interfaz de usuario pero mucho más potentes.

Los software más populares con estas características son NetMiner, que utiliza el lenguaje de scripting Python, la suite de paquetes Statnet que permite utilizar el lenguaje de programación estadística R, Igraph, que tiene paquetes para R y Python y muxViz, basado en el lenguaje estadístico y GNU Octave que se utiliza principalmente para el análisis y la visualización de las redes multicapa.

Además, son muy populares la librería para Python NetworkX, un paquete para la creación, la manipulación y estudio de la estructura, la dinámica y las funciones de las redes complejas, y el paquete de SNAP para el análisis de redes a gran escala en C++ y Python.

A continuación, una captura de pantalla del software NetMiner [42]

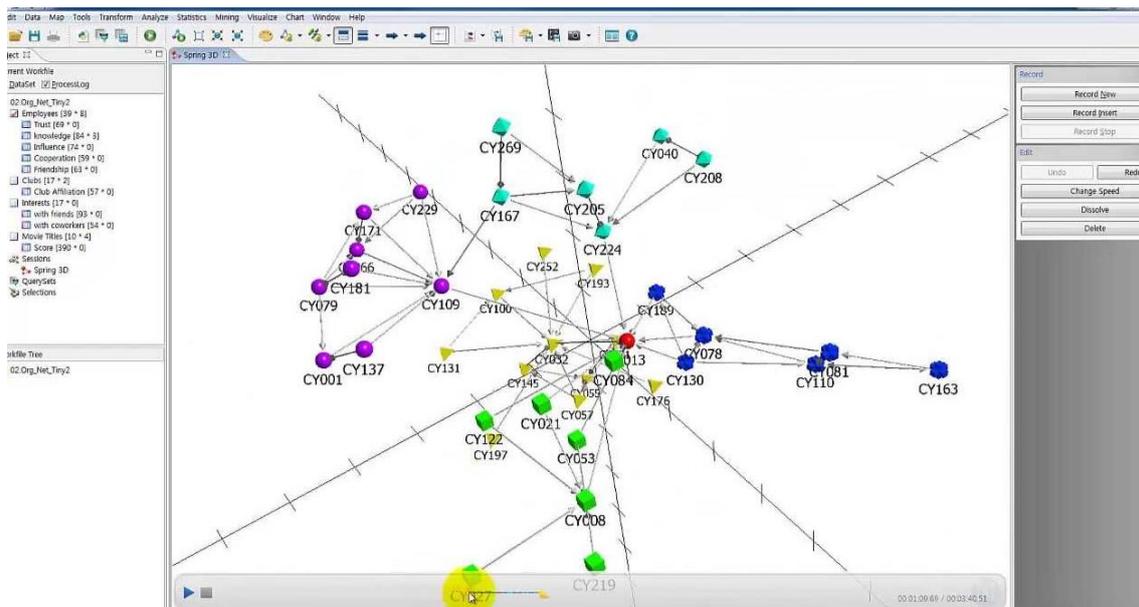


Ilustración 72 Captura de pantalla del software NetMiner

Las representaciones visuales de las redes sociales son importantes para entender los datos de red y transmitir el resultado del análisis. La visualización normalmente también facilita la interpretación cualitativa de datos de la red.

Con respecto a la visualización, herramientas de análisis de redes se utilizan para cambiar el diseño, los colores, el tamaño y otras propiedades de la representación de la red. Todas las herramientas descritas hasta ahora contienen capacidades de visualización.

Según Wikipedia [43], NetMiner, igraph, Cytoscape, muxViz y NetworkX tienen el más alto nivel de funcionalidad en términos de producción de gráficos de alta calidad.

5.4.3 Comparativas

A continuación se van a analizar dos comparativas entre programas independientes (stand-alone, que funcionan sin la necesidad de paquetes o utilidades externas), es decir Pajek y Gephi, y dos librerías, igraph y statnet.

5.4.3.1 Comparativa entre programas stand-alone: Pajek vs Gephi

5.4.3.1.1 Pajek

Pajek (que en esloveno significa araña) es una aplicación de software libre para el análisis y la visualización de grandes redes. Ha sido desarrollado por Vladimir Batagelj y Andrej Mrvar. El objetivo principal para que ha sido diseñado este programa es proporcionar unas potentes herramientas de visualización al usuario, como por ejemplo algoritmos eficientes para el análisis de grandes redes y la abstracción de grandes redes en redes más pequeñas.

Pajek tiene seis estructuras de datos para implementar los algoritmos: la red, la permutación, vectores, grupos, partición y jerarquía.

Mediante el uso de las estructuras de datos anteriormente mencionadas, el software tiene un conjunto de algoritmos que permiten realizar muchas funciones automáticamente.

Esto es muy importante ya que el usuario se ahorra mucho tiempo. A continuación se van a listar los algoritmos más utilizados:

- Particiones: grado, profundidad, núcleo, centros etc.
- Operaciones binarias: unión, intersección, diferencia etc.
- Componentes: fuertes y débiles, Bi conectados, simétrica etc.
- Rutas: Descubrir todas las rutas entre dos vértices, la ruta (s) más corta;
- Los flujos: flujo máximo entre dos vértices etc.
- Pesos de citas;
- Método del Camino Crítico
- Reducción: jerarquía, subdivisión, grado;
- Las simplificaciones y transformaciones: varias líneas, la eliminación de bucles, transformando arcos a los bordes.

Además, el software tiene también una herramienta de dibujo, donde los usuarios pueden editar la red manualmente, seleccionando solo una parte del grafo o por ejemplo seleccionando un vértice editar las relaciones que tiene.

Una vez proporcionados los datos, el diseño de la red puede generarse automáticamente y la imagen puede ser en formato 2D o en D3.

Es orientado sobre todo al data-mining. A continuación se pueden observar unas capturas de pantalla de Pajek.

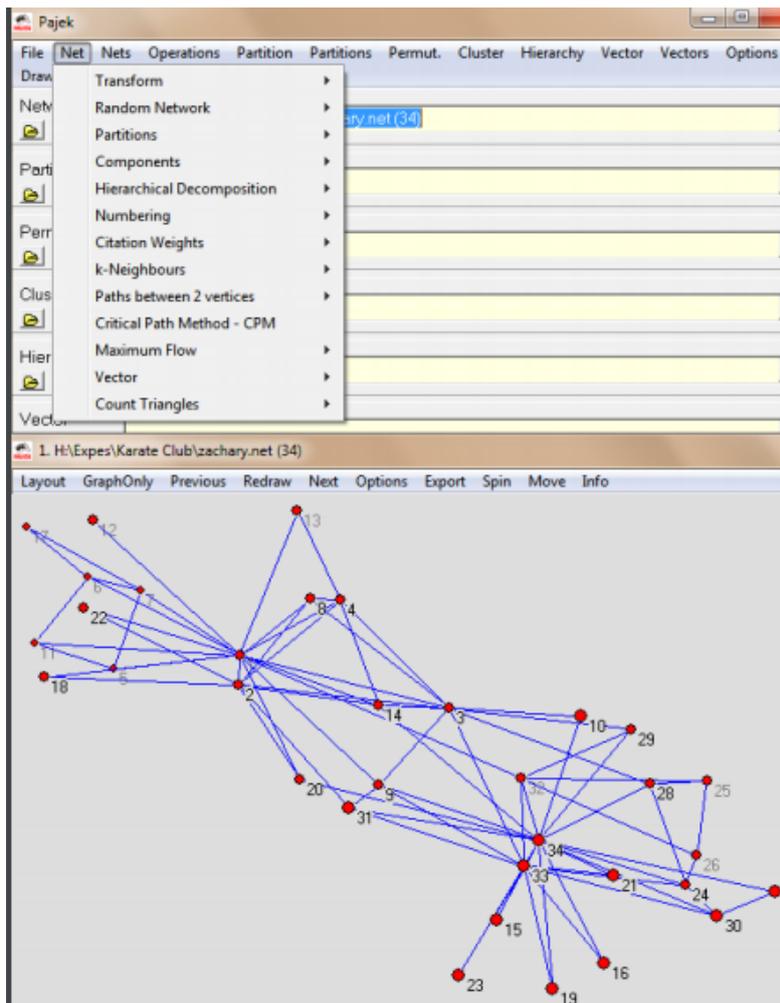


Ilustración 73 Capturas de pantalla de Pajek.

5.4.3.1.2 Gephi

Gephi es una plataforma open source dedicada a la visualización y la exploración interactiva. Se puede utilizar para todo tipo de redes, gráficos dinámicos y gráficos jerárquicos. El software se puede utilizar en Windows, Linux y Mac OS X.

Los objetivos de esta aplicación son apoyar a los analistas en hacer hipótesis, descubrir patrones, y descubrir fallos en las redes. Gephi pueden manejar redes de hasta 50.000 1.000.000 de bordes. Se dispone de algoritmos que permiten crear las gráficas dándole forma y además permite al usuario realizar ajustes para poder cambiar el diseño durante la ejecución.

Gephi proporciona dos tipos de algoritmos: Algoritmos force-based y algoritmos multinivel. Los layout ofrecen las más conocida y comunes métricas para el análisis de redes sociales, como por ejemplo intermediación, la cercanía, el diámetro, coeficiente de agrupamiento, el camino medio más corto, el PageRank, la detección de grupos y comunidades etc.

El usuario puede personalizar el color, el tamaño y las etiquetas para una mejor representación de la red. Gephi proporciona características con el fin de explorar grandes redes creadas con datos reales, como por ejemplo gráficos estructurados, comunidades sociales, vías bioquímicas o de gráficos de redes de tráfico.



El software puede también filtrar estructuras dinámicas, tales como redes sociales utilizando la función “línea temporal”.

Gephi utiliza Java, soporta muchos formatos de ficheros para poder importar los datos para las redes y, como ya dijimos, tiene una interfaz gráfica interactiva.

A continuación se puede observar una captura de pantalla de Gephi.

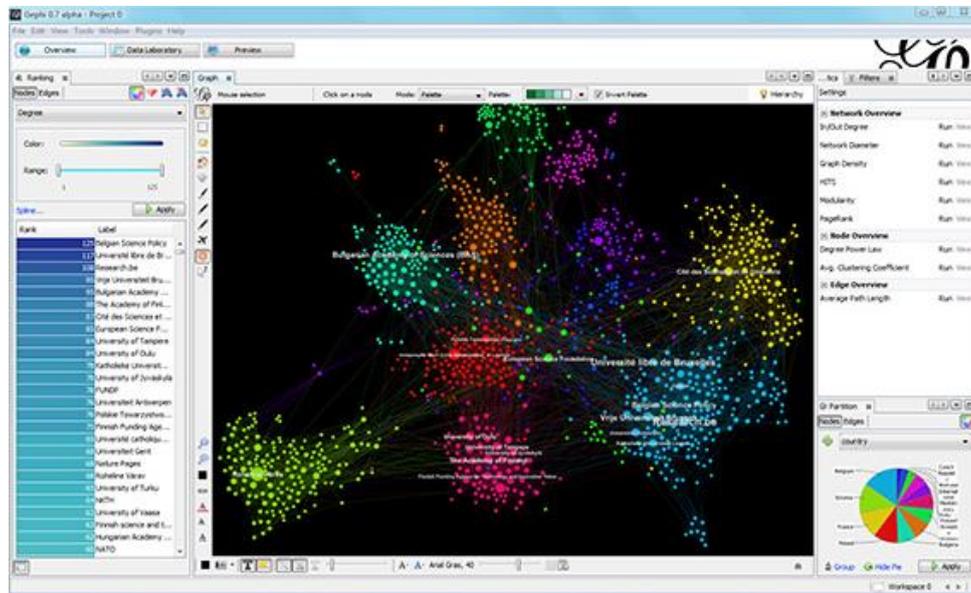


Ilustración 74 Captura de pantalla de Gephi.

5.4.3.2 Comparativa entre librerías, igraph vs Statnet vs D3.js

Igraph vs Statnet son dos librerías que permiten llevar a cabo el análisis de redes sociales utilizando el popular lenguaje de programación estadístico R. D3.js es una librería de Javascript utilizada principalmente para crear gráficos y visualizaciones interactivas en páginas web.

A continuación se van a analizar separadamente estas tres librerías.

5.4.3.2.1 Igraph

Igraph [44] es un paquete que permite realizar análisis de redes en el lenguaje y entorno de programación para análisis estadístico y gráfico R, además que en Python, C o C++.

El paquete proporciona rutinas y funciones para crear y manipular grafos con facilidad.

Uno de los puntos fuerte de igraph es que es de software libre, así que además de no pagar para utilizarlo, permite visualizar su código y personalizarlo.

La forma para pasar los datos a igraph es por matriz de adyacencia por parejas.

A continuación se visualiza una simple matriz de adyacencia.

| | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 | 0 |
| 3 | 1 | 1 | 1 | 0 | 0 |
| 4 | 1 | 1 | 0 | 0 | 0 |

Ilustración 75 Ejemplo de matriz de adyacencia

Los datos de esta matriz de deben pasar a igraph de la siguiente manera:

```
G <-GRAPH(C(0,1,0,2,2,0,2,1,3,0,3,1,3,2,4,0,4,1), N=4)
PNG("PRUEBA.PNG")
PLOT(G,MAIN="TITULO")
DEV.OFF()
```

Y ejecutando el programa, el gráfico resultante será el siguiente:

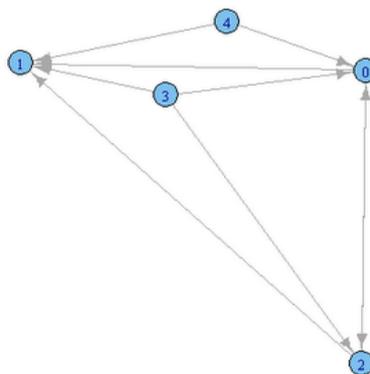


Ilustración 76 Gráfico realizado con la librería igraph

5.4.3.2.2 Statnet

Statnet [45] es una suite de paquete que ha sido desarrollada por un team de la Universidad de Washington.

El propósito principal por que ha sido desarrollada esta librería es el modelado estadístico. Sin embargo, en la librería se pueden encontrar unos paquetes dedicados al análisis de redes sociales, donde se incluyen funciones básicas que han permitido que esta librería sea muy popular.

Comparado con otros paquetes como UCINET o PAJEK, Statnet tiene un propósito diferente: La atención de Statnet se centra en el modelado estadístico de las redes de Datos.

Las capacidades de modelado estadístico de Statnet incluyen ERGMs (Exponential random graph models, que es una familia de modelos estadísticos utilizados para analizar datos de redes sociales u otro tipo de redes, modelos de espacio latente y modelos de cluster latente.

El paquete está escrito utilizando una combinación de R, uno de los lenguajes estadísticos Open Source más populares y C, y se ejecuta a través de la línea de comando de R.

Para poder ejecutar statnet y utilizar todas sus funcionalidades, se tiene que utilizar una interfaz de línea de comando, ya que no es presente una interfaz gráfica.

A continuación se visualizará un gráfico realizado con la librería Statnet, utilizando un dataset de ejemplo de los que va incluido con el paquete.

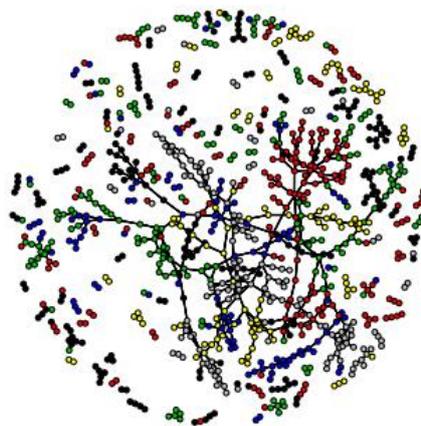


Ilustración 77 Gráfico realizado con la librería Statnet

5.4.3.3 D3.js

En este apartado se explicará que es D3.js, la librería con la cual se realizará la parte práctica de este trabajo.

D3.js [46] (Data-Driven Documents) o solamente D3, es una librería de Javascript utilizada principalmente para crear gráficos y visualizaciones interactivas en páginas web. La primera versión de esta librería fue distribuida en agosto de 2011 y ha sido desarrollada por Mike Bostock, Jeffrey Heer, Vadim Ogievetsky y el auxilio de la comunidad. Su principal función es la creación de visualizaciones complejas y gráficos interactivos.

Las componentes principales que esta herramienta utiliza son gráficos vectoriales (SVG), Javascript, HTML5 y CSS3. Además, D3.js permite usar grandes cantidades de datos en distintos formatos: JSON, CSV, GeoJSON, etc. Hay que destacar también que la librería es multiplataforma.

La librería D3.js utiliza funciones JavaScript predefinidas y el resultado se puede visualizar en una página HTML.

Gracias a estas funciones, se pueden fácilmente seleccionar grandes conjuntos de datos, a los cuales se le pueden dar diferentes estilos (utilizando herramientas propias de la librería o los CSS), además se pueden crear objetos SVG y añadirles por ejemplo transiciones, efectos dinámicos agregarle información.

El sitio web Data-Driven Journalism, en su artículo "Data-Driven Documents, Defined [47]" describe a D3.js como "un pedazo mágico de JavaScript que permite expresar tus datos visualmente en la web". En el mismo artículo se subraya que la librería D3 está explotando en popularidad en este momento, en parte porque funciona con cualquier navegador moderno, dispositivos móviles Android y iOS incluidos, pero también porque es muy potente.

Por esta razón no es muy basilar aprender a utilizar y se necesita de un tiempo para estudiarla y entender cómo funciona. Una vez que se sepan explotar todas sus funcionalidades, la librería D3 se convierte en una de las mejores herramientas para la visualización complejas de datos y sobre todo el diseño interactivo.

La posibilidad de crear gráfico interactivos es prácticamente la mejor ventaja de D3, ya que se convierte en una potente herramienta para los artículos periodísticos online ya que con una imagen se pueden visualizar muchísimos datos y se permite al usuario navegar en la visualización y focalizarse en lo que le interesa.

Gracias a las miles de visualizaciones que permite crear esta librería, se han empezado a utilizar nuevos métodos de narración visual. Por ejemplo, uno de los periódicos más conocidos del mundo, el New York Times, en la última década ha ido eliminando gradualmente Flash a favor de D3 hasta el punto de contratar en su equipo gráfico a Mike Bostok [48], que es el autor primario de la librería D3, aunque el mismo en mayo de 2016 dejó el New York Time ya que este trabajo le ralentizaba demasiado el desarrollo de sus proyectos Open Source [49].



5.4.3.3.1 Ejemplo de visualización: Paths to the White House

Una de las gráficas más populares creada con la librería D3.js es la gráfica “Paths to the White House [50]”, que se publicó en el New York Times en 2012, en la época de las elecciones para el presidente de los Estados Unidos de América entre Barak Obama y Mitt Romney.

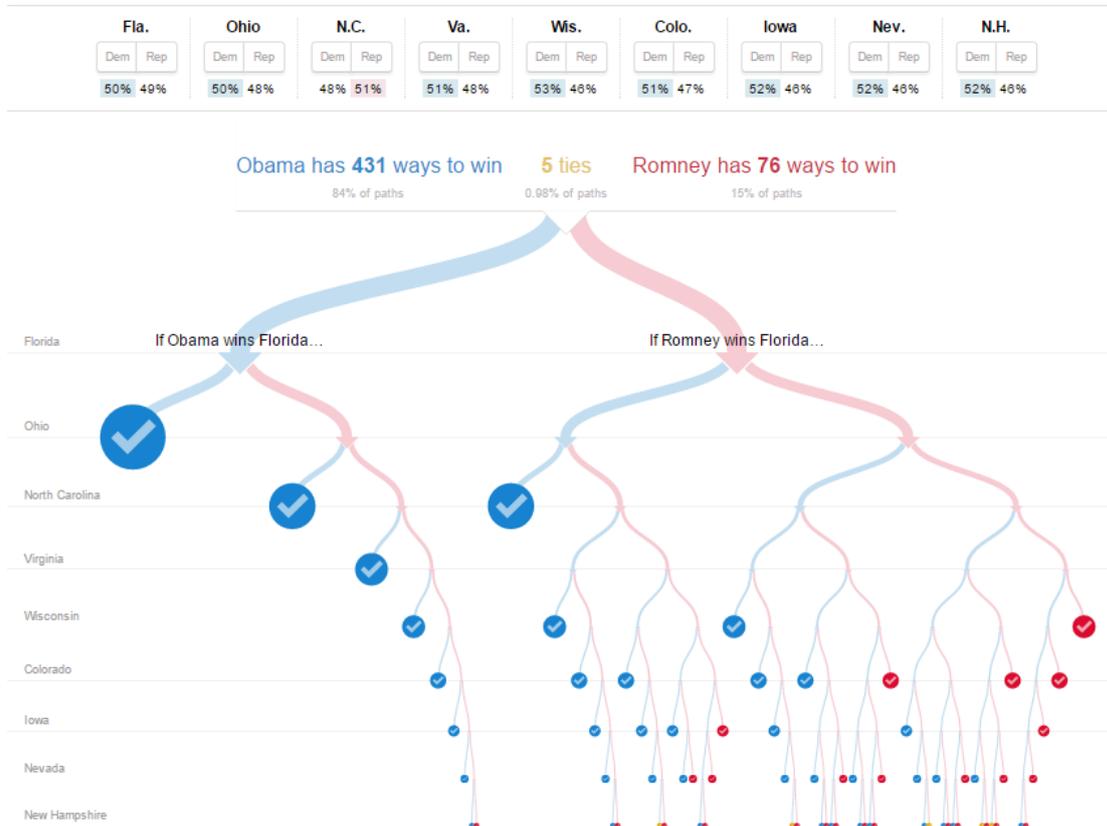


Ilustración 78 Visualización Paths to the White House, 30 de agosto de 2016, de <http://elections.nytimes.com/2012/results/president/scenarios>

El fin de esta gráfica era permitir al usuario, de manera fácil y atractiva, de poder “jugar” realizando hipótesis sobre quien ganara las elecciones en cada estado y viendo quien será el presidente.

En la parte superior se puede notar una simple lista horizontal con los principales estados, donde se permite al usuario de elegir el candidato que piensa que gane.



Ilustración 79 Elección del candidato que se piensa que gane en la visualización Paths to the White House, 30 de agosto de 2016, de <http://elections.nytimes.com/2012/results/president/scenarios>



Poniendo el resultado en todos los 9 estados, la aplicación calcula quien es el ganador y visualiza su cara en la parte principal.



Ilustración 80 Resultado de la elección del candidato que se piensa que gane en la visualización Paths to the White House, 30 de agosto de 2016, de <http://elections.nytimes.com/2012/results/president/scenarios>

Sin embargo, a la vez que se empieza a poner el resultado en algún estado, en la parte central el gráfico se mueve según las probabilidades de ganar de los dos candidatos dependiendo de los demás estados.

En el siguiente gráfico el ejemplo resultante de elegir que los democráticos ganaran en Virginia y los Republicanos en Florida y North Carolina.

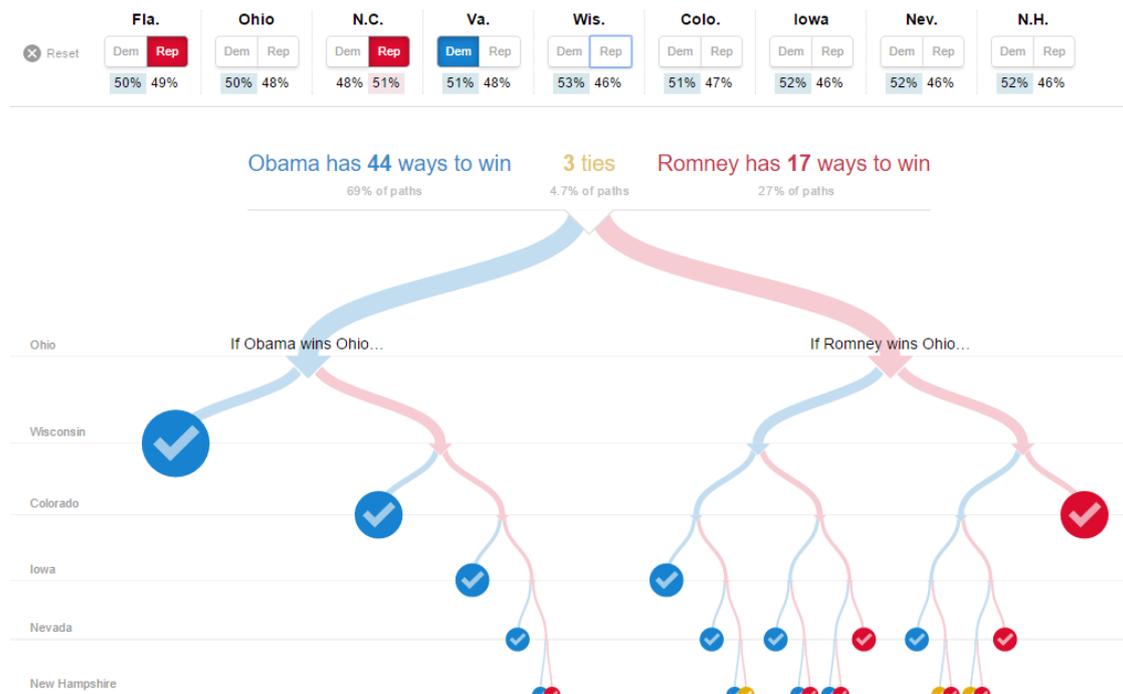


Ilustración 81 Ejemplo resultante de elegir que los democráticos ganaran en Virginia y los Republicanos en Florida y North Carolina en la visualización Paths to the White House, 30 de agosto de 2016, de <http://elections.nytimes.com/2012/results/president/scenarios>

En el ejemplo anterior se puede notar que la visualización sí que es compleja y presenta muchísimos datos y casuísticas, pero a la vista del lector parece muy fácil de utilizar y le permite comprobar todas las opciones.

Los datos pueden tender formatos variados, como dijimos antes los más utilizados son generalmente JSON, CSV o geoJSON, aunque es posible utilizar otros formatos de datos escribiendo propias funciones JavaScript.

En resumen, la librería permite manipular documentos que contienen datos brutos utilizando estándares abiertos de la web y los navegadores pueden crear visualizaciones complejas sin depender de ningún software propietario. La librería D3 es un claro reemplazo de Flash, con la diferencia principal de que sus desarrollos quedan abiertos y pueden ser actualizados o mejorados por otros desarrolladores.

A continuación se van a enseñar unos ejemplos creados con esta librería, ya que en la parte práctica de este trabajo se analizará su aspecto técnico y como funciona.

5.4.3.3.2 Ejemplo de visualización: El tamaño de la industria manufacturera china

En la siguiente gráfica, se quiere representar que China sigue dominando en el mercado de las industrias manufactureras, pero poco a poco las industrias empiezan a buscar en otros países.

El gráfico se refiere a un artículo del New York Times de 2013 [51] y sus conclusiones son que China sigue dominando en este mercado, pero algunos fabricantes empiezan a buscar en otros países. En detalle, las empresas multinacionales están buscando maneras de limitar la dependencia de sus fábricas que tienen en China.

Se visualiza un mapa geográfico donde el tamaño de los países se distorsiona de manera que cada país está dimensionado de acuerdo con su producción económica en 2012.

Los países están coloreadas por su tasa de crecimiento. Se puede notar que las economías más establecidas tienden a crecer más lentamente.

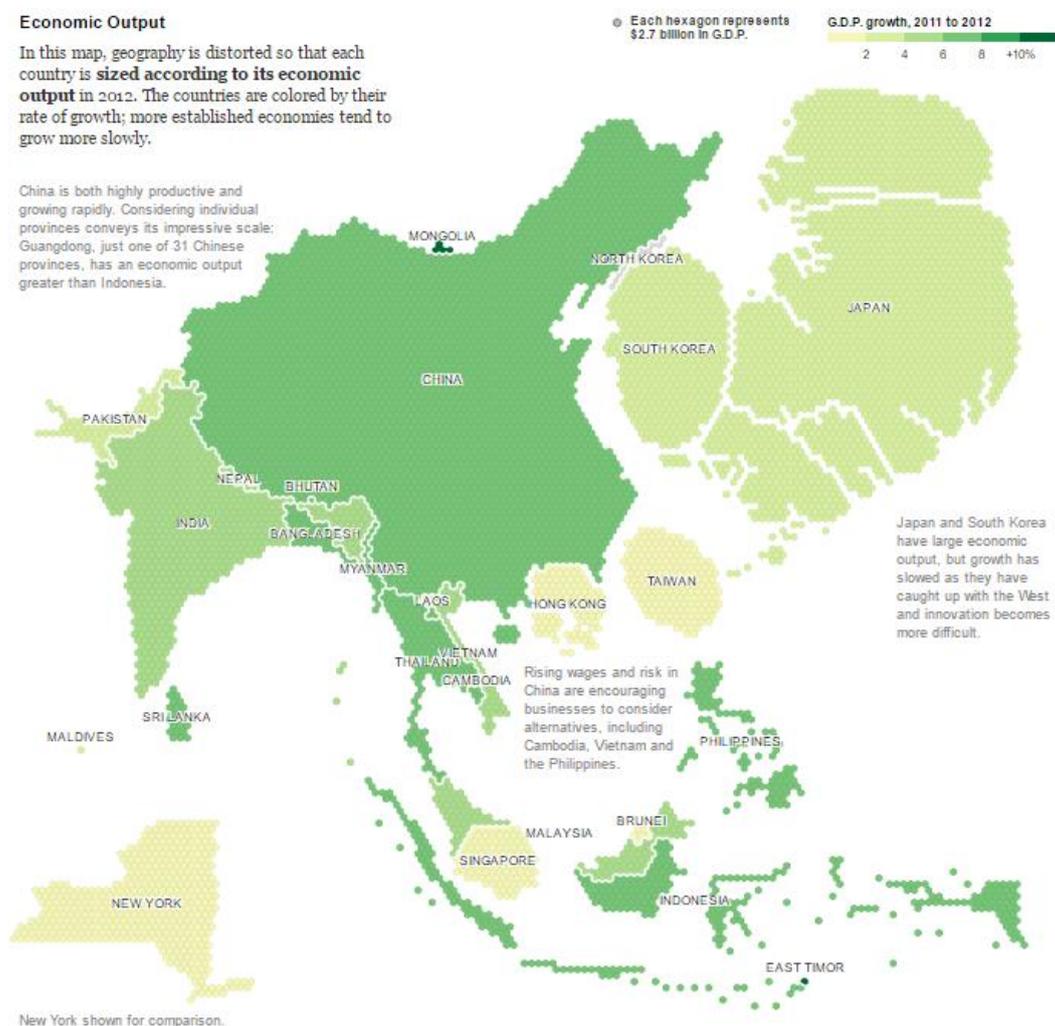


Ilustración 82 Visualización del tamaño de la industria manufacturera china, 30 de agosto de 2016, de http://www.nytimes.com/interactive/2013/04/08/business/global/asia-map.html?_r=0

En detalle, en el gráfico los países son coloreados según su tasa de crecimientos del año 2011 a 2012 y la silueta de los países es compuesta de hexágonos, cada uno que representa 2.7\$ de Producto Interno Bruto.

En la parte inferior del gráfico se ha analizado también New York.

Pasando el ratón sobre cada estado, aparece una tabla donde se indica el Producto Interno Bruto del país, el porcentaje de cambio respecto al 2011, el Producto Interno Bruto Bruto per cápita y la Población del país.



Ilustración 83 Selección de un país en la Visualización del tamaño de la industria manufacturera china, 30 de agosto de 2016, de http://www.nytimes.com/interactive/2013/04/08/business/global/asia-map.html?_r=0

5.4.3.3 Ejemplo de visualización: Conexiones entre los nominados a los Oscar de 2013

En el siguiente grafo se quieren visualizar las conexiones entre Actores, Productores y Directores que han sido nominados a los Oscar de 2013.

El gráfico ha salido en un artículo de Febrero de 2013 del New York Times [52].

Published: February 20, 2013

Among the Oscar Contenders, a Host of Connections

With few exceptions, this year's nominated actors, directors and producers have long worked on films with Oscar histories.

FACEBOOK TWITTER EMAIL SHARE

THE INCUMBENTS

The five major nominees for "Lincoln" have worked on at least 70 films with nominations. Kathleen Kennedy, a producer, has worked closely with Steven Spielberg since he gave Kennedy her first producer role, on "E.T. the Extra-Terrestrial."

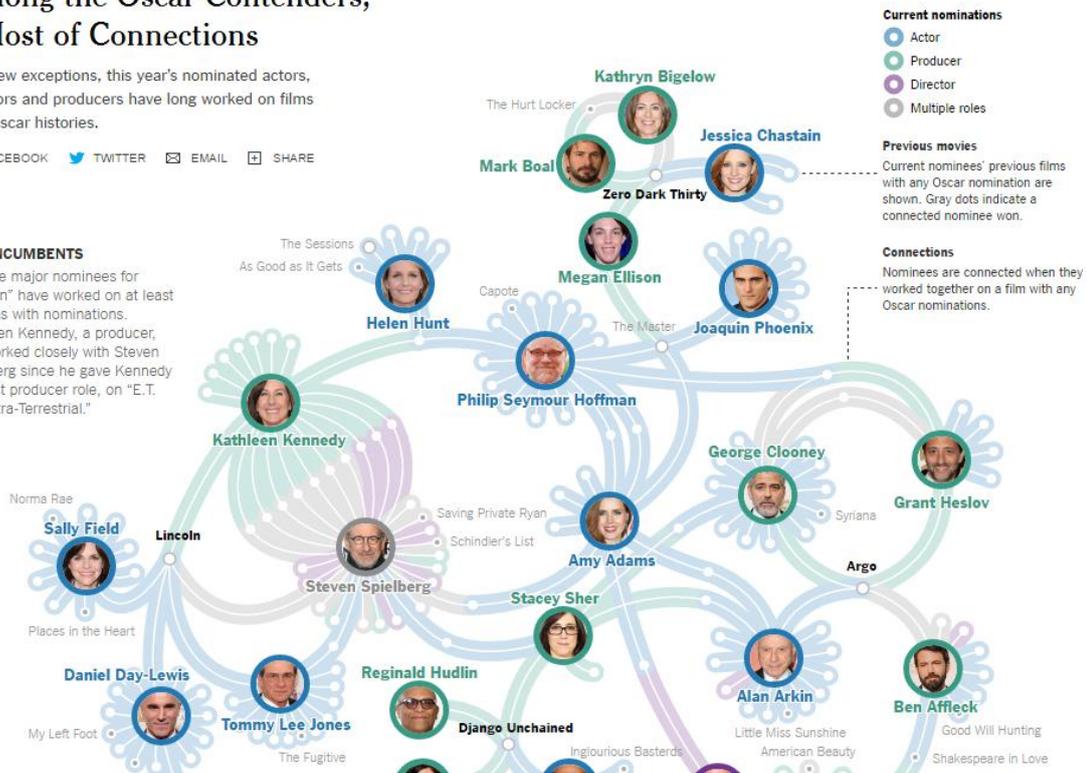


Ilustración 84 Visualización de las conexiones entre los nominados a los Oscar de 2013, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/02/20/movies/among-the-oscar-contenders-a-host-of-connections.html>

En el gráfico se pueden ver las personas (actores, productores o Directores) representados por su cara, y las películas, representadas solamente por su título. Pasando el ratón en el gráfico ocurren diferentes acciones:

- Pasando sobre la cara de una persona, se evidencian sus conexiones. En la siguiente imagen, el caso de Steven Spielberg

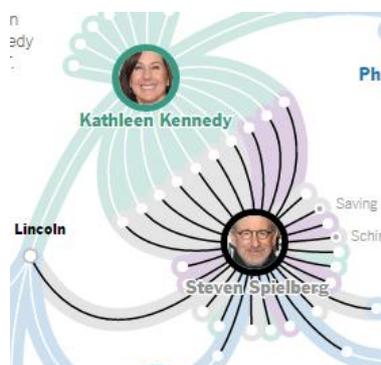


Ilustración 85 Selección de una persona en la Visualización de las conexiones entre los nominados a los Oscar de 2013, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/02/20/movies/among-the-oscar-contenders-a-host-of-connections.html>



- Pasando el ratón sobre el título de una película, se subrayan los miembros y sale una ventana con la lista de las personas que han participado. En la siguiente imagen se ha seleccionado la película “The Master”

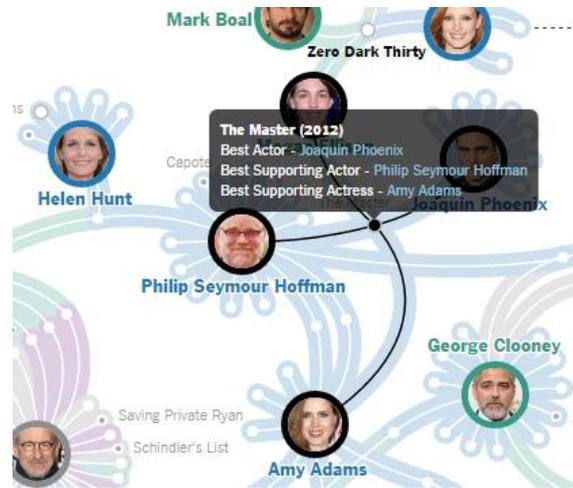


Ilustración 86 Pasando el ratón en una película en la Visualización de las conexiones entre los nominados a los Oscar de 2013, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/02/20/movies/among-the-oscar-contenders-a-host-of-connections.html>

5.4.3.3.4 Ejemplo de visualización: Comparación de las tasas efectivas tributadas entre las empresas del índice “500 S&P”

En la siguiente visualización se pretende mostrar los impuestos que tributan las empresas del 500 S&P, que son las 500 empresas de mayor tamaño de los Estados Unidos que cotizan en las bolsas de los Estados Unidos en el índice S&P 500. Es el índice de referencia que utiliza la empresa Standard & Poor's, una de las tres agencias más importantes de calificación de crédito [53].

Esta gráfica se creó para comparar el tamaño de las empresas con el porcentaje de impuestos que pagan, después de que se vio que grandes empresas como Apple tributan muy poco comparado con medianas empresas. Los datos han sido tomados desde el 2007 hasta el 2012 por Standard & Poor's.

Esta gráfica se creó para comparar el tamaño de las empresas con el porcentaje de impuestos que pagan, después de que se vio que grandes empresas como Apple tributan muy poco comparado con medianas empresas. Los datos han sido tomados desde el 2007 hasta el 2012. Ha sido publicada por el New York Times en un artículo del título “Across U.S. Companies, Tax Rates Vary Greatly” [54].

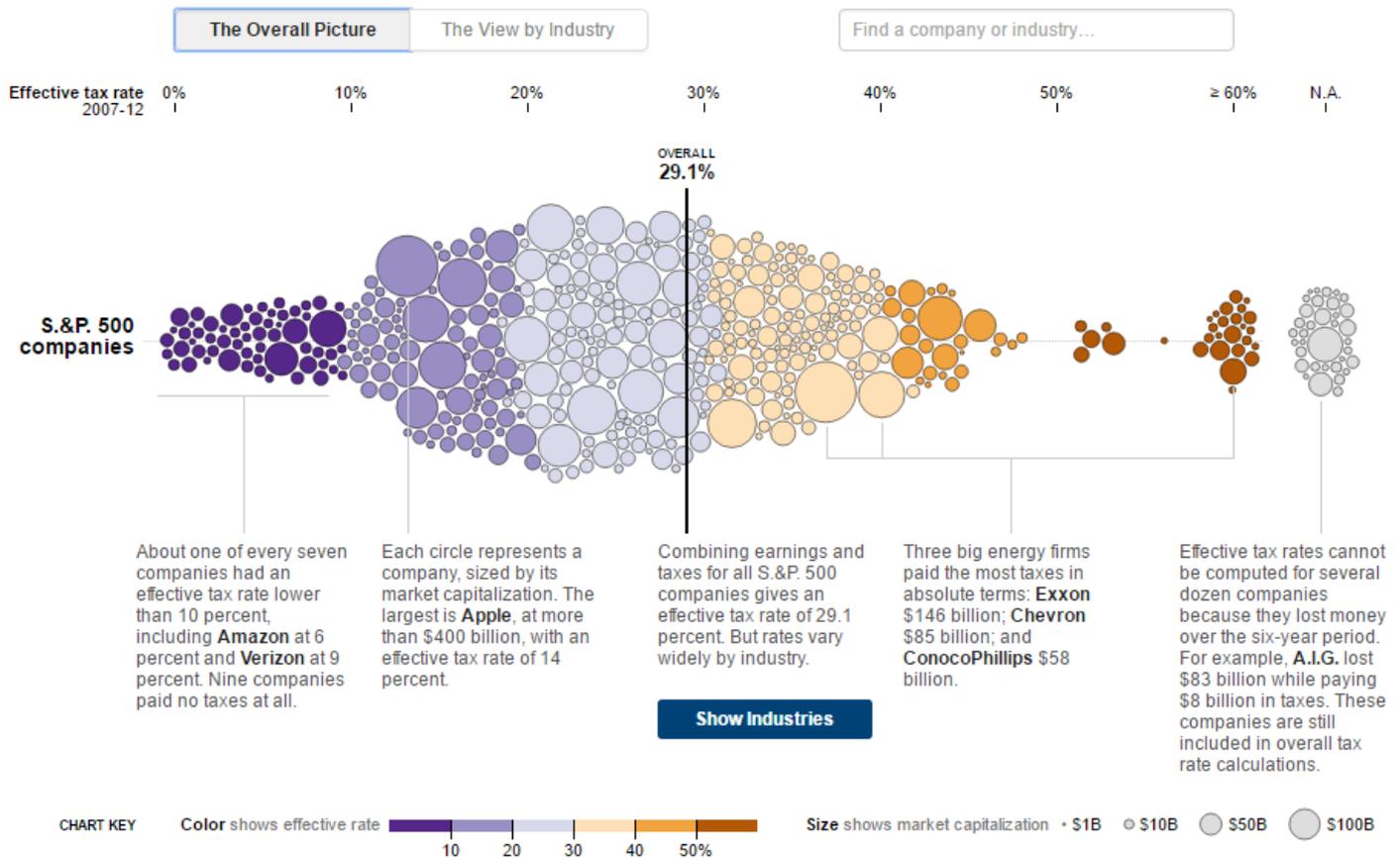


Ilustración 87 Visualización de las empresas del 500 S&P, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html>.

En la gráfica se visualizan las empresas de formas de círculos. El tamaño de cada círculo es directamente proporcional al capital de la empresa. Además, en la parte inferior hay una leyenda que enseña el tamaño del círculo de una empresa con un capital de 1B\$, 10B\$, 50B\$ y 100B\$.



Ilustración 88 Leyenda del tamaño de los círculos en la visualización sobre las empresas del 500 S&P

El color de cada círculo representa el porcentaje de tasas efectivas que tributa la empresa. Como en el caso anterior, en la parte inferior de la imagen se ha puesto una leyenda con los colores y sus correspondientes porcentajes.



Ilustración 89 Leyenda de los colores en la visualización sobre las empresas del 500 S&P, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html>.

De esta manera el usuario a primera vista puede ver varias tendencias, por ejemplo si las empresas que tributan poco suelen ser grandes o pequeñas.

En el medio del gráfico se represa también una línea vertical que indica la media del porcentaje de tasas efectivas pagadas, que nos permite ver como es la distribución de la

empresa. En este caso por ejemplo podemos notar que la mayoría de empresas están debajo de la media.

Pasando el ratón encima de cada círculo, sale una ventana que nos representa, además del nombre de la empresa, el porcentaje efectivo de las tasas que tributa la empresa, el importe de las tasas pagado por la empresa y las ganancias de dicha empresa. Esta funcionalidad es muy cómoda, ya que el usuario desplazando el ratón puede visualizar el porcentaje de tasas pagado por varias empresas en muy poco tiempo.

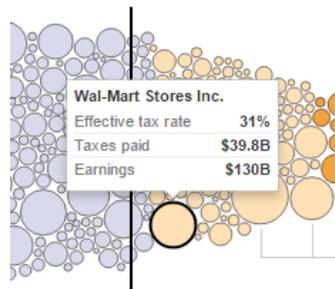


Ilustración 90 Zoom sobre una empresa en la visualización sobre las empresas del 500 S&P, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html>.

Por defecto, viene mostrada la versión general (The Overall Picture) de la visualización, que enseña todas las 500 empresas en el mismo gráfico. Sin embargo, haciendo click en el botón “The View by Industry”, la imagen cambia con una animación y las 500 empresas se dividen en categoría según su sector de pertenencia.

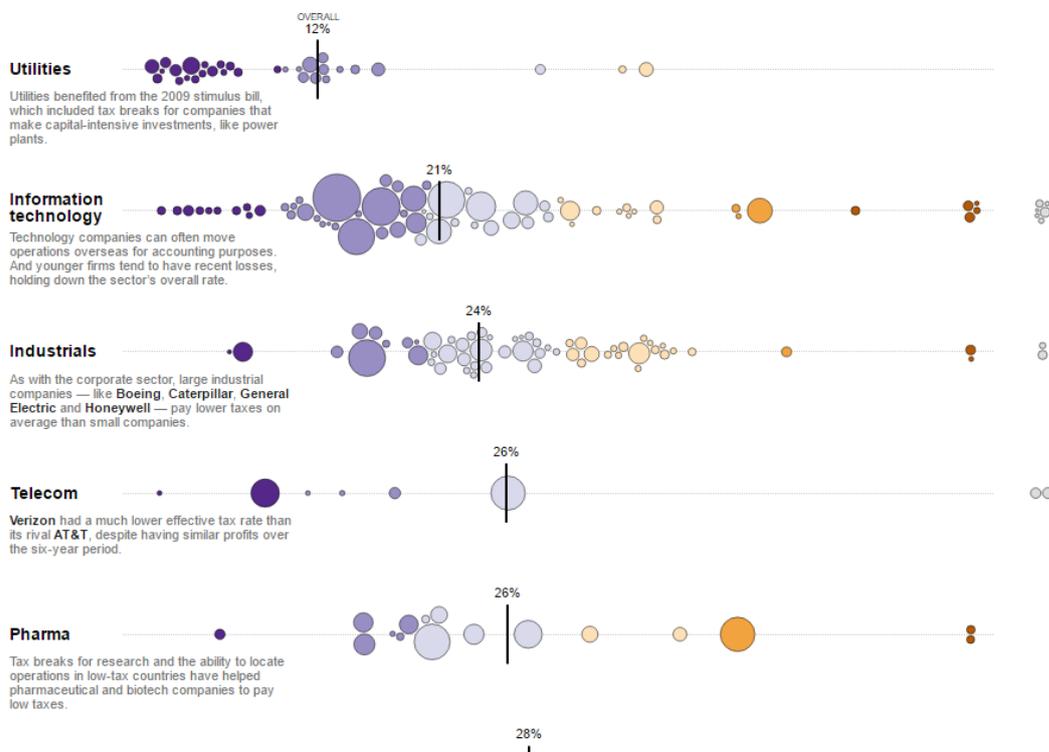


Ilustración 91 Visualización por sectores en la visualización sobre las empresas del 500 S&P, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html>.

Para cada categoría se visualiza su media y un comentario sobre el andamiaje del grupo de empresas.

Por ejemplo, en las empresa de Telecomunicaciones, se nos dice que a pesar de tener prácticamente los mismos beneficios en el periodo de tiempo que se está analizando, Verizon paga un porcentaje de tasas efectivas muy inferior respecto a AT&T, su rival directo.

Otra función muy interesante de esta gráfica es la búsqueda, que está ubicada en la parte superior derecha. Mientras vamos escribiendo el nombre de la empresa que vamos a buscar, la visualización resalta las empresas que contienen la palabra buscada.

Por ejemplo, si buscamos la palabra “Wal”, se evidencian las empresas que contienen en su nombre estas letras, es decir “The Walt Disney Company”, “Walgreen Co.” y “Wal-Mart Stores Inc.”. La búsqueda se realiza mientras se escribe la palabra, sin la necesidad de pulsar la tecla “Enter”.

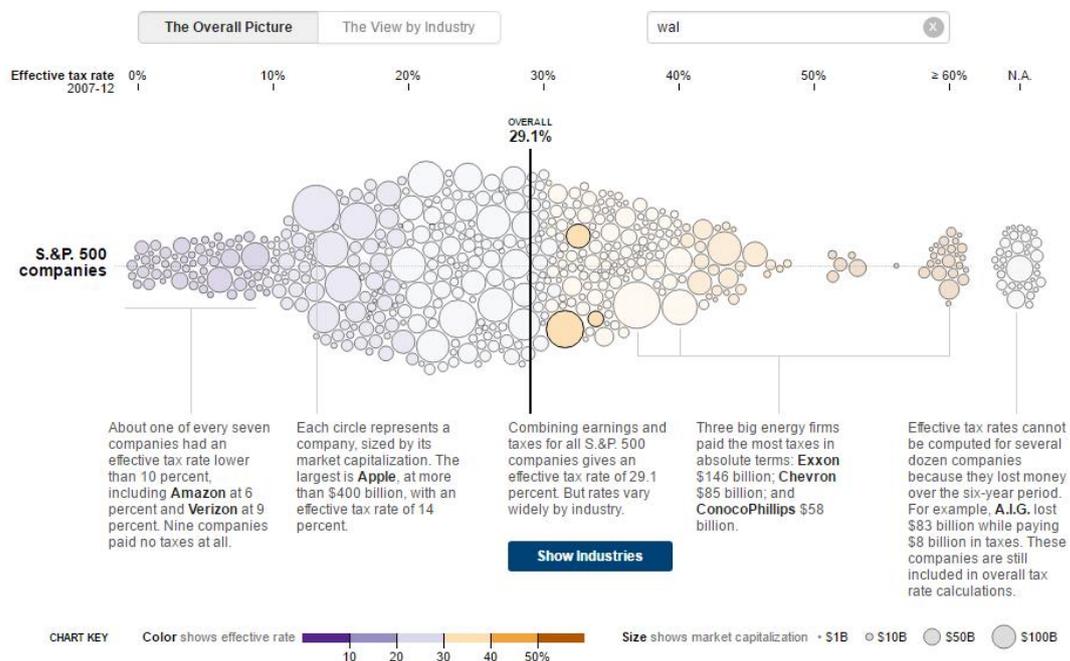


Ilustración 92 Ejemplo de búsqueda en la visualización sobre las empresas del 500 S&P, 30 de agosto de 2016, de <http://www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html>.

Este es un claro ejemplo de una visualización interactiva, ya que se permite al usuario de centrarse en lo que necesita, cambiando la visualización según sus requerimientos.

6 Aplicación de visualización de datos

En este apartado se va a aplicar todo lo que se ha analizado en la parte anterior generando unos ejemplos y comentando como se han creado.

Se van a explicar las problemáticas encontradas, lo que se pretendía hacer y lo que se ha conseguido desarrollar.

Para realizar estos ejemplos se ha utilizado la librería D3plus [55], una extensión de la conocida librería D3.js ya analizada precedentemente. La librería ha sido creada como “simplificación” de D3.js por un equipo del MIT (Massachusetts Institute of Technology).

En su página web, D3plus.org, se puede encontrar la guía de esta librería y varios ejemplos.

La librería dispone de una serie de visualizaciones interactivas, como diagramas de dispersión, gráfica de árboles, gráfica de barras, etc.

En la siguiente imagen, una captura de pantalla del apartado “Ejemplos” de la página web D3plus.org.

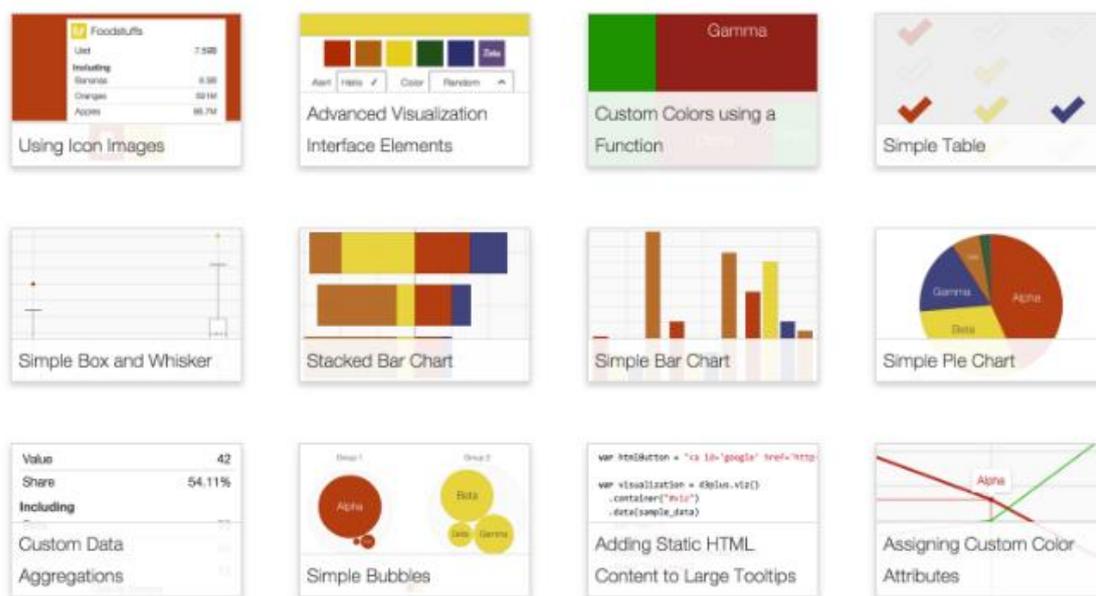


Ilustración 93 Captura de pantalla del apartado “Ejemplos” de la página web D3plus.org, 30 de agosto de 2016, de <http://d3plus.org/examples/>

6.1 Ejemplo 1: Tree Map con Timeline

En este ejemplo se pretende utilizar el tipo de gráfico “tree_map” que ofrece la librería D3plus.

6.1.1 El gráfico Tree Map de la librería D3plus

Los tree maps (Mapas de árboles) son una manera de mostrar un conjunto de datos en su totalidad, pudiendo comparar cada dato con los demás.

A diferencia de las Pie Charts (Gráfico de tarta), que tienen una gran cantidad de espacio en blanco no utilizados alrededor de los bordes de la visualización, los tree maps permiten que la visualización ocupe todo el espacio disponible.

En detalle, el gráfico está compuesto por cuadrados de diferentes colores, y su tamaño está en relación con el valor del dato. Los cuadrados están ordenados desde la esquina superior a la izquierda (el dato más alto) hasta la esquina inferior a la derecha (el dato con menos valor).

Dentro de los cuadrados, es posible visualizar un texto.

Además de eso, la librería ofrece la posibilidad de asociar a cada cuadrado un tooltip, es decir una ventanita que aparece cuando se pasa con el ratón encima del gráfico, donde se pueden visualizar textos con informaciones.

En la siguiente imagen se visualizará un ejemplo del tooltip en un Tree map.

Se puede notar que el dato Delta es el que más valor tiene (15 unidades, como se puede notar en el tooltip) respecto a sus “vecinos” Epsilon y Zeta, que son más pequeños.



Ilustración 94 Ejemplo de Tree Map con la librería D3plus

Si los elementos del gráfico tienen asociada una fecha, se puede además añadir una timeline a este gráfico, para poder ver la evolución de los datos en el tiempo.

En el panel de control de las fechas, que se coloca en la parte inferior del gráfico, se puede seleccionar la fecha que se quiere visualizar o, haciendo click en el botón con forma de triángulo a la izquierda, es posible ver las varias gráficas una detrás de otra automáticamente, para poder observar los cambios en el tiempo.

En la siguiente imagen, un ejemplo del apartado de gestión de las fechas en un Tree map.



Ilustración 95 Gestión de las fechas en un Tree map

6.1.2 Datos a disposición

Para realizar el gráfico, se ha decidido utilizar un conjunto de datos que representa el número de páginas web de escuelas para cada universidad española.

Se disponía solo de los datos del 2011, así que para poder crear la timeline, se han insertado números aleatorios para los años 2010 y 2009 ya que el fin de este apartado es analizar cómo funciona la librería gráfica.

Los datos los tenemos en una hoja Excel, una columna con el nombre de las universidades, y otras tres con los valores en los años 2011, 2010 y 2009.

En la siguiente imagen se puede ver un pantallazo de los datos abiertos con el programa Microsoft Excel.

| | A | B | C | D |
|---|-------------------------------------|-------------|-------------|-------------|
| 1 | Universidad | 2011 | 2010 | 2009 |
| 2 | Universidad de Castilla-La Mancha | 20 | 9 | 6 |
| 3 | Universidad del País Vasco | 18 | 19 | 7 |
| 4 | Universidad Politécnica de Madrid | 18 | 5 | 1 |
| 5 | Universidad Politécnica de Cataluña | 15 | 13 | 20 |
| 6 | Universidad de Oviedo | 12 | 11 | 11 |
| 7 | Universidad Politécnica de Valencia | 12 | 7 | 12 |

Ilustración 96 Fichero con los datos en el programa Microsoft Excel.

6.1.3 Código del gráfico “Tree Map” en la librería D3plus

En este apartado se analizará el código fuente del gráfico que se ha desarrollado.

El código completo se puede encontrar en el Anexo I.

Se trata de un fichero html donde se incrusta el gráfico.

En la cabecera del fichero, hay que importar las librerías que permiten dibujar el gráfico y utilizar las funciones de la librería D3plus. Se hace de esta manera:

```
<script src="http://www.d3plus.org/js/d3.js"></script>
```

```
<script src="http://www.d3plus.org/js/d3plus.js"></script>
```

Después, hay que insertar la sección donde vamos a insertar el gráfico con el tag html “div”. En este caso, hemos llamado esta sección “viz”. Dado eso, el código es el siguiente:

```
<div id="viz"></div>
```

Llegado a este punto, es necesario popular la librería con los datos que queremos representar. En la librería D3plus, se tiene que crear un array y popular cada elemento de la forma “Campo”: “Valor”, como se puede ver claramente en el siguiente ejemplo, donde se rellena el array “Datos_Escuelas” con una ocurrencia, en detalle se añade que la 'Universidad de Castilla-La Mancha' en el año 2011 poseía 20 páginas web de escuelas.

```
var Datos_Escuelas = [
```

```
{'Año': 2011, 'Nombre': 'Universidad de Castilla-La Mancha', 'Escuelas': 20},
```

```
....
```

```
]
```

Después de haber rellenado el array con todos los datos, por último hace falta crear el gráfico y configurarlo.

Esto se realiza en la parte final del fichero, creando una variable y asignándole el método viz() de la librería D3plus y sus configuraciones.

A continuación lo que se ha insertado en este ejemplo, con comentarios explicativos.

```
var visualization = d3plus.viz() → Se crea variable “Visualization” que será lo que se visualizará
```

```
.container("#viz") → Se asocia el nombre al container
```

```
.data(Datos_Escuelas) → Se le indica de que variable tiene que leer los datos (en este caso del array Datos_Escuelas)
```

```
.type("tree_map") → Se le indica que tipo de gráfico se requiere (en este caso el tree map)
```

`.id("Nombre")` → Se le indica que parámetro hay que visualizar dentro de los cuadrados (en este caso el Nombre)

`.size("Escuelas")` → Se le indica dependiendo de que parámetro hay que dibujar el tamaño de los cuadrados, en este caso del parámetro "Escuelas"

`.tooltip(["Nombre", "Escuelas", "Año"])` → Se le indica que parámetros hay que visualizar en la ventanita que sale cuando se pasa el ratón en los cuadrados (tooltip), en este caso el Nombre, la Escuela y el Año

`.time({"value": "Año", "solo": 2011})` → Se habilita la timeline, y se pone que por defecto se visualizan los datos de 2011

`.draw()` → Se activa la visualización

Como se ha visto anteriormente en la descripción del código, para "pasar" los datos a la librería hay que indicárselos de la forma "campo": "valor", "campo": "valor", etc.

Sin embargo, nosotros disponemos de los datos en dos columnas de un fichero Excel. Si se tratase de pocos datos, se podría insertarlos a mano en el código. En nuestro caso sin embargo necesitamos insertar muchos valores, y esta opción entonces no aplica.

Por esta razón vamos a crear un mapa para convertir el formato de los valores.

6.1.4 Mapping con el software Ebimap

Para resolver este problema se ha decidido realizar un “mapa” que sepa convertir los datos en el formato de input en los datos con el formato esperado por la librería.

6.1.4.1 El software Ebimap

Se ha decidido recurrir a esta solución ya que el “mapping” es una solución muy útil en la gestión de la información, ya que a menudo la información no es estructurada en el mismo formato que se espera para poderla tratar.

Para crear el mapa, se ha utilizado el software Ebimap [56].

EbiMap (Edicom Business Intergratior Mapping Tool) es el software mapeador desarrollado por la empresa valenciana Edicom [57] para realizar transformación de estructuras de datos en múltiples formatos a partir de la información contenida en los sistemas informáticos de origen y destino a integrar en una transacción comercial o de cualquier tipo.

La aplicación permite administrar esquemas de transformación de estructuras de datos, que garantizan una perfecta adaptación de cualquier fichero propietario (txt, csv, idoc, etc.) a cualquier estándar XML/EDI (EDIFACT, X21, XML, VDA, TRADACOM...) y viceversa.

En la siguiente imagen de ejemplo, se puede ver el software Ebimap en acción, en particular un mapeo de conversión de un pedido en formato EDIFACT [58] a un documento CSV.

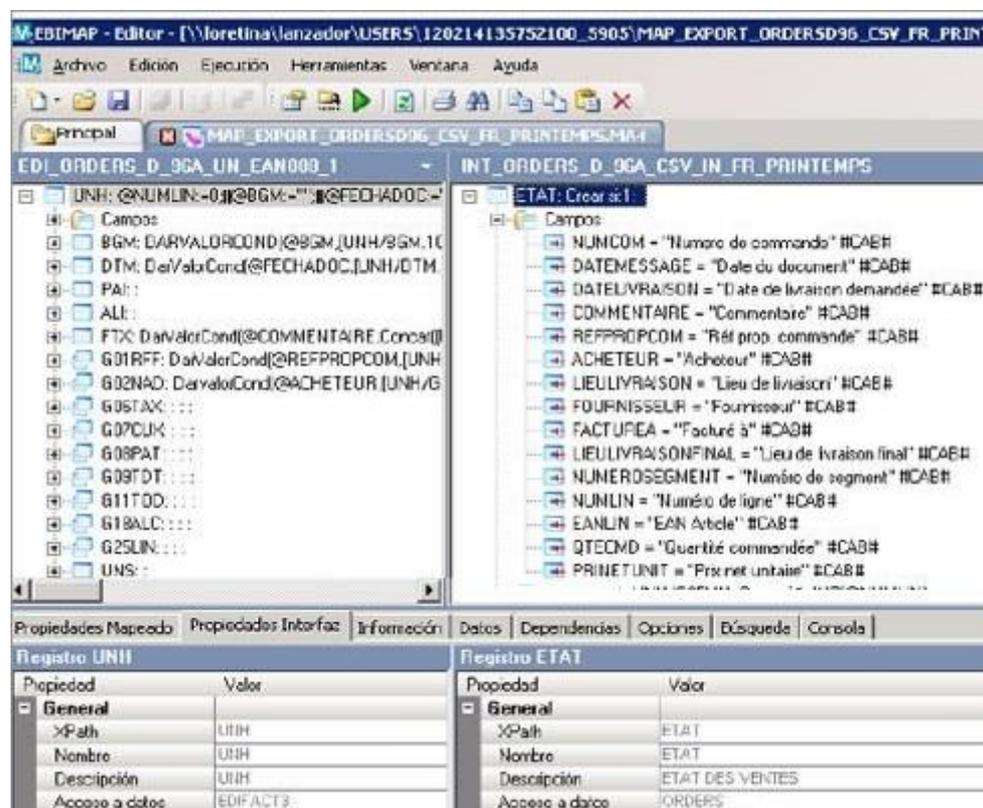
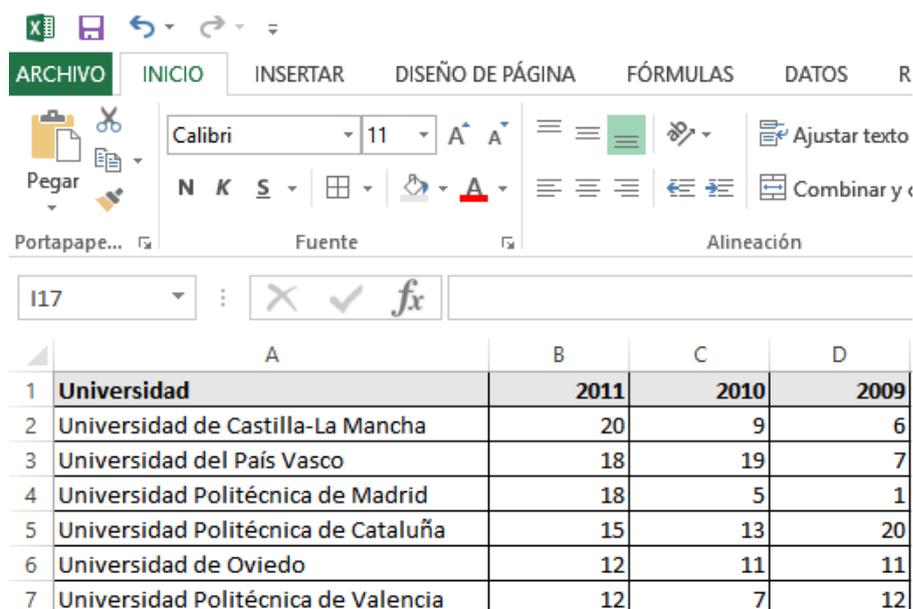


Ilustración 97 Ejemplo de mapeo con el software Ebimap

6.1.4.2 Interfaz Origen y descripción de los componentes del mapa

En nuestro caso, se ha desarrollado un mapa con interfaz de origen capaz de leer un fichero CSV (datos en cada línea con separador “;”) y como interfaz destino un archivo de texto con los datos en el formato requerido (“campo”: “valor”, “campo”: “valor”, etc.).

En detalle, las columnas del CSV de input, como se vio anteriormente, son “Nombre Universidad”, “Escuelas en el año 2011”, “Escuelas en el año 2010” y “Escuelas en el año 2009”.



| | A | B | C | D |
|---|-------------------------------------|-------------|-------------|-------------|
| 1 | Universidad | 2011 | 2010 | 2009 |
| 2 | Universidad de Castilla-La Mancha | 20 | 9 | 6 |
| 3 | Universidad del País Vasco | 18 | 19 | 7 |
| 4 | Universidad Politécnica de Madrid | 18 | 5 | 1 |
| 5 | Universidad Politécnica de Cataluña | 15 | 13 | 20 |
| 6 | Universidad de Oviedo | 12 | 11 | 11 |
| 7 | Universidad Politécnica de Valencia | 12 | 7 | 12 |

Ilustración 98 Fichero con los datos en el programa Microsoft Excel.

Guardando el Excel en CSV y especificando en las opciones que se quiere utilizar el “;” como separador, apendo el fichero con un editor de texto los datos están estructurados de la siguiente manera:

```
1 Universidad de Castilla-La Mancha;20;9;6
2 Universidad del País Vasco;18;19;7
3 Universidad Politécnica de Madrid;18;5;1
4 Universidad Politécnica de Cataluña;15;13;20
5 Universidad de Oviedo;12;11;11
6 Universidad Politécnica de Valencia;12;7;12
```

Ilustración 99 Fichero con los datos abiertos con un editor de texto (Notepad++)

La interfaz origen del mapa entonces es la siguiente:



| Nombre del interfaz | Campos del registro: HEADER | | | | | | | |
|---------------------|-----------------------------|-------------|------------|----------|---------|---------|--------|-------|
| HEADER | Nombre | Descripción | Tipo Datos | Long.Max | Long... | Decimal | Estado | Lista |
| | Nombre | | X | 70 | 0 | 0 | C | |
| | Escuelas2011 | | N | 5 | 0 | 0 | C | |
| | Escuelas2010 | | N | 5 | 0 | 0 | C | |
| | Escuelas2009 | | N | 5 | 0 | 0 | C | |

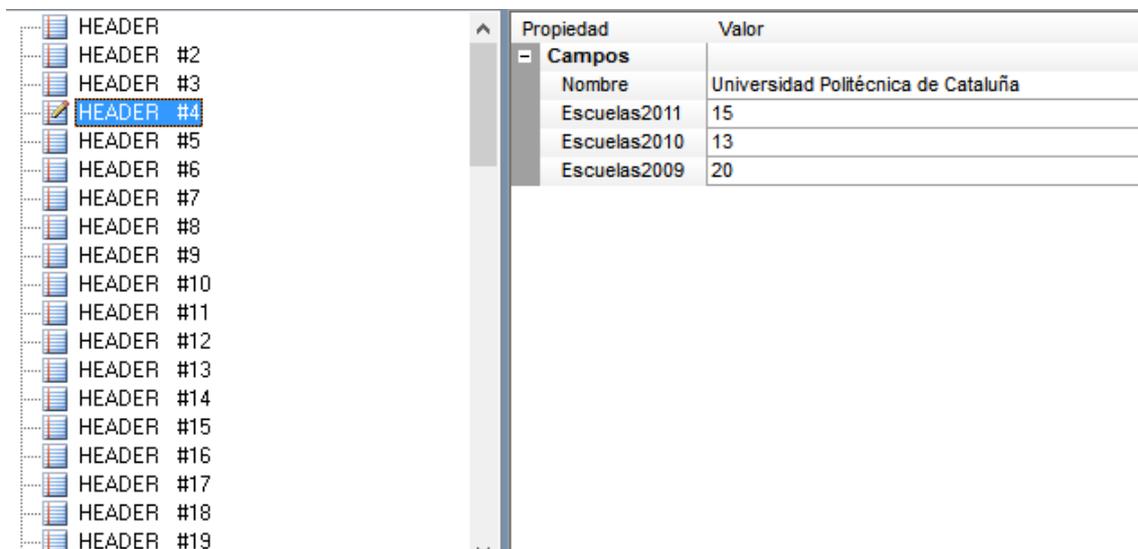
Ilustración 100 Interfaz de origen del mapa en Ebimap

Se puede ver como se ha creado solo un registro (HEADER) ya que en el fichero de origen solo hay un tipo de línea.

Este registro está compuesto por cuatro campos:

1. Campo "Nombre", alfanumérico de longitud 70, donde se almacenará el nombre de la facultad.
2. Campo "Escuelas2011", numérico de longitud 5, donde se almacenará el número de escuelas de la facultad en el año 2011.
3. Campo "Escuelas2010", numérico de longitud 5, donde se almacenará el número de escuelas de la facultad en el año 2010.
4. Campo "Escuelas2019", numérico de longitud 5, donde se almacenará el número de escuelas de la facultad en el año 2019.

En la siguiente imagen se representa una captura de pantalla de la función "Ver datos" del Ebimap, que permite visualizar el fichero a través de la interfaz, y se puede notar como es leído correctamente por el programa (todos los datos están en el campo correspondiente).



| Propiedad | Valor |
|---------------|-------------------------------------|
| Campos | |
| Nombre | Universidad Politécnica de Cataluña |
| Escuelas2011 | 15 |
| Escuelas2010 | 13 |
| Escuelas2009 | 20 |

Ilustración 101 Fichero de origen CSV abierto con la interfaz del mapa en el software Ebimap

Ya que el resultado de nuestro mapa no es ningún formato standard, no se ha creado ninguna interfaz destino (en alternativa se podía crear una interfaz destino DUMMY, es decir sin registros ni campos) y se ha utilizado la función del Ebimap denominada "Escribir", que escribe texto en un fichero según unas condiciones.

Esta función admite los siguientes parámetros:

Escribir(Nombre fichero, texto, condición(Facultativo)).

En el primer parámetro hay que indicar el nombre del fichero, en el segundo el texto que se quiere escribir en el fichero y en el tercer parámetro un booleano (el valor 1 que indica positivo o 0 que indica negativo) como condición de ejecución de la función.

Esta función se ha insertado en el Script final del registro HEADER (es decir, después de haber leído la línea) y es la siguiente:

```
ESCRIBIRF("DATA.TXT","{'Año': 2011, 'Nombre':'&[HEADER.NOMBRE]&', 'Escuelas':  
"&[HEADER.ESCUELAS2011]&"},",1);
```

```
ESCRIBIRF("DATA.TXT","{'Año': 2010, 'Nombre':'&[HEADER.NOMBRE]&', 'Escuelas':  
"&[HEADER.ESCUELAS2010]&"},",1);
```

```
ESCRIBIRF("DATA.TXT","{'Año': 2009, 'Nombre':'&[HEADER.NOMBRE]&', 'Escuelas':  
"&[HEADER.ESCUELAS2009]&"},",1);
```

Con esta función, para cada línea del fichero CSV, se escriben tres líneas en el fichero de destino "data.txt", una para cada año. En cada línea, se han insertado los cuatro campos de la interfaz origen (HEADER.NOMBRE, HEADER.ESCUELAS2011, HEADER.ESCUELAS2010 y HEADER.ESCUELAS2009).

En detalle, tomando por ejemplo las primeras dos líneas:

| | | | | |
|---|-----------------------------------|----|----|---|
| 1 | Universidad de Castilla-La Mancha | 20 | 9 | 6 |
| 2 | Universidad del País Vasco | 18 | 19 | 7 |

El resultado en el fichero de texto será el siguiente, convertido en el formato que se espera la librería:

```
1 {'Año': 2011, 'Nombre':'Universidad de Castilla-La Mancha', 'Escuelas': 20},  
2 {'Año': 2010, 'Nombre':'Universidad de Castilla-La Mancha', 'Escuelas': 9},  
3 {'Año': 2009, 'Nombre':'Universidad de Castilla-La Mancha', 'Escuelas': 6},  
4 {'Año': 2011, 'Nombre':'Universidad del País Vasco', 'Escuelas': 18},  
5 {'Año': 2010, 'Nombre':'Universidad del País Vasco', 'Escuelas': 19},  
6 {'Año': 2009, 'Nombre':'Universidad del País Vasco', 'Escuelas': 7},
```

6.1.5 Visualización del gráfico

Copiando el contenido de este fichero en el fichero index.html con nuestra gráfica, la librería crea el tree map con las universidades representada por cuadrados de diferentes colores, cuyo tamaño se corresponde al número de escuelas.

Además, pasando el ratón encima de los cuadrados, se puede leer el nombre de la universidad, el año y el número de escuela.

En la parte inferior de la página, es posible seleccionar el año que se quiere ver.

En las siguientes 3 capturas de pantalla se puede ver el resultado de este gráfico, seleccionando los tres años:



Ilustración 102 Gráfico Tree Map con los datos del 2011



Ilustración 103 Gráfico Tree Map con los datos del 2010

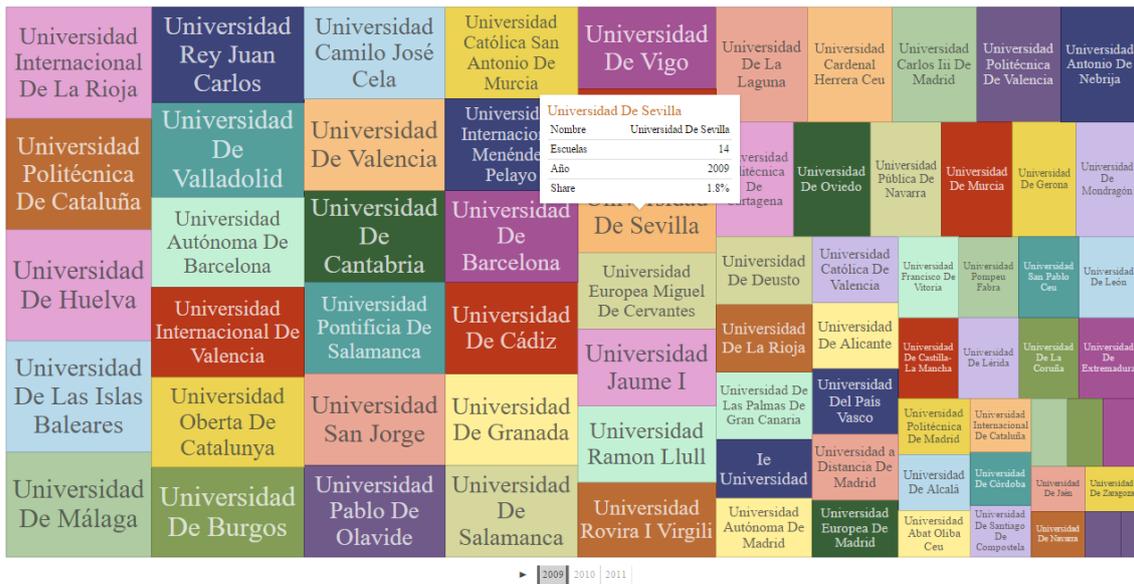


Ilustración 104 Gráfico Tree Map con los datos del 2009

6.1.6 Errores de la librería

Como se puede notar de las tres imágenes anteriores, la única mejora que se podría realizar a la librería es en el caso de tener un cuadrado demasiado pequeño donde el texto no cabe (por ejemplo en el margen inferior a la derecha).

En este caso la librería deja el cuadrado vacío, se podría poner algún mensaje o icono para que se avise el usuario de esto. De todas formas, pasando el ratón encima del cuadrado, el Nombre sigue saliendo en el tooltip, como se puede ver en la siguiente imagen.



Ilustración 105 Error en el gráfico Tree Map



6.2 Ejemplo 2: Red de enlaces entre 9 de las Universidades Españolas más importantes

En este segundo ejemplo se va a crear un gráfico de red (gráfico “network” de la librería D3plus) que conecta nueve de las universidades Españolas más importantes.

Para establecer cuáles son las universidades más importantes, se ha considerado el Ranking Mundial “Ranking Web (Webometrics) de Universidades” [59].

El Ranking Web (Webometrics) de Universidades proporciona la clasificación más completa y actualizada de instituciones de educación superior de todo el mundo. El Laboratorio de Cibermetría (Consejo Superior de investigaciones Científicas, CSIC, España) publica desde 2004, cada seis meses, un ranking académico independiente con el objetivo de suministrar información fiable, multidimensional, actualizada y útil sobre las universidades de todo el mundo teniendo en cuenta su presencia e impacto en la Web.

En la actualidad el Ranking incluye más de 21.000 universidades con dos ediciones, una que aparece a finales de Enero y una segunda disponible al terminar Julio.

Dicho esto, las 10 universidades Españolas más importantes son las siguientes, en orden de ranking [60]:

1. Universitat Politècnica de Catalunya BarcelonaTech (www.upc.edu), posición 110 en el ranking mundial.
2. Universitat de Barcelona (www.ub.edu), posición 116 en el ranking mundial.
3. Universidad Complutense de Madrid (www.ucm.es), posición 151 en el ranking mundial.
4. Universitat Autònoma de Barcelona (www.uab.cat), posición 168 en el ranking mundial.
5. Universitat de València (ww.uv.es), posición 170 en el ranking mundial.
6. Universidad de Granada (www.ugr.es), posición 178 en el ranking mundial.
7. Universidad Politécnica de Valencia (www.upv.es), posición 227 en el ranking mundial.
8. Universidad Autónoma de Madrid (www.uam.es), posición 230 en el ranking mundial.
9. Universidad de Sevilla (www.us.es), posición 234 en el ranking mundial.
10. Universidad Politécnica de Madrid (www.upm.es), posición 265 en el ranking mundial.

En este ejemplo no vamos a considerar la número 1, es decir www.upc.edu, ya que no se han podido descargar sus backlinks de Majestic SEO porque se han superado el número máximo de backlinks descargados con el tipo de account que disponíamos.

El gráfico que se pretende realizar es un gráfico de bolas conectadas con las siguientes características:

- El tamaño de cada bola corresponde al tamaño de la página web de la universidad
- Las aristas del grafo representan los enlaces entre las dos universidades. El sentido representa “quien enlaza a quien” y el grosor la cantidad de enlaces.



6.2.1 Descubrir el tamaño de las páginas web

Para empezar a crear el gráfico, necesitamos saber el tamaño de las páginas web para poder establecer el tamaño de las bolas en el gráfico. Para esto, vamos a utilizar el comando “site:” de Google [61]. Este comando permite listar solo las páginas indexadas en la totalidad del dominio que se le pasa por parámetro.

Por ejemplo, “site:upm.es” lista las páginas indexadas en la totalidad del portal upm.es, comprendidos todos los sub-dominios. Será entonces resultado de la búsqueda la página www.upm.es/directorio pero también la página www.cesvima.upm.es/ o www.etsisi.upm.es/.

Es posible también utilizar este comando atacando un subdominio (poniendo el url del subdominio en lugar del URL del dominio).

En este ejemplo vamos a considerar solo las páginas contenidas en el dominio raíz, sin subdominios.

El valor que se tomará será el que se indica justo debajo de la SERP (Search engine results page o Página de resultados del buscador) de Google, como se puede notar en la siguiente imagen.

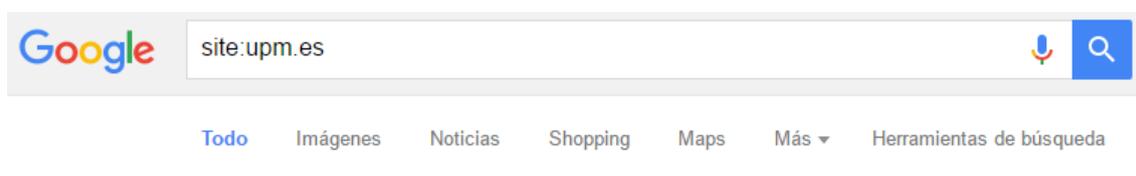


Ilustración 106 Búsqueda con el comando “site:” de Google en el dominio upm.es

El resultado de las queries sobre las 9 universidades españolas más importantes se puede ver en la siguiente captura de pantalla del fichero Excel donde se han apuntado.

En la primera columna se puede ver el URL de la universidad, en la segunda la sintaxis del comando “site:” y en la tercera el resultado.

| DOMINIO | QUERY | RESULTADO |
|---------|--------------|-----------|
| ub.edu | site:ub.edu | 1.900.000 |
| ucm.es | site:ucm.es | 1.870.000 |
| uab.cat | site:uab.cat | 2.030.000 |
| uv.es | site:uv.es | 1.560.000 |
| ugr.es | site:ugr.es | 1.670.000 |
| upv.es | site:upv.es | 1.470.000 |
| uam.es | site:uam.es | 690.000 |
| us.es | site:us.es | 3.050.000 |
| upm.es | site:upm.es | 1.170.000 |

Ilustración 107 Resultado de las queries con el comando de Google “site:”

6.2.2 Calcular la cantidad de enlaces desde una universidad hacia otra

Llegados a este punto, nos queda descubrir los enlaces entre las universidades. Es decir, cuantos enlaces hay desde la universidad A hacia la universidad B, y viceversa, para las 10 universidades.

Para realizar esta operación, antes de todo, necesitamos saber los backlinks de cada universidades, es decir las páginas web que enlazas nuestra página.

Más en detalle, los Backlinks, o vínculos externos de respaldo, son los enlaces que recibe una página web desde otras páginas web. El número de backlinks es la cantidad de páginas que la enlazan a través de un vínculo (puede ser en texto o gráfico) [62].

En nuestro caso, las operaciones que realizaremos serán las siguientes:

1. Sacar los backlinks que enlazan a la universidad A. Esta operación nos devuelve una lista de enlaces.
 - a. Buscar en esta lista de enlaces, cuántos son de la universidad B. Este valor en el gráfico será el grosor de la artista que enlaza A con B (en este sentido).
 - b. Repetir el punto a. Para todas las demás universidades (Enlaces de A con C, A con D etc.)
2. Sacar los backlinks de las demás universidades y repetir el punto 1 y sus subpuntos.



6.2.2.1 Obtener los backlinks de una página web con Majestic SEO

Vamos a utilizar la herramienta Majestic SEO para obtener los backlinks de nuestras páginas web.

Como se define en su propia página web [63], Majestic SEO es una herramienta que inspecciona y mapea Internet, y ha creado la mayor base de datos comercial de inteligencia de enlaces del mundo.

Este mapa de Internet es utilizado por servicios de posicionamiento en buscadores (SEO), especialistas en nuevos medios, gerentes de asociados y expertos en marketing en línea, con diversos fines relacionados con la prominencia en Internet, como la creación de enlaces, la administración de la reputación, el desarrollo del tráfico en Internet, el análisis de la competencia y el seguimiento de noticias.

Para descargar los backlinks desde Majestic SEO hay que realizar las siguientes operaciones:

1. Ir al apartado informes de Majestic SEO, que se encuentra en <https://es.majestic.com/reports>. Necesitamos crear un informe para cada página web de la cual queremos sacar los backlinks, ya que con la opción básica para calcular los backlinks de un dominio, Majestic SEO permite solo descargar los primeros 5000. La posibilidad de crear informe es una opción de pago, sin embargo se dispone de un account Premium así que se puede realizar esta operación.
2. En la ventana “Opciones de informe”, seleccionamos “Crear informe” y escribimos el dominio que necesitamos, como se ve en la siguiente captura de pantalla.

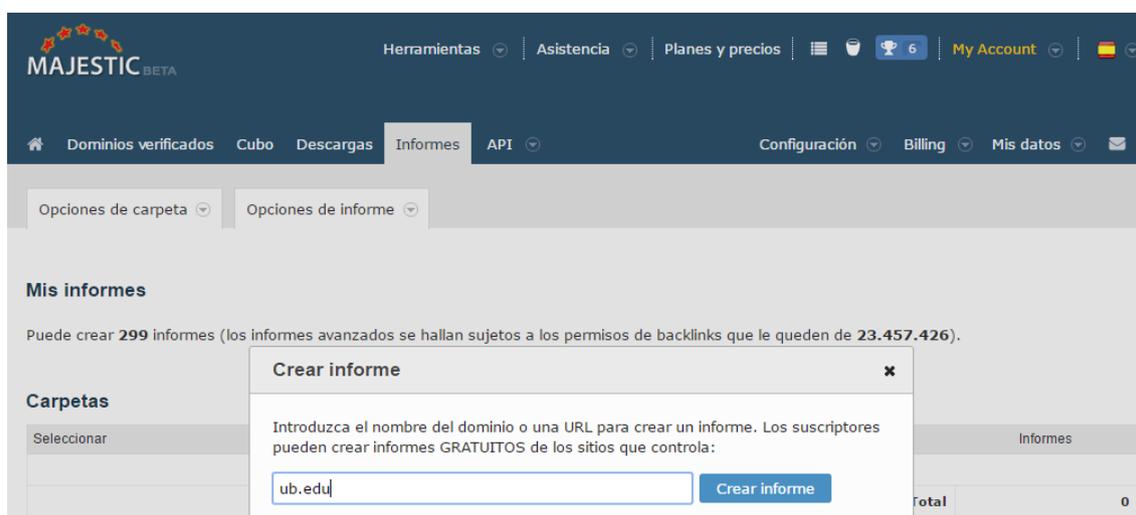


Ilustración 108 Creamos un informe en Majestic SEO

- Elegimos el informe avanzado, ya que en informe estándar se ofrece información detallada pero solamente de los enlaces más importantes a su sitio web, y elegimos Domain.

Tipo de informe

| | |
|--|--|
|  <p><input type="radio"/> Informe estándar</p> <p>Ofrece información detallada de los enlaces más importantes a su sitio web.</p> |  <p><input checked="" type="radio"/> Informe avanzado</p> <p>Ofrece todos los datos de backlinks para su dominio raíz, subdominio o URL (dentro de su concesión).</p> |
|--|--|

Coste del informe

| | |
|--|---|
| <p><input checked="" type="radio"/> Domain</p> <p>ub.edu</p> <p>Consumirá:</p> <ul style="list-style-type: none"> • 1 de sus 299 informes • 2.063.084 unidades de análisis (aprox. 9% de su concesión) <p style="text-align: right;">Crear informe</p> | <p><input type="radio"/> Subdomain</p> <p>www.ub.edu</p> <p>Consumirá:</p> <ul style="list-style-type: none"> • 1 de sus 299 informes • 1.442.631 unidades de análisis (aprox. 7% de su concesión) |
|--|---|

Ilustración 109 Selección de descarga de informe avanzado

- Cuando el informe esté listo desde las opciones de descarga seleccionamos la opción “Descargar Backlinks”. En el apartado Descargas encontraremos nuestro CSV con la lista de todos los backlinks. Cuando esté listo, podemos descargarlo en formato CSV.

En la siguiente imagen se puede ver el fichero resultante CSV de los backlinks de la página web “upv.es” abierto con Microsoft Excel. Los backlinks están en la segunda columna, denominada “Source URL”. En la primera columna, “Target URL”, se puede ver la URL de destino, en este caso upv.es.

| | A | B | C | D | E |
|----|---------------|--|-----------------|-------------------|-------------------------|
| 1 | Target URL | Source URL | Anchor Text | Source Crawl Date | Source First Found Date |
| 2 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 3 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 4 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 5 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 6 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 7 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 8 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 9 | http://upv.es | http://emb.at/_php/webbuddy/kategorien.php?FUNC= | findige spanier | 12/07/2016 | 12/07/2016 |
| 10 | http://upv.es | http://www.emb.at/_php/webbuddy/kategorien.php? | findige spanier | 11/08/2016 | 11/08/2016 |
| 11 | http://upv.es | http://www.emb.at/_php/webbuddy/kategorien.php? | findige spanier | 11/08/2016 | 11/08/2016 |
| 12 | http://upv.es | http://www.emb.at/_php/webbuddy/kategorien.php? | findige spanier | 06/08/2016 | 06/08/2016 |
| 13 | http://upv.es | http://www.emb.at/_php/webbuddy/kategorien.php? | findige spanier | 06/08/2016 | 06/08/2016 |
| 14 | http://upv.es | http://www.emb.at/_php/webbuddy/kategorien.php? | findige spanier | 02/08/2016 | 02/08/2016 |
| 15 | http://upv.es | http://www.emb.at/_php/webbuddy/kategorien.php? | findige spanier | 02/08/2016 | 02/08/2016 |

Ilustración 110 Fichero CSV generado por Majestic SEO con los backlinks de www.upv.es



Los ficheros CSV con los backlinks de las universidades se pueden consultar en los adjuntos de este trabajo.

A este punto, hay que buscar los backlinks entre nuestras nueve páginas web y crear la tabla de los enlaces.

Vamos a analizar esta operación para la página web “upv.es”(es decir ver en las demás página web cuantos links hacia www.upv.es hay) y de esta manera se explicarán en detalle los varios paso.

1. Abrimos el fichero CSV con los backlinks de upv.es con el Notepad++
2. Realizamos 8 búsquedas, una para cada universidad. Por ejemplo buscamos la cadena .ucm.es y contamos los resultados de la búsqueda (por ejemplo, un resultado es <https://www.ucm.es/universidades-publicas-espanolas>).
3. Anotamos en un Excel la cantidad de resultados de cada búsqueda (con el comando Contar de la ventana de búsqueda de Notepad++).

En la siguiente imagen se puede ver por ejemplo la tabla con los backlinks de ub.edu provenientes de las demás 8 universidades.

| ORIGEN | DESTINO | BACKLINKS |
|---------|---------|-----------|
| ucm.es | ub.edu | 647 |
| uab.cat | ub.edu | 1064 |
| uv.es | ub.edu | 386 |
| ugr.es | ub.edu | 110 |
| upv.es | ub.edu | 304 |
| uam.es | ub.edu | 115 |
| us.es | ub.edu | 68 |
| upm.es | ub.edu | 9 |

Ilustración 111 Backlinks del dominio ub.edu que vienen de las 8 universidades

Para los ficheros CSV con los backlinks de 4 universidades, ya que eran mayores de 350Mb, no se ha podido utilizar Notepad++ ya que salía el error que el fichero era demasiado grande. Para solucionar este problema, se ha utilizado el programa gratuito EdidPadLite [64].

Ahora que tenemos los datos, podemos proceder a crear el gráfico con la librería D3plus.

6.2.3 Código del gráfico “Network” en la librería D3plus

En este apartado se analizará el código fuente del gráfico que se ha desarrollado. Se trata de un fichero html donde se incrusta el gráfico.

El código completo se puede encontrar en el Anexo II.

Este código tiene muchas partes en común con el código del gráfico “Tree Map” que se ha analizado precedentemente.

En la cabecera del fichero, hay que importar las librerías que permiten dibujar el gráfico y utilizar las funciones de la librería D3plus. Se puede hacer de esta manera:

```
<script src="http://www.d3plus.org/js/d3.js"></script>
```

```
<script src="http://www.d3plus.org/js/d3plus.js"></script>
```

Después, hay que insertar la sección donde vamos a insertar el gráfico con el tag html “div”. En este caso, hemos llamado esta sección “viz”. Dado eso, el código es el siguiente:

```
<div id="viz"></div>
```

6.2.3.1 Creación de los nodos del gráfico “network”

Llegados a este punto, es necesario popular la librería con los datos que queremos representar. En la librería D3plus, se tiene que crear un array y popularlo de la forma “Campo”: “Valor”.

En nuestro caso, los campos que necesitamos son el nombre de la universidad y su tamaño de páginas. Este último valor ya lo hemos calculado en el apartado 6.2.1.

Como vimos anteriormente, disponemos del dato en el siguiente formato, es decir un fichero CSV separado por “,”

| DOMINIO | QUERY | RESULTADO |
|----------------|--------------|------------------|
| ub.edu | site:ub.edu | 1.900.000 |
| ucm.es | site:ucm.es | 1.870.000 |
| uab.cat | site:uab.cat | 2.030.000 |
| uv.es | site:uv.es | 1.560.000 |
| ugr.es | site:ugr.es | 1.670.000 |
| upv.es | site:upv.es | 1.470.000 |
| uam.es | site:uam.es | 690.000 |
| us.es | site:us.es | 3.050.000 |
| upm.es | site:upm.es | 1.170.000 |

Ilustración 112 Resultado de las queries con el comando de Google "site:" abierto con Microsoft Excel

Abriendo el fichero con el Notepad++ (e ignorando la cabecera), se presenta de la siguiente manera:

```
1 ub.edu;site:ub.edu;1.900.000
2 ucm.es;site:ucm.es;1.870.000
3 uab.cat;site:uab.cat;2.030.000
4 uv.es;site:uv.es;1.560.000
5 ugr.es;site:ugr.es;1.670.000
6 upv.es;site:upv.es;1.470.000
7 uam.es;site:uam.es;690.000
8 us.es;site:us.es;3.050.000
9 upm.es;site:upm.es;1.170.000
```

Ilustración 113 Resultado de las queries con el comando de Google "site:" abierto con el Notepad++

En la librería D3plus, para pasar los datos al gráfico de tipo “network” se utiliza la misma modalidad del gráfico “tree map”, es decir se crea un array y cada elemento se rellena de manera “campos”:“valor”, etc.

El array con el tamaño de las diez universidades será el siguiente:

```
var Datos_universidades = [  
  {"Nombre": "ub.edu", "Paginas": 1900000},  
  {"Nombre": "ucm.es", "Paginas": 1870000},  
  {"Nombre": "uab.cat", "Paginas": 2030000},  
  {"Nombre": "uv.es", "Paginas": 1560000},  
  {"Nombre": "ugr.es", "Paginas": 1670000},  
  {"Nombre": "upv.es", "Paginas": 1470000},  
  {"Nombre": "uam.es", "Paginas": 690000},  
  {"Nombre": "us.es", "Paginas": 3050000},  
  {"Nombre": "upm.es", "Paginas": 1170000}  
]
```

6.2.3.2 Creación de las artistas del gráfico “Network”

Ahora vamos a crear las artistas de gráfico.

El gráfico de tipo network de la librería D3plus necesita el siguiente formato para crear un enlace entre dos nodos: un array con cada elemento que contenga por lo menos los campos “source” (nodo de origen) y “target” (nodo de destino).

Nosotros vamos a pasarle también el campo “strength”, que es el grosor de la artista, que representa la cantidad de enlaces que hay entre dos páginas web.

En el siguiente trozo de código, vamos a crear el array “connections” y añadir manualmente el primer elemento, es decir la arista que va desde el nodo “ucm.es” al nodo “ub.edu” con valor 647:

```
var connections = [  
    {source: 'ucm.es', target: 'ub.edu', strength: 647}  
    ...  
]
```

Sin embargo, no podemos realizar esta operación manualmente ya que tenemos que crear muchas aristas y sobre todo para dejar para un futuro una herramienta para crear gráficos con más de miles de aristas.

Para resolver este problema vamos a crear otro mapeo, como ya hicimos para el gráfico “tree node”.

En la siguiente imagen se explica lo que tiene que realizar el mapa, en la parte izquierda vemos el fichero de origen y en la derecha el fichero destino resultado, con la colocación de los datos en los campos correspondiente (source y target). Además, si no hay enlaces entre nodos, no se tiene que crear la arista.

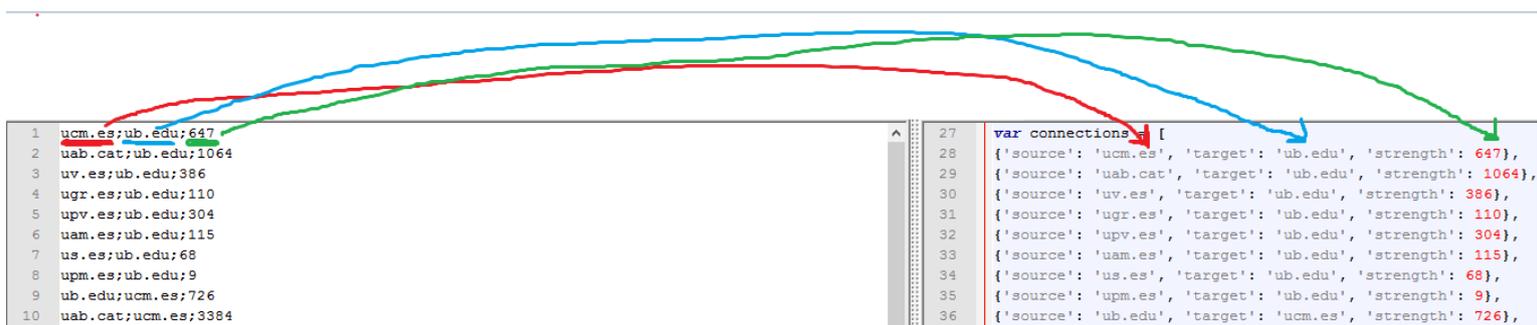


Ilustración 114 Modelo de lo que tiene que realizar el mapa creado para formatear las aristas

Podemos entonces adaptar el mapa creado anteriormente, ya que la interfaz origen es muy parecida (3 campos, separados por “;”) y en el editor cambiamos el contenido del Script que escribe en el fichero de destino.

En detalle, la interfaz de origen será la siguiente:

| Interfaz | | Campos de registro activo | | | | | | |
|----------|-------------|---------------------------|---------|---------|---------|--------|-------|--|
| Nombre | Descripción | Tipo Datos | Long... | Long... | Decimal | Estado | Lista | |
| Origen | | X | 70 | 0 | 0 | C | | |
| Destino | | X | 70 | 0 | 0 | C | | |
| Enlaces | | N | 70 | 0 | 0 | C | | |

Ilustración 115 Interfaz de origen del mapa usado en el gráfico network

En el Script de fin del mapa, vamos a insertar este simple código, que os permite colocar los campos donde necesitamos.

```
ESCRIBIRF("DATA.TXT","{'SOURCE': '&[HEADER.ORIGEN]&', 'TARGET': '&[HEADER.DESTINO]&', 'STRENGTH': '&[HEADER.ENLACES]&'}',[HEADER.ENLACES]<>0);
```

Utilizamos la función Escribif que hemos utilizado anteriormente, la única diferencia con el ejemplo anterior, además del texto, es que en el último parámetro que pasamos a la función, es decir la condición de ejecución, le especificamos que imprima la línea solo si el campo “enlaces” no es cero (se puede ver en negrita en el código anterior).

En la siguiente imagen, se puede ver el código que ha generado el mapa, con las aristas que tienen por lo menos un enlace.

```
var connections = [
  {'source': 'ucm.es', 'target': 'ub.edu', 'strength': 647},
  {'source': 'uab.cat', 'target': 'ub.edu', 'strength': 1064},
  {'source': 'uv.es', 'target': 'ub.edu', 'strength': 386},
  {'source': 'ugr.es', 'target': 'ub.edu', 'strength': 110},
  {'source': 'upv.es', 'target': 'ub.edu', 'strength': 304},
  {'source': 'uam.es', 'target': 'ub.edu', 'strength': 115},
  {'source': 'us.es', 'target': 'ub.edu', 'strength': 68},
  {'source': 'upm.es', 'target': 'ub.edu', 'strength': 9},
  {'source': 'ub.edu', 'target': 'ucm.es', 'strength': 726},
  {'source': 'uab.cat', 'target': 'ucm.es', 'strength': 3384},
  {'source': 'uv.es', 'target': 'ucm.es', 'strength': 2231},
  {'source': 'ugr.es', 'target': 'ucm.es', 'strength': 300},
  {'source': 'upv.es', 'target': 'ucm.es', 'strength': 290},
  {'source': 'uam.es', 'target': 'ucm.es', 'strength': 846},
  {'source': 'us.es', 'target': 'ucm.es', 'strength': 155},
  {'source': 'upm.es', 'target': 'ucm.es', 'strength': 399},
  {'source': 'ucm.es', 'target': 'uab.cat', 'strength': 15},
  {'source': 'ub.edu', 'target': 'uab.cat', 'strength': 439},
  {'source': 'uv.es', 'target': 'uab.cat', 'strength': 54},
  {'source': 'ugr.es', 'target': 'uab.cat', 'strength': 44},
  {'source': 'upv.es', 'target': 'uab.cat', 'strength': 10},
  {'source': 'uam.es', 'target': 'uab.cat', 'strength': 74},
  {'source': 'us.es', 'target': 'uab.cat', 'strength': 47},
  {'source': 'upm.es', 'target': 'uab.cat', 'strength': 11},
  {'source': 'ucm.es', 'target': 'uv.es', 'strength': 91},
```

Ilustración 116 Aristas en el formato requerido por la librería D3plus en el gráfico "network"

La parte final del código se parece bastante a la que se explicó en el gráfico tree map. A continuación se va a enseñar el código con comentarios.

`var visualization = d3plus.viz()` → Se crea la variable "Visualization" que será lo que se visualizará

`.container("#viz")` → Se asocia el nombre al container

`.data(Datos_universidades)` → Se le indica de que array tiene que leer los datos (en este caso del array `Datos_universidades`)

`.type("network")` → Se le indica que tipo de gráfico se requiere (en este caso el `network`)

`.id("Nombre")` → Se le indica que parámetro hay que visualizar dentro de los círculos (en este caso el `Nombre`)

`.size("size")` → Se le indica dependiendo de que parámetro hay que dibujar el tamaño de los cuadrados, en este caso "size", que es el tamaño de la página web

`.edges(connections)` → Se le pasa el array con la definición de las aristas

`.edges({`

`"size": "strength",`

`"value": connections,`

`"arrows": true`

`})` → Se configuran algunos parámetros de las aristas. En este caso, le decimos que el grosor de las flechas depende de la variable "strength", que tome el valor del array "connections" y habilitamos la punta de la flecha que especifica el sentido.

`.draw()` --> Se activa la visualización

6.2.4 Gráficos resultantes

A continuación, el gráfico resultado que ha generado la librería.

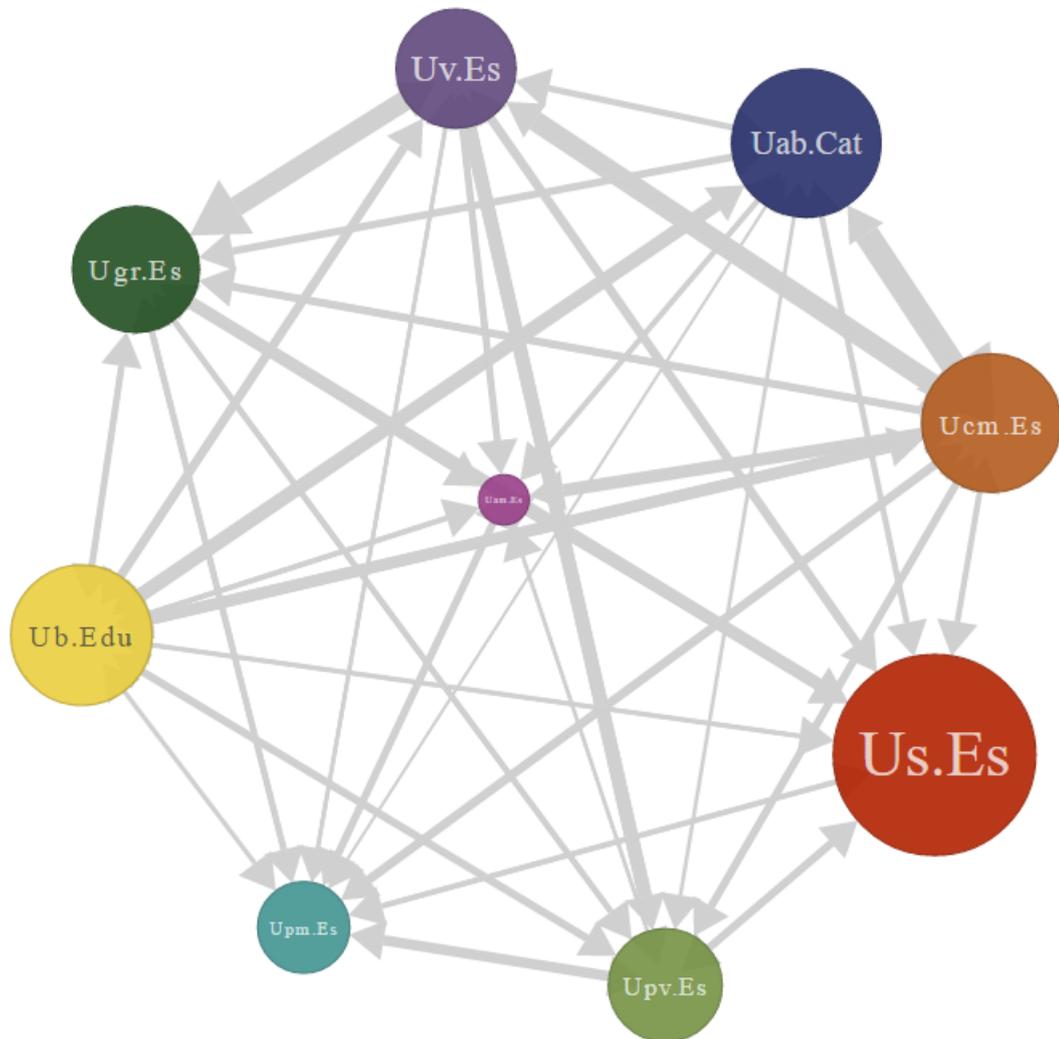


Ilustración 117 Gráfico de tipo network generado con la librería D3plus que representa los enlaces entre las páginas web de las 10 universidad española más importantes, sin considerar subdominios (ej. www.upv.es/)*

A primera vista se puede notar el nombre de la página web en cada círculo, el grosor de las aristas depende de la cantidad de enlaces que hay y las puntas de las flechas indican la dirección de los enlaces.

Además, pasando el ratón encima de los círculos, además de visualizar una ventanita con su nombre y su tamaño, también se evidencian las aristas que involucran dicho nodo, como se puede ver en la siguiente imagen.

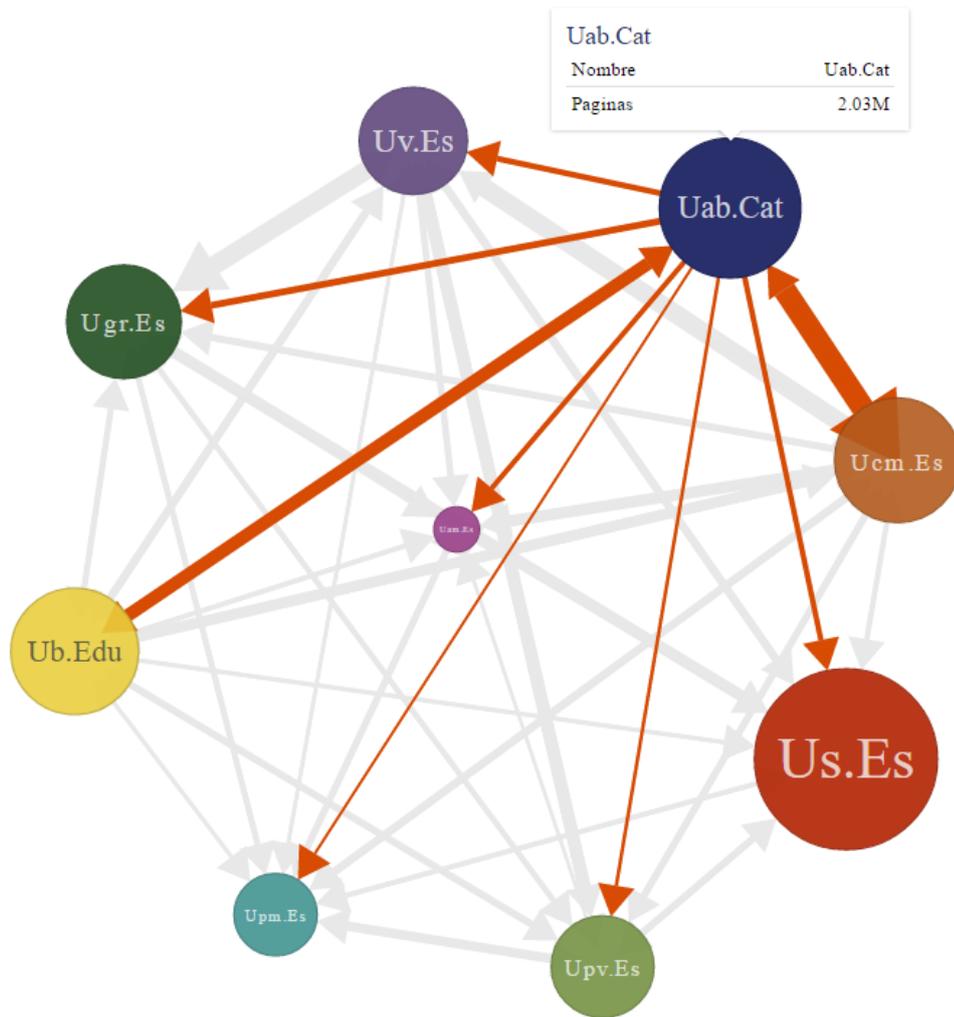
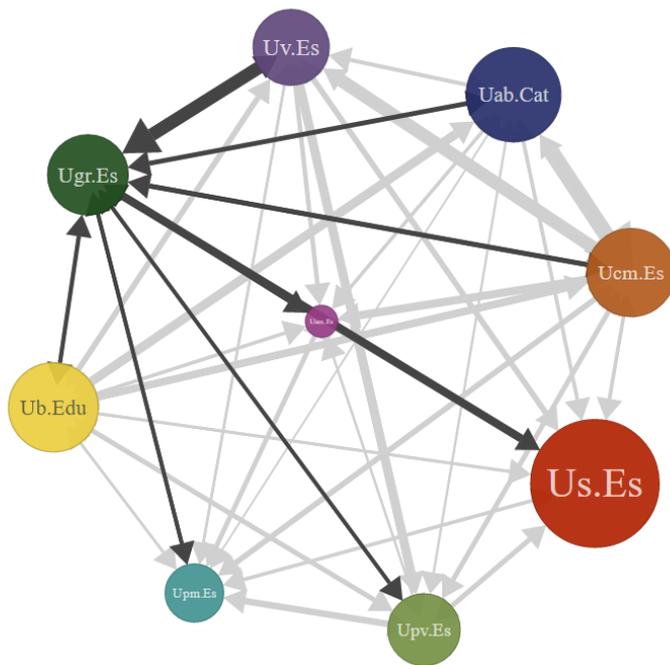


Ilustración 118 Pasando el ratón en un nodo, se evidencian las aristas que lo involucran

Otra característica muy interesante de este tipo de gráfico es que si se hace click en un nodo, además de evidenciar las aristas que lo involucran, se hace “zoom” en este nodo y aparece una ventana con una lista de todas sus conexiones.

En la lista de conexiones si se hace click encima de algún elemento de la lista, se pasa a enfocar dicho elemento.

En la siguiente imagen se puede ver esta operación, es decir se ha hecho click encima del nodo ugr.es y el gráfico ha hecho zoom en este mismo nodo, ha evidenciado las aristas y ha creado la ventana con las conexiones.



| Ugr.Es | |
|---|---------|
| Nombre | Ugr.Es |
| Páginas | 1.67M |
| Primary Connections | |
| ● | Ub.Edu |
| ● | Ucm.Es |
| ● | Uab.Cat |
| ● | Uv.Es |
| ● | Ucm.Es |
| ● | Uab.Cat |
| ● | Uv.Es |
| ● | Ub.Edu |
| ● | Upv.Es |
| ● | Uam.Es |
| ● | Us.Es |
| ● | Upm.Es |
| ● | Upv.Es |
| ● | Uam.Es |
| ● | Us.Es |
| ● | Upm.Es |

Ilustración 119 Comportamiento del gráfico "network" cuando se hace click en un nodo

Otra de las muchas posibilidades que ofrece la librería es poner encima de cada arista el número correspondiente a su grosor, en nuestro caso el número de enlaces.

Sin embargo la librería posiciona este número en el medio de la línea y si hay dos aristas entre dos nodos (en las dos direcciones) los números se superponen.

En la siguiente imagen se puede ver este fenómeno.

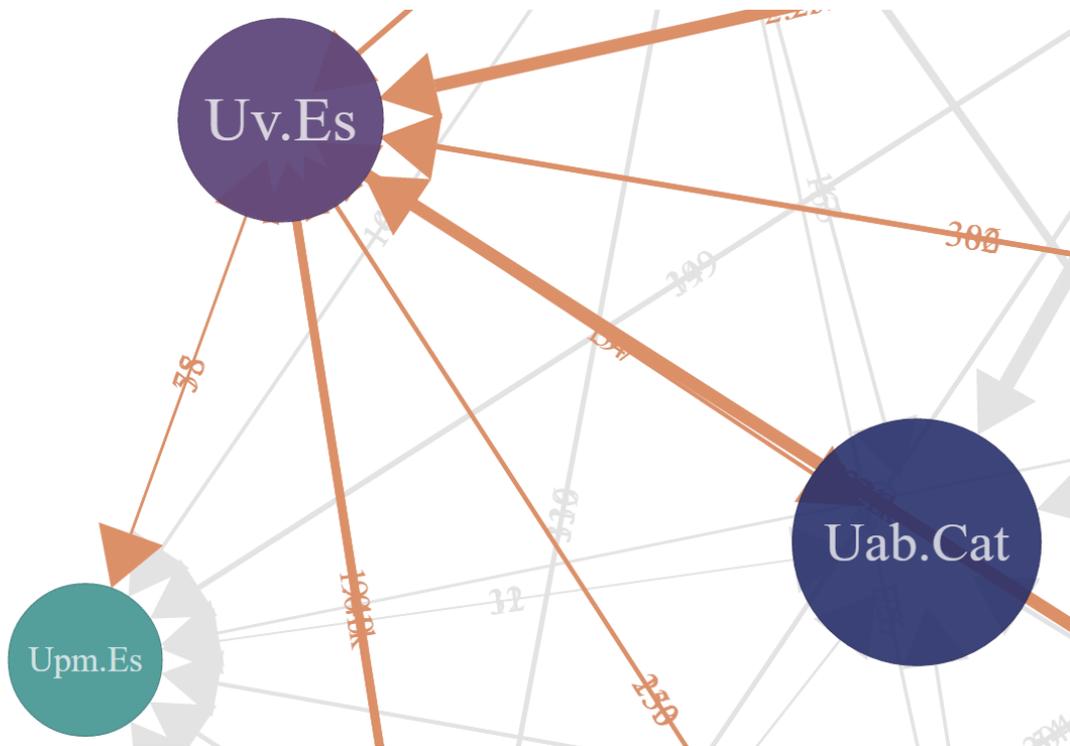


Ilustración 120 Gráfico con la opción "label" de las aristas activada.

Para activar esta opción, es necesario añadir el parámetro label en el objeto ".edges", especificándole que el label es el grosor (en nuestro caso la variable "strength"), como se puede ver del siguiente código.

```
.edges ({  
  "size": "strength",  
  "label": "strength",  
  "value": connections,  
  "arrows": true  
})
```

Otra opción interesante es la posibilidad de fijar la posición de los nodos.

Para realizar esta operación, había que crear un array con, por cada elemento, los campos nombre, con el nombre del nodo, y las coordenada x e y, por ejemplo de esta manera:

```
var positions = [  
  {"Nombre": "upc.edu", "x": 10, "y": 15},  
  {"Nombre": "ub.edu", "x": 12, "y": 24},  
  {"Nombre": "ucm.es", "x": 16, "y": 18},  
  {"Nombre": "uab.cat", "x": 26, "y": 21},  
  {"Nombre": "uv.es", "x": 13, "y": 4},  
  {"Nombre": "ugr.es", "x": 31, "y": 13},  
  {"Nombre": "upv.es", "x": 19, "y": 8},  
  {"Nombre": "uam.es", "x": 24, "y": 11}  
]
```

Al final del código había que insertar este array en las propiedades de los nodos de la siguiente manera:

```
.nodes(positions)
```

Sin embargo, como se especifica en la página web de D3plus (<http://d3plus.org/examples/advanced/9956648/>), si la posición de los nodos no se especifica, D3plus determinará un diseño que tiene como objetivo reducir la superposición de los nodos y de las aristas (Automatic Node Positions in Network).

Estos cálculos se realizan cada vez que se está dibujando el gráfico, “on the fly”, así que actualizando la página la disposición de los componentes del gráfico puede cambiar.

En la siguiente imagen vemos un ejemplo de este fenómeno.

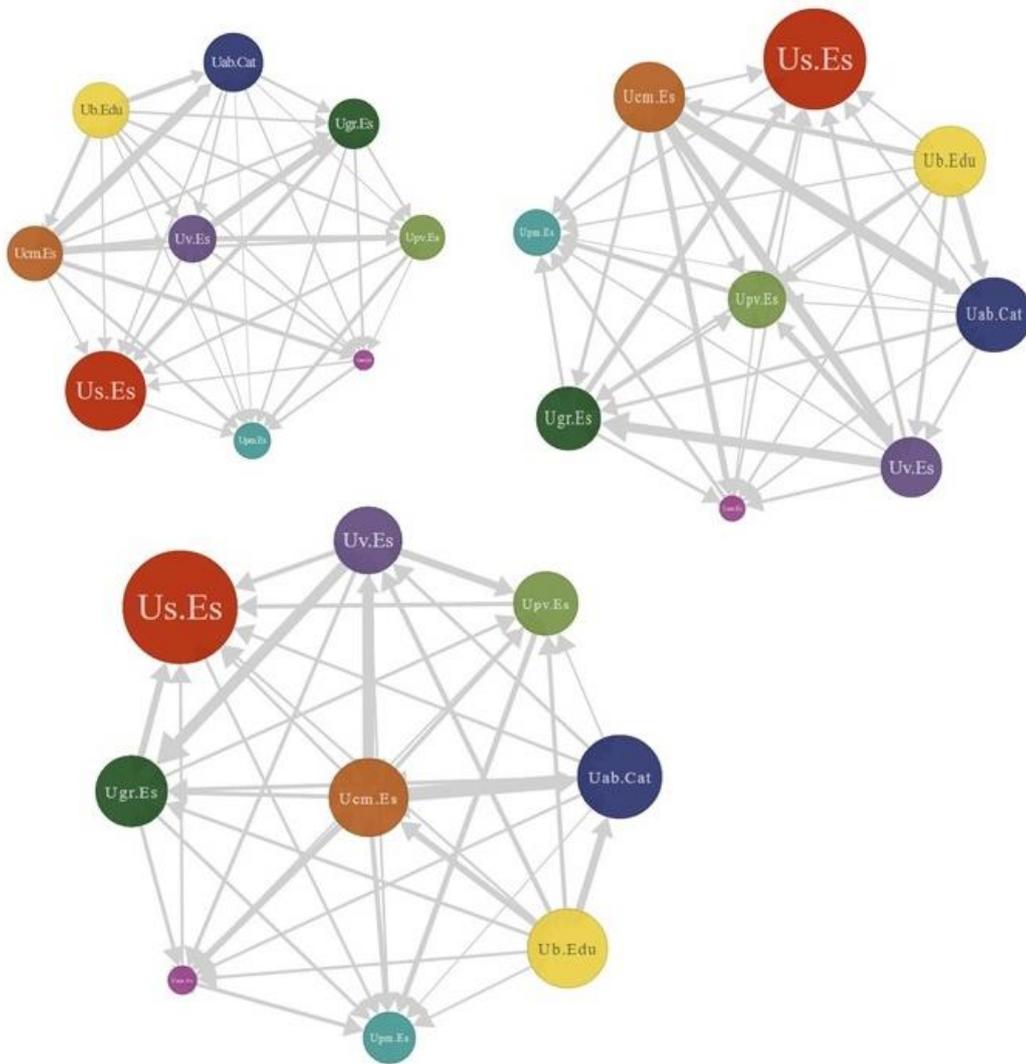


Ilustración 121 Ejemplos de cómo actúa la librería D3plus para posicionar dinámicamente los nodos de un gráfico de tipo "network"

6.3 Ejemplo 3: Creación de una simple red multinivel. Red de enlaces entre nueve de las Universidades Españolas más importantes y entre sus departamentos de matemáticas.

En este ejemplo se pretende crear una red multinivel. Para realizar esta operación, se va a utilizar la red creada en el ejemplo anterior (red de enlaces entre 9 de las universidades españolas más importantes) y se creará otra subred.

Como se ha especificado en el ejemplo anterior, las páginas web de las 9 Universidades españolas seleccionadas son las siguientes:

- ub.edu
- ucm.es
- uab.cat
- uv.es
- ugr.es
- upv.es
- uam.es
- us.es
- upm.es

Se ha decidido crear la subred entre los departamento de matemáticas de estas 9 universidades, ya que se intentó inicialmente crearla con los departamentos de informática, sin embargo varias universidades tienen las facultades de informática dentro del departamento de matemática.

Se intentó también realizar la subred entre las bibliotecas de estas universidades, sin embargo cada universidad tiene muchas web de bibliotecas y era difícil elegir una.

Para crear la red, inicialmente se han realizado los pasos de ejemplo anterior, y luego se han realizado otras operaciones para que las dos redes conviviesen en el mismo gráfico.

Como se puede observar, se trata de una red social multinivel ya que en un primer nivel están los nodos “Padres”, es decir las páginas web de las universidades, mientras en el segundo nivel están los nodos que representan las páginas web de sus departamentos de matemáticas.

Entre un nodo de una universidad y los nodos que representa su departamento de matemáticas, existe la relación padre-hijo.

En la siguiente imagen un zoom hacia el nodo que representa la universidad “ub.es” y el nodo “hijo”, que representa su departamento de matemática, “mat.ub.es”.

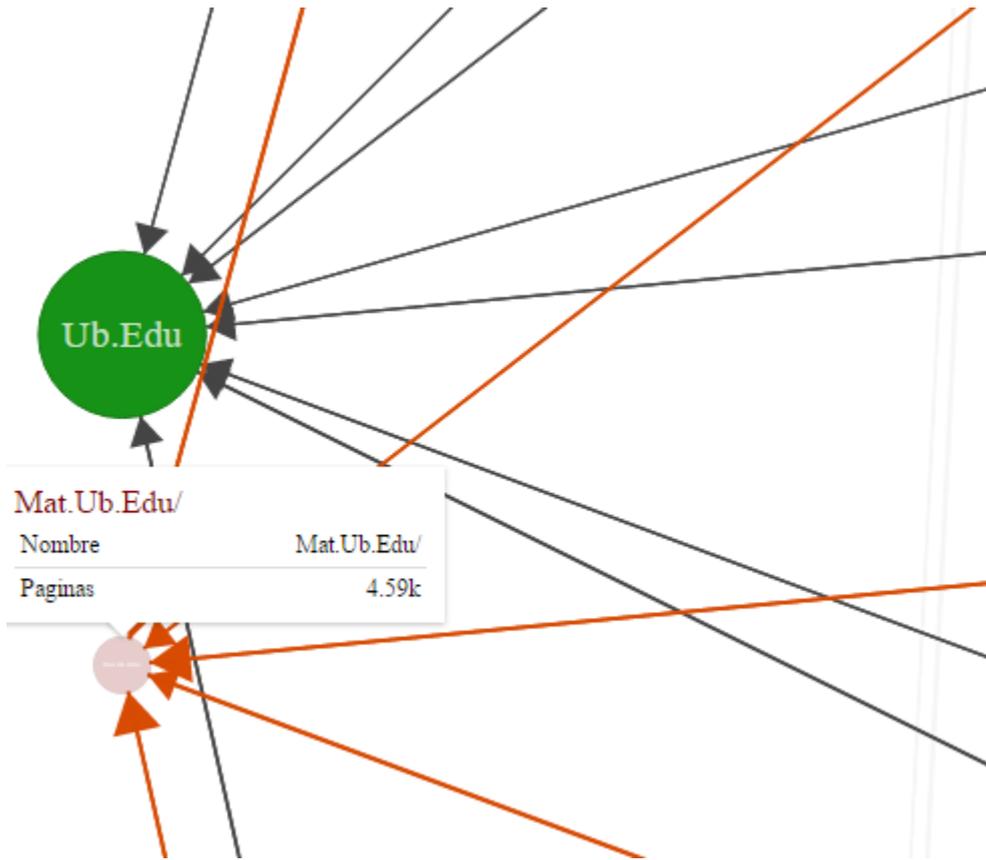


Ilustración 122 Zoom hacia dos nodos de dos niveles diferentes

6.3.1 Creación de los nodos de los departamentos de matemáticas

Como en el ejemplo anterior, se van a crear los nodos correspondientes a las URL de los departamentos de matemáticas de las 9 universidades españolas de la lista anterior.

Después de haber buscado en internet, se han encontrado 7 páginas web del departamento de matemáticas de las universidades, mientras en dos casos no se ha encontrado dicha página web, ya que las páginas con información sobre el departamento de matemática estaban en el nodo raíz (como si fuesen artículos de un blog).

El tamaño de los círculos será la cantidad de las páginas web de cada departamento.

Para obtener este valor, se va a utilizar como se ha visto precedentemente el comando de Google “Site:”, como se puede ver en la siguiente imagen.

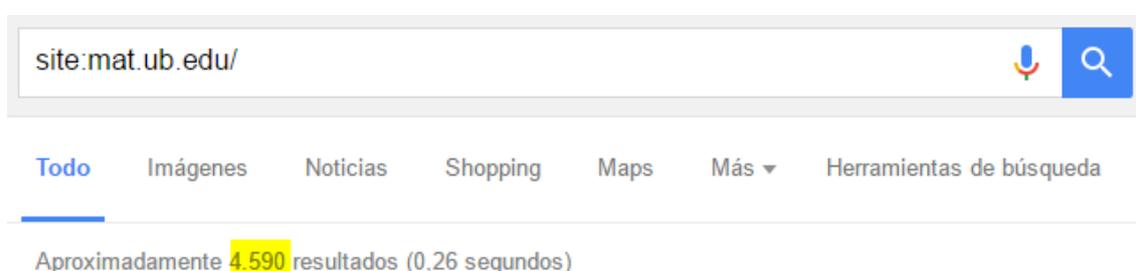


Ilustración 123 Ejecución de la query site:mat.ub.edu/

En la siguiente imagen se pueden ver las URL de cada departamento de matemáticas y el resultado de la query “Site:”.

| URL UNIVERSIDAD | URL DEPARTAMENTO MATEMATICAS | QUERY GOOGLE SITE: | RESULTADO DE LA QUERY |
|-----------------|-------------------------------------|---------------------------------|-----------------------|
| ub.edu | mat.ub.edu/ | site:mat.ub.edu/ | 4.590 |
| ucm.es | matematicas.ucm.es/ | site:matematicas.ucm.es/ | 4.650 |
| uab.cat | uab.cat NO HAY WEBSITE DEPARTAMENTO | NO HAY WEBSITE DEPARTAMENTO | 0 |
| uv.es | uv.es/uvweb/matematicas/ | site:uv.es/uvweb/matematicas/ | 853 |
| ugr.es | analisismatematico.ugr.es/ | site:analisismatematico.ugr.es/ | 374 |
| upv.es | upv.es/entidades/DMAA/ | site:upv.es/entidades/DMAA/ | 441 |
| uam.es | verso.mat.uam.es/ | site:verso.mat.uam.es/ | 2.550 |
| us.es | matematicas.us.es/ | site:matematicas.us.es/ | 2.950 |
| upm.es | upm.es NO HAY WEBSITE DEPARTAMENTO | NO HAY WEBSITE DEPARTAMENTO | 0 |

Ilustración 124 URL de cada departamento de matemáticas y el resultado de la query “Site:”

Llegados a este punto, se tienen que insertar estos datos en el array del gráfico con los nodos.

En el ejemplo anterior el color de los nodos era aleatorio, mientras en este ejemplo nos interesa distinguir las páginas web de las universidades de las de los departamentos de matemáticas.



Por esta razón se ha procedido en la siguiente manera:

- A cada nodo del gráfico, es decir a cada elemento del array “Datos_Universidades”, se ha añadido el campo “Tipo”, y en las universidades se ha dado el valor “Universidad” a este campo, mientras en el caso de los departamentos de matemáticas se le ha dado el valor “Matematicas”.

La declaración del array “Datos_Universidades” con todos los nodos se ha quedado de la siguiente forma:

```
var Datos_Universidades = [  
  {"Nombre": "ub.edu", "Paginas": 1900000, "Tipo": "Universidad"},  
  {"Nombre": "ucm.es", "Paginas": 1870000, "Tipo": "Universidad"},  
  {"Nombre": "uab.cat", "Paginas": 2030000, "Tipo": "Universidad"},  
  {"Nombre": "uv.es", "Paginas": 1560000, "Tipo": "Universidad"},  
  {"Nombre": "ugr.es", "Paginas": 1670000, "Tipo": "Universidad"},  
  {"Nombre": "upv.es", "Paginas": 1470000, "Tipo": "Universidad"},  
  {"Nombre": "uam.es", "Paginas": 690000, "Tipo": "Universidad"},  
  {"Nombre": "us.es", "Paginas": 3050000, "Tipo": "Universidad"},  
  {"Nombre": "upm.es", "Paginas": 1170000, "Tipo": "Universidad"},  
  {"Nombre": "mat.ub.edu/", "Paginas": 4590, "Tipo": "Matematicas"},  
  {"Nombre": "matematicas.ucm.es/", "Paginas": 4650, "Tipo": "Matematicas"},  
  {"Nombre": "uv.es/uvweb/matematicas/", "Paginas": 853, "Tipo": "Matematicas"},  
  {"Nombre": " analisismatematico.ugr.es/", "Paginas": 374, "Tipo": "Matematicas"},  
  {"Nombre": "upv.es/entidades/DMAA/", "Paginas": 441, "Tipo": "Matematicas"},  
  {"Nombre": "verso.mat.uam.es/", "Paginas": 2550, "Tipo": "Matematicas"},  
  {"Nombre": "matematicas.us.es/", "Paginas": 2950, "Tipo": "Matematicas"}  
]
```

- En la parte final de la página, se ha añadido el siguiente trozo de código, que permite asignar a los nodos el color verde (#008800 en hexadecimal) si el atributo “Tipo” es igual a “Universidad”, por lo contrario se le asigna el color rojo (#880000 en hexadecimal).

```
.color(function(d){  
  return d.Tipo == "Universidad" ? "#008800" : "#880000";  
})
```

Todavía no es posible enseñar los nodos, ya que en un gráfico de tipo “Network” de la librería D3plus si un nodo no tiene aristas no se visualiza.



6.3.2 Creación de las aristas del gráfico “network”

En este punto se va a proceder a crear las aristas del gráfico, que representan la cantidad y el sentido de los enlaces que hay entre los departamentos de matemáticas de las 7 universidades que estamos analizando.

6.3.2.1 Obtención de los valores de las aristas

Para obtener estos valores, se ha procedido de la misma manera que el ejemplo anterior, es decir:

- A través de la herramienta Majestic SEO, se han sacado los backlinks de las páginas web de los departamentos de matemáticas, es decir la lista de todas las páginas web que enlazan nuestros sitios web.
- Una vez obtenidos las 7 listas, se realizan 6 búsquedas en cada lista. En cada una de estas, se busca la URL de otro departamento de matemática, para ver si lo enlaza.

Por ejemplo, en el fichero con los backlinks de la página web mat.ub.edu/ se ha realizado la búsqueda “matematicas.ucm.es”, que da como resultados todos los enlaces provenientes del departamento de matemáticas de la UCM.ES que enlazan al departamento de matemática de la UB.EDU. Con el comando “contar”, se obtienen el número de los enlaces que se han encontrado.

En la siguiente imagen se puede ver una captura de pantalla tomada durante el proceso que se acaba de describir. Se puede notar que no se ha encontrado ningún enlace

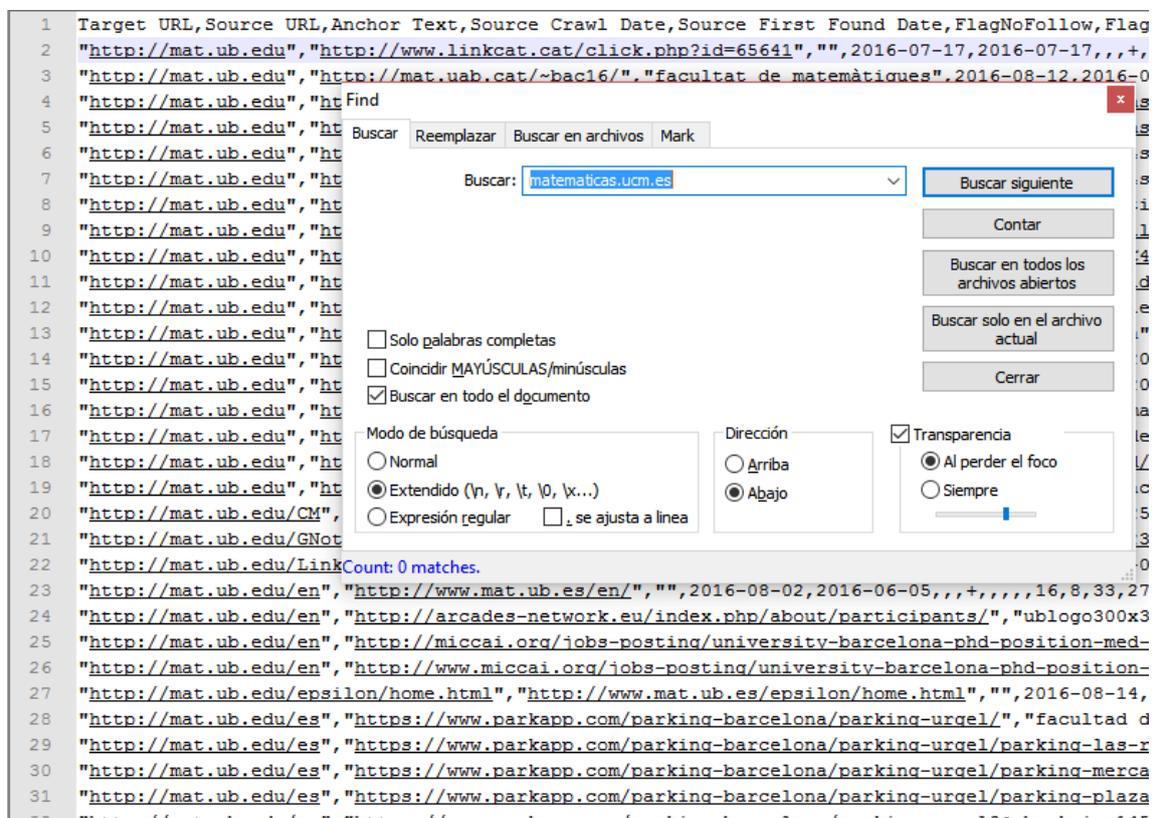


Ilustración 125 Búsqueda en el fichero con los backlinks de la página web mat.ub.edu/ de todos los enlaces provenientes de matematicas.ucm.es



Sin embargo, todas estas búsquedas han dado resultado cero, es decir no existen enlaces entre los departamentos de matemáticas de las 7 universidades en estudio.

Para poder seguir realizando el gráfico con la subred, se ha decidido asignar un valor random entre 300 y 3000 a cada arista del gráfico. Para poder generar los números aleatorios, se ha utilizado la función ALEATORIO.ENTRE(min,max) de Microsoft Excel.

En la siguiente imagen se puede ver una captura de pantalla de cómo queda el fichero Excel con los datos de los enlaces.

A continuación la descripción de las columnas:

- URL ORIGEN → URL de donde sale en enlace
- URL DESTINO → URL donde llega el enlace
- BACKLINKS_ALEATORIOS → Numero aleatorio generado con la función ALETORIO.ENTRE(300,3000) que representa el número de enlaces desde la URL ORIGEN hacia la URL DESTINO
- BACKLINKS_REALES → Numero de backlinks obtenido con la metodología explicada precedentemente

| URL ORIGEN | URL DESTINO | BACKLINKS_ALEATORIOS | BACKLINKS_REALES |
|----------------------------|----------------------------|----------------------|------------------|
| matematicas.ucm.es/ | mat.ub.edu/ | 1445 | 0 |
| uv.es/uvweb/matematicas/ | mat.ub.edu/ | 655 | 0 |
| analisismatematico.ugr.es/ | mat.ub.edu/ | 1253 | 0 |
| upv.es/entidades/DMAA/ | mat.ub.edu/ | 1908 | 0 |
| verso.mat.uam.es/ | mat.ub.edu/ | 702 | 0 |
| mat.ub.edu/ | matematicas.ucm.es/ | 1598 | 0 |
| uv.es/uvweb/matematicas/ | matematicas.ucm.es/ | 796 | 0 |
| analisismatematico.ugr.es/ | matematicas.ucm.es/ | 602 | 0 |
| upv.es/entidades/DMAA/ | matematicas.ucm.es/ | 2098 | 0 |
| verso.mat.uam.es/ | matematicas.ucm.es/ | 2655 | 0 |
| mat.ub.edu/ | uv.es/uvweb/matematicas/ | 2797 | 0 |
| matematicas.ucm.es/ | uv.es/uvweb/matematicas/ | 1960 | 0 |
| analisismatematico.ugr.es/ | uv.es/uvweb/matematicas/ | 2026 | 0 |
| upv.es/entidades/DMAA/ | uv.es/uvweb/matematicas/ | 2686 | 0 |
| verso.mat.uam.es/ | uv.es/uvweb/matematicas/ | 919 | 0 |
| mat.ub.edu/ | verso.mat.uam.es/ | 742 | 0 |
| uv.es/uvweb/matematicas/ | verso.mat.uam.es/ | 1010 | 0 |
| analisismatematico.ugr.es/ | verso.mat.uam.es/ | 2449 | 0 |
| upv.es/entidades/DMAA/ | verso.mat.uam.es/ | 2151 | 0 |
| matematicas.ucm.es/ | verso.mat.uam.es/ | 936 | 0 |
| mat.ub.edu/ | upv.es/entidades/DMAA/ | 972 | 0 |
| uv.es/uvweb/matematicas/ | upv.es/entidades/DMAA/ | 2493 | 0 |
| analisismatematico.ugr.es/ | upv.es/entidades/DMAA/ | 677 | 0 |
| matematicas.ucm.es/ | upv.es/entidades/DMAA/ | 1909 | 0 |
| verso.mat.uam.es/ | upv.es/entidades/DMAA/ | 1280 | 0 |
| mat.ub.edu/ | analisismatematico.ugr.es/ | 2004 | 0 |
| uv.es/uvweb/matematicas/ | analisismatematico.ugr.es/ | 1529 | 0 |
| upv.es/entidades/DMAA/ | analisismatematico.ugr.es/ | 2920 | 0 |

Ilustración 126 Captura de pantalla del Excel con los valores de las aristas de la subred de websites de departamentos de matemáticas

6.3.2.2 Modificación del formato con un mapa para que se puedan insertar en la librería

A este punto, utilizando el mapa creado para los ejemplos precedentes, se han añadido al array connections[] las aristas de los nodos de la subred de departamentos de matemática.

En detalle, el mapa realiza la operación descrita en la siguiente imagen:

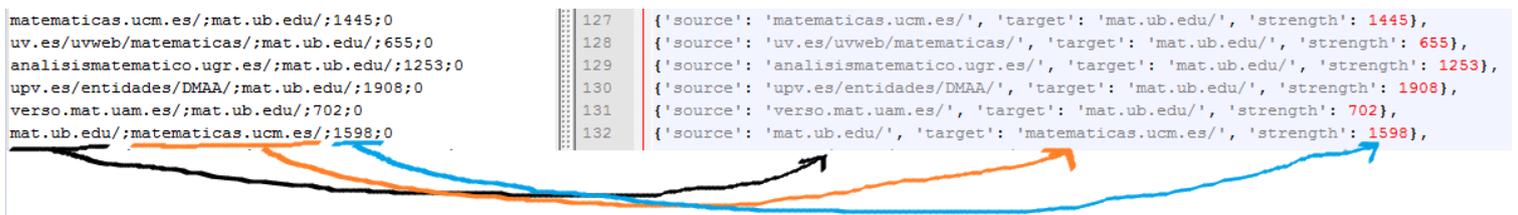


Ilustración 127 Explicación de la función del mapa para insertar las aristas en el array

Las nuevas aristas se han añadido en el array connections[] después de las aristas de las web de la universidades, como se puede ver en la siguiente captura de pantalla

```
{'source': 'uab.cat', 'target': 'upm.es', 'strength': 32},
{'source': 'uv.es', 'target': 'upm.es', 'strength': 75},
{'source': 'ugr.es', 'target': 'upm.es', 'strength': 154},
{'source': 'upv.es', 'target': 'upm.es', 'strength': 552},
{'source': 'uam.es', 'target': 'upm.es', 'strength': 274},
{'source': 'us.es', 'target': 'upm.es', 'strength': 108},
{'source': 'ub.edu', 'target': 'upm.es', 'strength': 94},
{'source': 'matematicas.ucm.es/', 'target': 'mat.ub.edu/', 'strength': 1445},
{'source': 'uv.es/uvweb/matematicas/', 'target': 'mat.ub.edu/', 'strength': 655},
{'source': 'analisismatematico.ugr.es/', 'target': 'mat.ub.edu/', 'strength': 1253},
{'source': 'upv.es/entidades/DMAA/', 'target': 'mat.ub.edu/', 'strength': 1908},
{'source': 'verso.mat.uam.es/', 'target': 'mat.ub.edu/', 'strength': 702},
{'source': 'mat.ub.edu/', 'target': 'matematicas.ucm.es/', 'strength': 1598},
{'source': 'uv.es/uvweb/matematicas/', 'target': 'matematicas.ucm.es/', 'strength': 796},
{'source': 'analisismatematico.ugr.es/', 'target': 'matematicas.ucm.es/', 'strength': 602},
{'source': 'upv.es/entidades/DMAA/', 'target': 'matematicas.ucm.es/', 'strength': 2098},
{'source': 'verso.mat.uam.es/', 'target': 'matematicas.ucm.es/', 'strength': 2655},
{'source': 'mat.ub.edu/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 2797},
{'source': 'matematicas.ucm.es/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 1960},
{'source': 'analisismatematico.ugr.es/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 2026},
{'source': 'upv.es/entidades/DMAA/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 2686},
{'source': 'verso.mat.uam.es/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 919},
{'source': 'mat.ub.edu/', 'target': 'matematicas.ucm.es/', 'strength': 843}
```

Ilustración 128 Aristas del array connections[]

La parte final del código html es exactamente igual que en el ejemplo anterior, así que no se explicará en este apartado.

El código completo se puede encontrar en el Anexo III.

6.3.3 Gráficos resultantes

En la siguiente imagen se puede ver la representación de los gráfico network que acabamos de crear.

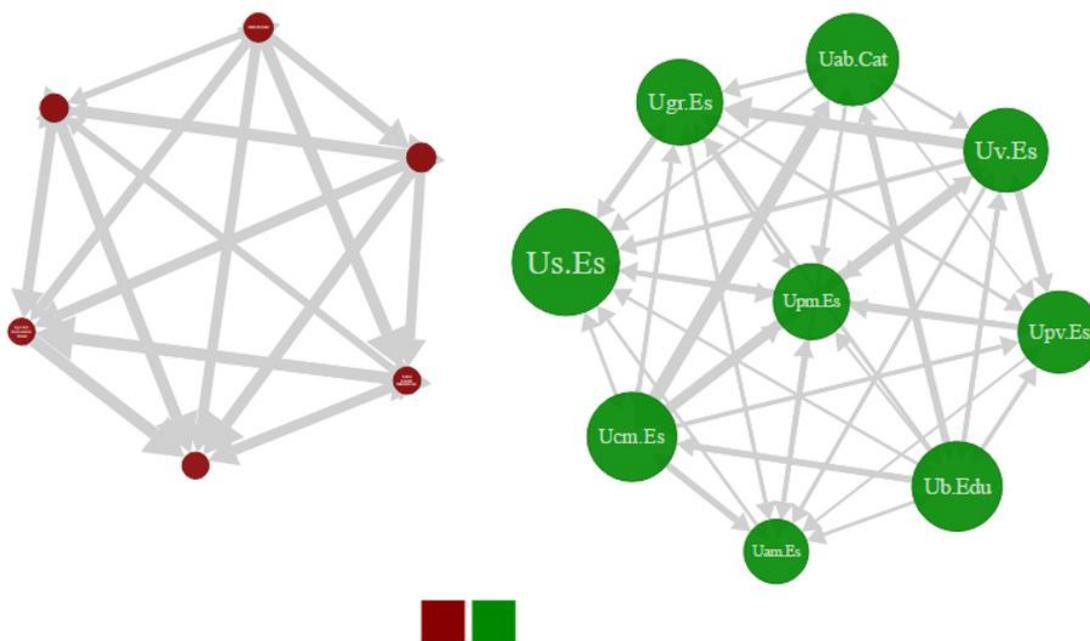


Ilustración 129 Gráficos de tipo network de los enlaces entre las 9 universidades y sus departamentos de matemáticas

La librería gráfica D3plus ha generado los dos gráficos. En rojo, como configuramos en las opciones del gráfico, están los nodos de los departamentos de matemáticas con sus enlaces, mientras en verde el gráfico del ejemplo anterior, es decir las web de las 9 universidades en estudio.

Como ya se observó anteriormente, ya que no hemos especificado una posición fija de los nodos, la librería ha generado los dos gráficos por separado, ya que tiene como objetivo evitar que las aristas y los nodos se superpongan.

Como en el caso anterior, cuando se pasa el ratón encima de un nodo, se evidencian sus las aristas que llegan y salen de dicho nodo y sale una pequeña ventana indicando su nombre y el número de páginas que contiene la página web.

En la siguiente imagen se puede ver esta operación.

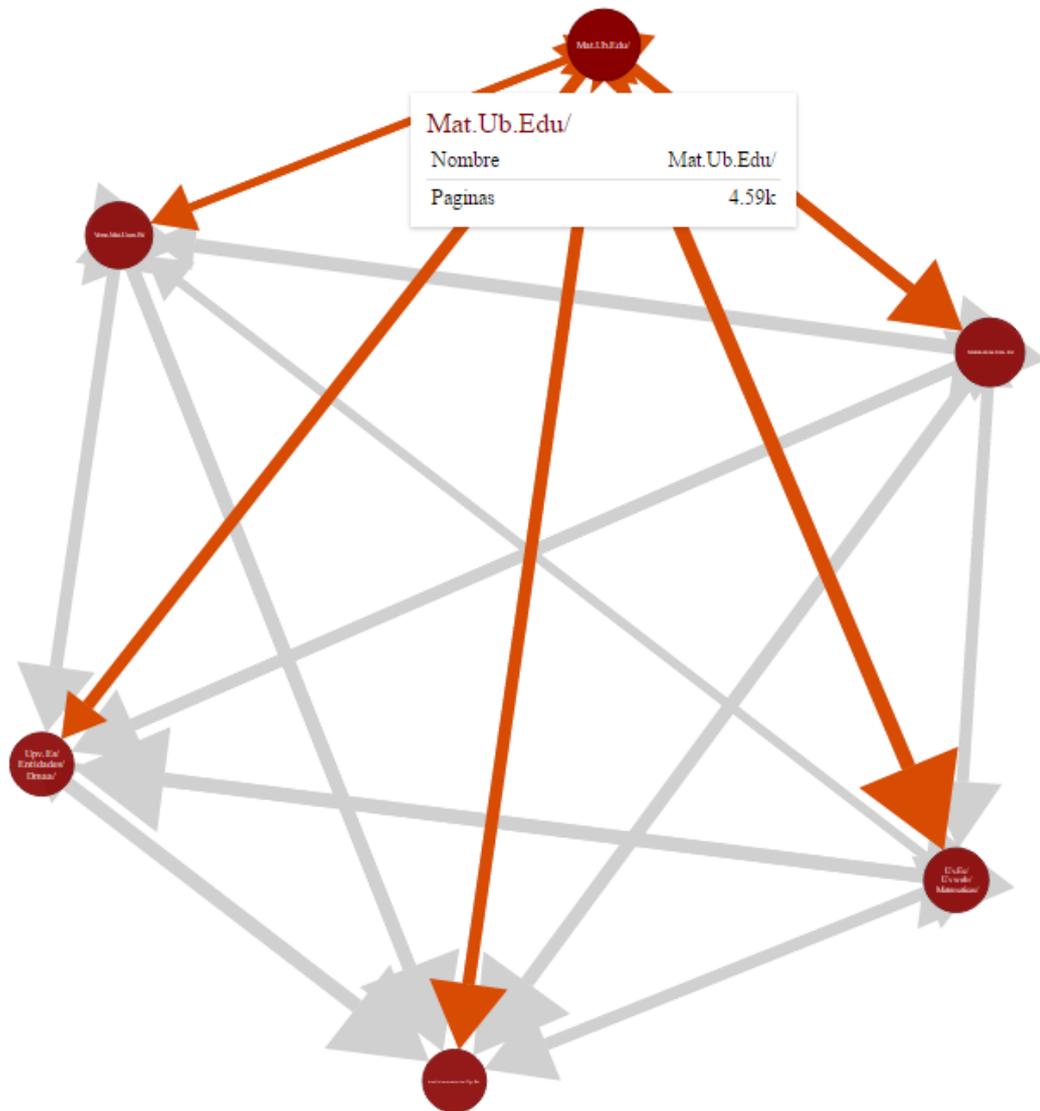


Ilustración 130 Respuesta del gráfico de los departamentos de matemáticas cuando se pasa el ratón encima de un nodo, en este caso mat.ub.edu

Como en el caso anterior, cuando se hace click encima de un nodo, además de hacer un zoom y de evidenciar sus conexiones, sale una pequeña ventana con la lista de los nodos con los cuales es conectado, como se puede ver en la siguiente captura de pantalla.

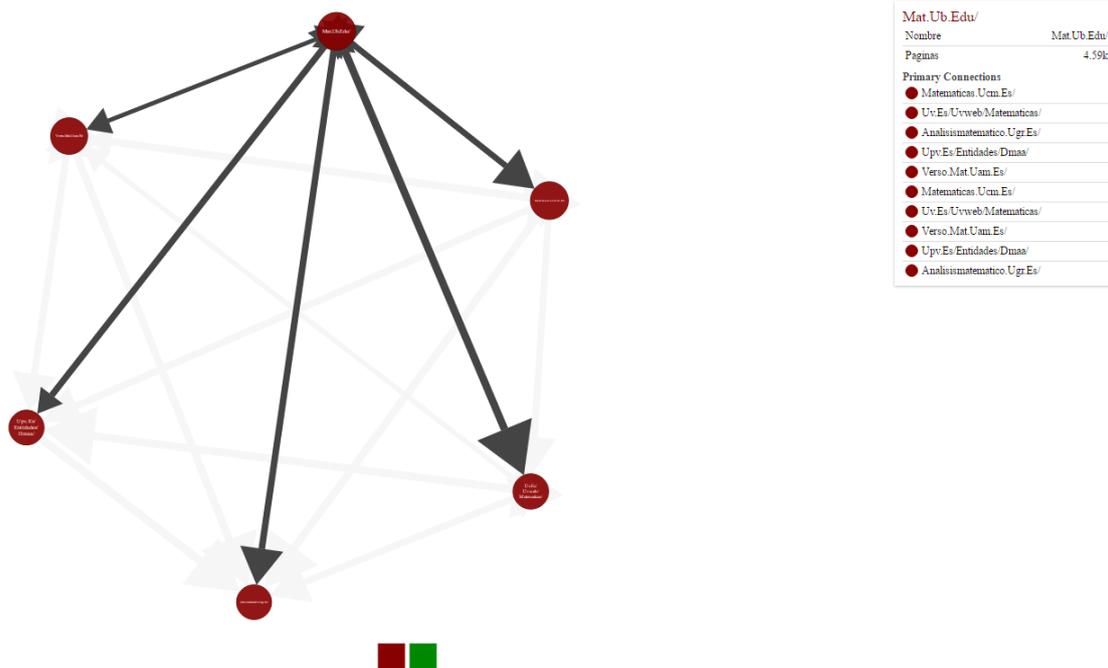


Ilustración 131 Respuesta del gráfico de los departamentos de matemáticas cuando se hace click encima de un nodo, en este caso mat.ub.edu

Ya que hay dos redes en el mismo gráfico, la librería añade en la parte inferior de la pantalla dos cuadraditos de los colores de los nodos, que tienen la función de leyenda.

Pasando el ratón encima de ellos, aparece una ventanita con la lista de los nodos de este color, como se puede ver en la siguiente imagen.

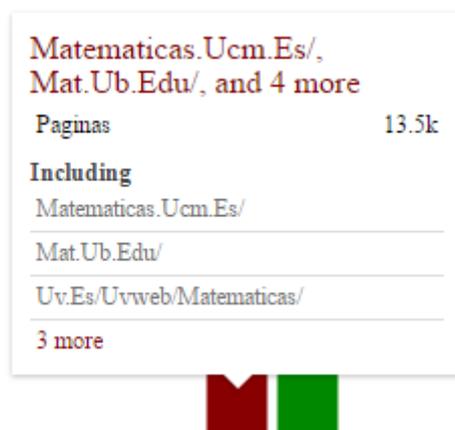


Ilustración 132 Leyenda de los colores de las redes

6.3.3.1 Superposición de los gráficos fijando las posiciones de los nodos

Como se nota en la imagen 125, la librería ha creado los gráficos correctamente, sin embargo no se ve con claridad que son redes en dos niveles diferentes, es decir en lugar de ser redes que contienen información parecida, se visualizan como si fuesen redes totalmente diferente y es casi imposible cruzar la información.

Para resolver este problema, se ha decidido fijar los nodos. Como ya se dijo anteriormente, la librería ha intentado de “alejar” las dos redes para evitar que los nodos y las aristas se superpusiesen.

Vamos a intentar fijar los nodos a coordenadas que se le pasan, para poder ver los gráficos en el mismo sitio y poder realizar comparaciones.

Para realizar esta operación, hay que crear el array “positions[]” y añadir los campos “x” e “y” para cada nodo. Se ha decidido que los nodos de los departamentos de matemáticas tienen el mismo valor en el axis x pero una unidad más en el axis y, de esta manera los nodos de la misma universidad se colocan cerca.

Hay que observar que la librería empieza a contar los valores de los axis x e y desde el margen de arriba a la izquierda, así que por ejemplo incrementando el valor del axis y, el nodo bajará.

En la siguiente imagen, los axis con la orientación observada por esta librería.

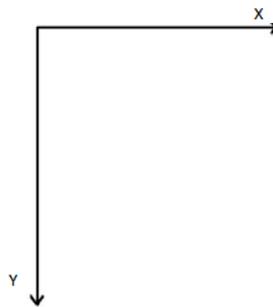


Ilustración 133 Orientación de los axis observada por la librería D3plus

El nuevo array positions[] queda entonces de la siguiente manera:

```
var positions = [  
  {"Nombre": "ub.edu", "x": 10, "y": 15},  
  {"Nombre": "ucm.es", "x": 12, "y": 24},  
  {"Nombre": "uab.cat", "x": 16, "y": 18},  
  {"Nombre": "uv.es", "x": 26, "y": 21},  
  {"Nombre": "ugr.es", "x": 13, "y": 4},  
  {"Nombre": "upv.es", "x": 31, "y": 13},  
  {"Nombre": "uam.es", "x": 19, "y": 8},  
  {"Nombre": "us.es", "x": 24, "y": 11},  
  {"Nombre": "upm.es", "x": 20, "y": 5},  
  {"Nombre": "mat.ub.edu/", "x": 10, "y": 16},  
  {"Nombre": "matematicas.ucm.es/", "x": 12, "y": 25},  
  {"Nombre": "uv.es/uvweb/matematicas/", "x": 26, "y": 22},  
  {"Nombre": "analisismatematico.ugr.es/", "x": 13, "y": 5},  
  {"Nombre": "upv.es/entidades/DMAA/", "x": 31, "y": 14},  
  {"Nombre": "verso.mat.uam.es/", "x": 19, "y": 9}  
]
```

Ilustración 134 Declaración del Array positions[]

En la parte final del fichero html, se pasará este array a los nodos con el siguiente código

`.nodes(positions)`

El gráfico con las dos redes superpuestas queda de la siguiente manera:

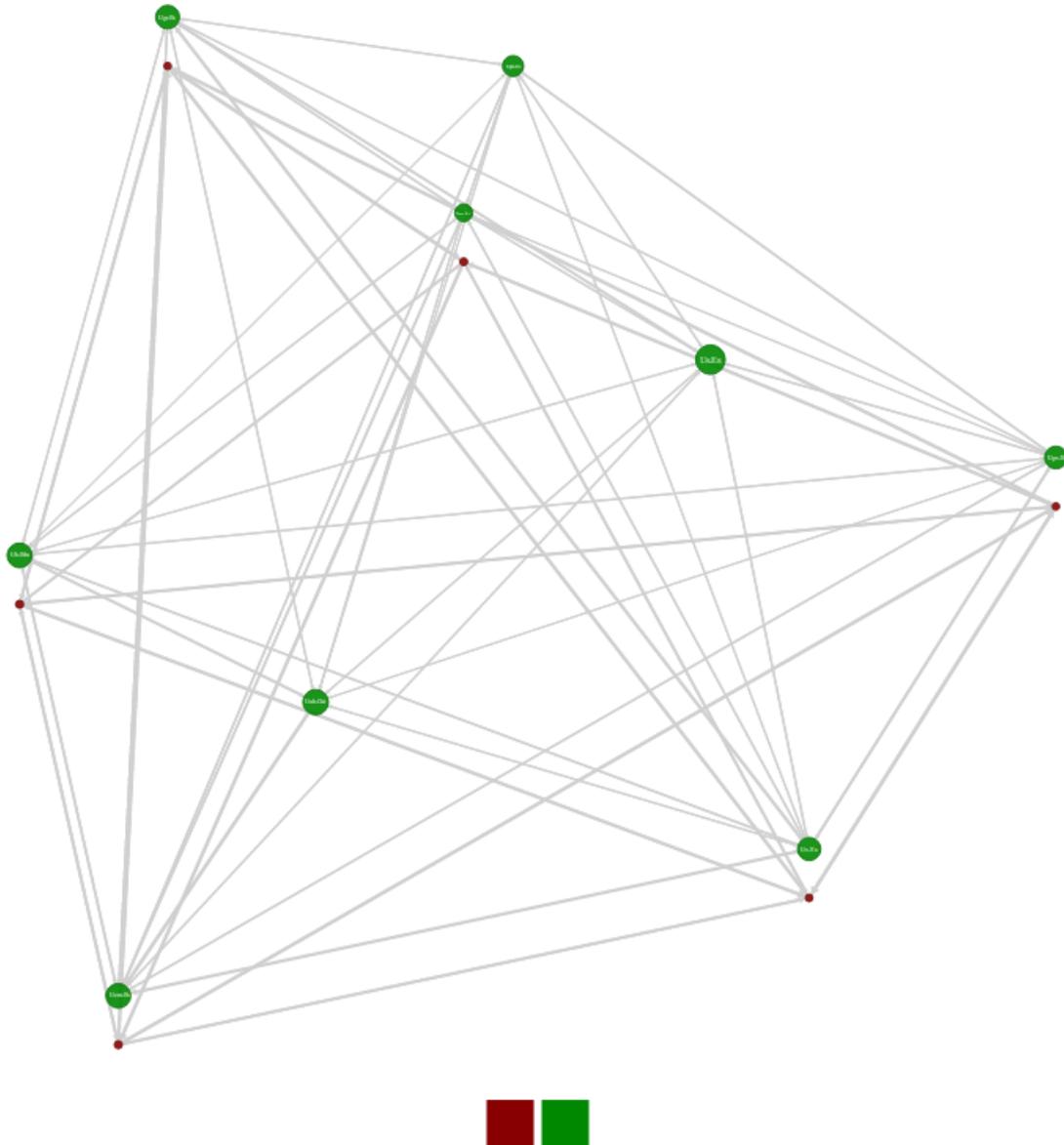


Ilustración 135 Gráfico con las dos redes superpuestas

De esta manera se pueden comparar los elementos de las dos redes.

Por ejemplo, para comparar hacia donde van los enlaces de la upv.es y su departamento de matemáticas upv.es/entidades/dmaa, solo hay que hacer click en uno de los dos nodos y pasar el ratón encima del otro.

El resultado se puede observar en la siguiente imagen.

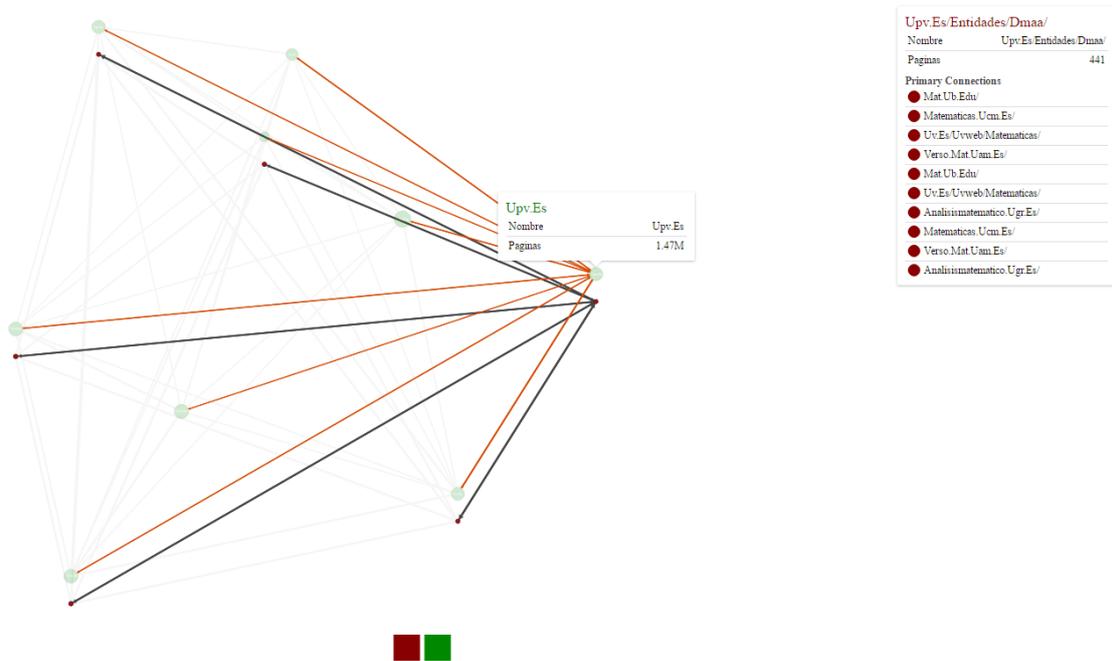


Ilustración 136 Comparación entre el nodo upv.es y su departamento de matemáticas, el nodo upv.es/entidades/dmaa

6.3.3.2 Errores de la librería

Como se puede notar en las imágenes anteriores, cuando se superponen dos gráficos de network la librería ya no visualiza el grosor de las aristas, aunque siguen habilitadas con el comando "size" (en negrita en el siguiente trozo de código).

```
.edges{{
    "size": "strength",
    "label": "strength",
    "value": connections,
    "arrows": true
}}
```

Además, se ha intentado también dibujar el valor de las aristas con el comando "label", sin embargo la librería tampoco responde correctamente.

7 Conclusiones

El objetivo general de este trabajo es identificar la tipología de visualización de datos más adecuada al análisis de redes sociales multinivel.

Este objetivo se ha cumplido, ya que se ha encontrado un tipo de visualización adapta para representar las redes sociales multinivel, como se vio en los ejemplos durante la aplicación de la visualización de datos.

Para poder llegar al cumplimiento de este objetivo, se han tenido que realizar las siguientes fases, cada una necesaria para poder desarrollar la fase siguiente.

Estas fases se describieron en los objetivos específicos, que a continuación se analizarán singularmente.

El primer objetivo específico es la descripción de la evolución de la visualización de la información. Este objetivo era necesario para poder informarse y analizar todo lo que podía ofrecer la visualización de información, la manera de visualizar ciertos tipos de datos en la época actual y en el pasado.

Este objetivo específico se ha cumplido, realizando en la primera parte del trabajo una amplia análisis de la visualización de la información. Se empezó describiendo la información en sí, sobre todo enfocándose hacia el fenómeno del Big Data y analizando la manera de visualizarlo, orientándose hacia los diferentes tipos de infografías y de gráficos.

El segundo objetivo específico era la comparación de las características de los programas para el análisis de redes sociales.

Este objetivo también se cumplió. Antes de todo se definieron las redes sociales, para luego poder realizar un estudio comparativo de los principales programas para el análisis de las mismas.

En tercer y último objetivo específico implicaba la realización de unos ejemplos de visualización gráfica interactiva de redes sociales multinivel. Este objetivo también se ha cumplido, en particular se ha utilizado la librería gráfica D3plus ya que ha resultado ser la mejor para nuestro objetivo.

Se ha elegido la tipología de gráfico network gracias a todo el trabajo realizado anteriormente, ya que se han podido comparar todas las tipologías de visualización de datos analizadas y se ha podido elegir la mejor.

En particular, se ha elegido una visualización interactiva ya que de esta manera el usuario puede navegar en el gráfico y enfocar sobre los datos que le interesan, cosa que no permite una visualización estática o una infografía, que enseñan al usuario lo que ha querido enfocar su creador.



Los ejemplos se han desarrollado en el contexto del mundo universitario español, en particular se ha creado una red social multinivel que representa la cantidad de enlaces entre las páginas web de las universidades y, en el último ejemplo, se han añadido a la red sus departamentos de matemáticas, para crear una visualización de red social interactiva multinivel.

Se han utilizado las web de nueve de las universidades españolas más importantes para poder estar seguros que tenían un amplio volumen de tráfico de datos, para poderlas analizar.

Esta elección resulto ser correcta, sobre todo porque en algunos casos no se han encontrado suficientes datos para generar un ejemplo y se han tenido que generar números aleatorios. Esto hace pensar que con pginas web de universidades menores, se habrían tenido muchas más dificultades para visualizar los datos.

Se puede decir que sobre uno de los argumentos principales de la tesis, las redes sociales multinivel, no se ha encontrado mucha información en internet.

Por esta razón se ha querido profundizar la visualización interactiva de estos tipos de redes desarrollando y analizando un ejemplo práctico, y me gustaría en investigaciones futuras seguir analizando este argumento, ya que como se ha dicho hay muy pocos trabajos de investigación y es un tipo de gráfico con el cual se pueden analizar muchos casos reales.

Por ejemplo, una red social multinivel podría ser compuesta por las sociedades españolas de telefonía, y en un nivel más bajo sus sub-empresas de telefonía móvil y proveedores de internet. Las aristas de esta red entre los nodos podrían ser el número de menciones de las empresas en artículos de blog o de revistas especializadas.

En este caso, en una red social multinivel, cada nodo principal sería una empresa de telefonía, y haciendo un “zoom” se podrían ver como sub-nodos las empresas de su propiedad.

Con este tipo de gráfico se podrían realizar muchos análisis de mercado, por ejemplo se vería claramente si hay alianzas entre algunas empresas.

8 Referencias

- [1] Geodan Madrid, «Geodan,» 02 2015. [En línea]. Available: <http://www.geodan.es/big-data-y-gis/>. [Último acceso: 02 2016].
- [2] Cisco Systems, Inc, «cisco.com,» Enero 2015. [En línea]. Available: http://www.cisco.com/c/en/us/solutions/service-provider/global-cloud-index-gci/index.html?CAMPAIGN=GCI+2015&COUNTRY_SITE=us&POSITION=Press+Release&REFERRING_SITE=PR&CREATIVE=PR+to+GCI+page&_ga=1.268100174.1385830356.1459120871. [Último acceso: Febrero 2016].
- [3] Forbes, «Forbes Europe,» 2015. [En línea]. Available: <http://www.forbes.com/forbes/welcome/#36c2cdde32d9>. [Último acceso: Febrero 2016].
- [4] jobandtalent.com, «jobandtalent.com,» Octubre 2015. [En línea]. Available: <http://blog.jobandtalent.com/profesionales-big-data-demandados/>. [Último acceso: Febrero 2016].
- [5] A. Kirk, «visualisingdata.com,» Junio 2012. [En línea]. Available: <http://www.visualisingdata.com/2012/06/article-the-8-hats-of-data-visualisation-design/>. [Último acceso: April 2016].
- [6] Wikipedia, «El mapa del cólera de John Snow,» Enero 2016. [En línea]. Available: https://es.wikipedia.org/wiki/John_Snow#El_mapa_del_c.C3.B3lera_de_John_Snow. [Último acceso: April 2016].
- [7] Wikipedia, «Infografía,» April 2016. [En línea]. Available: <https://es.wikipedia.org/wiki/Infograf%C3%ADa>. [Último acceso: April 2016].
- [8] pinterest, «pinterest.com/,» [En línea]. Available: <https://es.pinterest.com/>. [Último acceso: 01 09 2016].
- [9] D. McCandless, La Información Es Bella, Primera ed., RBA, 2014.
- [10] B. Shaler, «<http://datarealization.com/>,» Octubre 2015. [En línea]. [Último acceso: April 2016].
- [11] wikipedia.org, «Edward Tufte,» [En línea]. Available: https://es.wikipedia.org/wiki/Edward_Tufte. [Último acceso: 01 09 2016].
- [12] Medialab-Prado, «Edward Tufte: Envisioning Information. Capítulo 1. Escapando de la Planicie | Mediación Medialab-Prado,» [En línea]. Available: <http://mediacion.medialab-prado.es/2007/11/21/edward-tufte-envisioning-information-capitulo-1-escapando-de-la-planicie/>. [Último acceso: 01 09 2016].
- [13] Target Marketing, «"Power Corrupts. PowerPoint Corrupts Absolutely."--Edward Tufte,» [En línea]. Available: <http://www.targetmarketingmag.com/article/power->



- corrupts-powerpoint-corrupts-absolutely-edward-tufte-23351/. [Último acceso: 01 09 2016].
- [14] Wikipedia, «Chartjunk,» [En línea]. Available: <https://en.wikipedia.org/wiki/Chartjunk>. [Último acceso: 01 09 2016].
- [15] E. Tufte, *The Visual Display of Quantitative Information*, Graphics Pr; 2nd edition (January 2001), 1983.
- [16] Wikipedia, «Sparkline,» [En línea]. Available: <https://en.wikipedia.org/wiki/Sparkline>. [Último acceso: 01 09 2016].
- [17] Es.wikipedia.org, «Siniestro del transbordador espacial Challenger,» [En línea]. Available: https://es.wikipedia.org/wiki/Siniestro_del_transbordador_espacial_Challenger. [Último acceso: 01 09 2016].
- [18] E. Tufte, «Edward Tufte forum: PowerPoint Does Rocket Science--and Better Techniques for Technical Reports,» [En línea]. Available: http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001yB. [Último acceso: 01 09 2016].
- [19] Infovis-wiki.net, «Few, Stephen - InfoVis:Wiki,» [En línea]. Available: http://www.infovis-wiki.net/index.php?title=Few%2C_Stephen. [Último acceso: 01 09 2016].
- [20] S. Few, *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, Gazelle Book Services, 2004.
- [21] CIORReview, «cioreview.com,» Enero 2016. [En línea]. Available: <http://magazine.cioreview.com/magazines/January2016/Data-Visualization/>. [Último acceso: April 2016].
- [22] Ernesto Olivares Visual Information, «Ernesto Olivares Visual Information: INFOGRAFIAS,» [En línea]. Available: <https://ernestoolivares.es/>. [Último acceso: 01 09 2016].
- [23] IBM, «IBM,» Enero 2015. [En línea]. Available: <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>. [Último acceso: April 2016].
- [24] Fundacionseres.org, «Big data: una "revolución industrial" en la gestión de los datos digitales,» [En línea]. Available: <http://www.fundacionseres.org/Paginas/Campus/Informes.aspx?IDi=960>. [Último acceso: 01 09 2016].
- [25] podio.com, «The daily routines of famous creative people,» 2015. [En línea]. Available: <https://podio.com/site/creative-routines>.

- [26] Wikipedia, «Statistical graphics,» 2015. [En línea]. Available: https://en.wikipedia.org/wiki/Statistical_graphics. [Último acceso: 09 Mayo 2016].
- [27] INe, «<http://www.ine.es>,» 2016. [En línea]. Available: http://www.ine.es/explica/explica_pasos_tipos_graficos.htm.
- [28] C. L. & G. D. F. Streeter, «Social Network Analysis,» *Journal of Social Service Research*, 1992.
- [29] K. F. Stanley Wasserman, «Social Network Analysis: Methods and Applications (Structural Analysis in the Social Sciences),» 1994.
- [30] S. Y. David Knoke, «Social Network Analysis (Quantitative Applications in the Social Sciences),» 2007.
- [31] Wikipedia, «Boundary (topology),» [En línea]. Available: [https://en.wikipedia.org/wiki/Boundary_\(topology\)](https://en.wikipedia.org/wiki/Boundary_(topology)). [Último acceso: 01 09 2016].
- [32] M. y. M. P. Castells, «Prologue to the Special Section. International Journal of Communication,» de *NETWORK MULTIDIMENSIONALITY IN THE*, 2011.
- [33] F. Requena Santos, «El concepto de red social,» Reis, 1989, p. 141.
- [34] Á. RAYÓN, «Deusto BigData,» 12 2015. [En línea]. Available: <https://blogs.deusto.es/bigdata/analisis-de-redes-sociales-el-poder-de-la-teoria-de-grafos/>. [Último acceso: 2016].
- [35] Wikipedia, «Sociometría,» 2016. [En línea]. Available: <https://es.wikipedia.org/wiki/Sociometr%C3%ADa>. [Último acceso: 2016].
- [36] Wikipedia, «Jacob Levy Moreno,» 2016. [En línea]. Available: https://es.wikipedia.org/wiki/Jacob_Levy_Moreno. [Último acceso: 2016].
- [37] D. C. A. F. HARARY, «STRUCTURAL BALANCE: A GENERALIZATION,» de *THE PSYCHOLOGICAL REVIEW*, 1956.
- [38] UCINET Software, «UCINET Software,» [En línea]. Available: <https://sites.google.com/site/ucinetsoftware/home>. [Último acceso: 09 01 2016].
- [39] Networks / Pajek, «Networks / Pajek,» [En línea]. Available: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>. [Último acceso: 01 09 2016].
- [40] GUESS: The Graph Exploration System , «GUESS: The Graph Exploration System,» [En línea]. Available: <http://graphexploration.cond.org/>. [Último acceso: 01 09 2016].
- [41] «Gephi - The Open Graph Viz Platform,» [En línea]. Available: <https://gephi.org/>. [Último acceso: 01 09 2016].

- [42] Netminer.com, «NetMiner - Social Network Analysis Software,» [En línea]. Available: <http://www.netminer.com/main/main-read.do>. [Último acceso: 01 09 2016].
- [43] Wikipedia, «Social network analysis software,» [En línea]. Available: https://en.wikipedia.org/wiki/Social_network_analysis_software. [Último acceso: 01 09 2016].
- [44] Igraph.org, «Welcome to igraph's new home,» [En línea]. Available: <http://igraph.org/redirect.html>. [Último acceso: 01 09 2016].
- [45] Statnet, «Statnet.org,» [En línea]. Available: <http://www.statnet.org/>. [Último acceso: 01 09 2016].
- [46] D3js.org, «D3.js - Data-Driven Documents,» [En línea]. Available: <https://d3js.org/>. [Último acceso: 01 09 2016].
- [47] Datadrivenjournalism.net, «Data-Driven Documents, Defined,» [En línea]. Available: http://datadrivenjournalism.net/resources/data_driven_documents_defined. [Último acceso: 30 08 2016].
- [48] reddit, «Hi! I'm Mike Bostock, creator of D3.js and a former graphics editor for The New York Times. I do data visualization, design and open source.,» [En línea]. Available: https://www.reddit.com/r/dataisbeautiful/comments/3k3if4/hi_im_mike_bostock_creator_of_d3js_and_a_former/. [Último acceso: 01 09 2016].
- [49] Nieman Lab, «Data visualization rock star Mike Bostock is leaving The New York Times,» [En línea]. Available: <http://www.niemanlab.org/reading/data-visualization-rock-star-mike-bostock-is-leaving-the-new-york-times/>. [Último acceso: 01 09 2016].
- [50] The New York Times, «The New York Times,» [En línea]. Available: <http://elections.nytimes.com/2012/results/president/scenarios>. [Último acceso: 30 08 2016].
- [51] New York Times, «New York Times,» [En línea]. Available: http://www.nytimes.com/interactive/2013/04/08/business/global/asia-map.html?_r=0. [Último acceso: 30 08 2016].
- [52] Nytimes.com, «Among the Oscar Contenders, a Host of Connections,» [En línea]. Available: <http://www.nytimes.com/interactive/2013/02/20/movies/among-the-oscar-contenders-a-host-of-connections.html>. [Último acceso: 30 08 2016].
- [53] Wikipedia, «Wikipedia,» [En línea]. Available: https://en.wikipedia.org/wiki/List_of_S%26P_500_companies. [Último acceso: 30 08 2016].



- [54] The New York Times, «The New York Times,» 25 05 2013. [En línea]. Available: <http://www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html>. [Último acceso: 30 08 2016].
- [55] D3plus, «D3plus.org,» [En línea]. Available: <https://d3plus.org/>. [Último acceso: 30 08 2016].
- [56] EDICOM, «Edicomgroup.com,» [En línea]. Available: http://www.edicomgroup.com/es_AR/solutions/edi/software/Edicom-Mapping-Tool. [Último acceso: 30 08 2016].
- [57] Edicom, «Edicom España | Edicom,» [En línea]. Available: http://www.edicomgroup.com/es_ES/home.html. [Último acceso: 01 09 2016].
- [58] Wikipedia, «EDIFACT,» [En línea]. Available: <https://es.wikipedia.org/wiki/EDIFACT>. [Último acceso: 01 09 2016].
- [59] Webometrics, «Webometrics.info,» [En línea]. Available: <http://www.webometrics.info>. [Último acceso: 30 08 2016].
- [60] Webometrics.info, «Webometrics.info,» [En línea]. Available: http://www.webometrics.info/es/europe_es/espa%C3%B1a. [Último acceso: 30 08 2016].
- [61] Google, «Support.google.com,» [En línea]. Available: <https://support.google.com/websearch/answer/2466433?hl=es>. [Último acceso: 30 08 2016].
- [62] Limoncomunicacion.com, «Limoncomunicacion.com,» [En línea]. Available: <http://limoncomunicacion.com/2016/01/02/que-es-un-backlink-y-para-que-sirve/>. [Último acceso: 30 08 2016].
- [63] Majestic, «Es.majestic.com,» [En línea]. Available: <https://es.majestic.com/support/about>. [Último acceso: 30 08 2016].
- [64] Editpadlite.com, «EditPad Lite - Free Text Editor for Windows,» [En línea]. Available: <https://www.editpadlite.com/>. [Último acceso: 30 08 2016].
- [65] Webometrics.info, «Webometrics.info,» [En línea]. Available: http://www.webometrics.info/es/europe_es/espa%C3%B1a. [Último acceso: 30 08 2016].



Anexo I: Código del gráfico Tree map con timeline que representa el número de páginas web de escuelas de las universidades españolas en años diferentes

Fichero index.html

```
<!doctype html>
<meta charset="utf-8">
<script src="http://www.d3plus.org/js/d3.js"></script>
<script src="http://www.d3plus.org/js/d3plus.js"></script>
<div id="viz"></div>
<script>
var Datos_Escuelas = [
    {'Año': 2011, 'Nombre': 'Universidad de Castilla-La Mancha', 'Escuelas': 20},
    {'Año': 2011, 'Nombre': 'Universidad del País Vasco', 'Escuelas': 18},
    {'Año': 2011, 'Nombre': 'Universidad Politécnica de Madrid', 'Escuelas': 18},
    {'Año': 2011, 'Nombre': 'Universidad Politécnica de Cataluña', 'Escuelas': 15},
    {'Año': 2011, 'Nombre': 'Universidad de Oviedo', 'Escuelas': 12},
    {'Año': 2011, 'Nombre': 'Universidad Politécnica de Valencia', 'Escuelas': 12},
    {'Año': 2011, 'Nombre': 'Universidad de La Coruña', 'Escuelas': 12},
    {'Año': 2011, 'Nombre': 'Universidad de Valladolid', 'Escuelas': 10},
    {'Año': 2011, 'Nombre': 'Universidad de Vigo', 'Escuelas': 9},
    {'Año': 2011, 'Nombre': 'Universidad de La Laguna', 'Escuelas': 8},
    {'Año': 2011, 'Nombre': 'Universidad de Granada', 'Escuelas': 8},
    {'Año': 2011, 'Nombre': 'Universidad de Alcalá', 'Escuelas': 8},
    {'Año': 2011, 'Nombre': 'Universidad de Sevilla', 'Escuelas': 8},
    {'Año': 2011, 'Nombre': 'Universidad de Málaga', 'Escuelas': 8},
    {'Año': 2011, 'Nombre': 'Universidad de Santiago de Compostela', 'Escuelas': 6},
    {'Año': 2011, 'Nombre': 'Universidad de Salamanca', 'Escuelas': 6},
    {'Año': 2011, 'Nombre': 'Universidad de Las Palmas de Gran Canaria', 'Escuelas': 6},
    {'Año': 2011, 'Nombre': 'Universidad Politécnica de Cartagena', 'Escuelas': 6},
    {'Año': 2011, 'Nombre': 'Universidad de Cádiz', 'Escuelas': 6},
    {'Año': 2011, 'Nombre': 'IE Universidad', 'Escuelas': 6},
    {'Año': 2011, 'Nombre': 'Universidad de Zaragoza', 'Escuelas': 6},
```

{'Año': 2011, 'Nombre': 'Universidad Rovira i Virgili', 'Escuelas': 5},
{'Año': 2011, 'Nombre': 'Universidad de Cantabria', 'Escuelas': 5},
{'Año': 2011, 'Nombre': 'Universidad Complutense de Madrid', 'Escuelas': 5},
{'Año': 2011, 'Nombre': 'Universidad de Burgos', 'Escuelas': 5},
{'Año': 2011, 'Nombre': 'Universidad de Córdoba', 'Escuelas': 4},
{'Año': 2011, 'Nombre': 'Universidad de Extremadura', 'Escuelas': 4},
{'Año': 2011, 'Nombre': 'Universidad de Jaén', 'Escuelas': 4},
{'Año': 2011, 'Nombre': 'Universidad Autónoma de Madrid', 'Escuelas': 4},
{'Año': 2011, 'Nombre': 'Universidad de León', 'Escuelas': 4},
{'Año': 2011, 'Nombre': 'Universidad de Valencia', 'Escuelas': 4},
{'Año': 2011, 'Nombre': 'Universidad Pompeu Fabra', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad de Lérida', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad Pública de Navarra', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad Ramon Llull', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad de Huelva', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad de Murcia', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad Rey Juan Carlos', 'Escuelas': 3},
{'Año': 2011, 'Nombre': 'Universidad de Gerona', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad de Navarra', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad Europea de Madrid', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad de Vic', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad Nacional de Educación a Distancia', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad Pontificia de Salamanca', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad de Almería', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad San Jorge', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad de Alicante', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad de las Islas Baleares', 'Escuelas': 2},
{'Año': 2011, 'Nombre': 'Universidad Internacional de Cataluña', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Carlos III de Madrid', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad de Mondragón', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Europea Miguel de Cervantes', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Autónoma de Barcelona', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Jaume I', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Alfonso X El Sabio', 'Escuelas': 1},

{'Año': 2011, 'Nombre': 'Universidad Antonio de Nebrija', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Católica San Antonio de Murcia', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad de La Rioja', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Miguel Hernández de Elche', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Pablo de Olavide', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad Pontificia de Comillas', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad San Pablo CEU', 'Escuelas': 1},
{'Año': 2011, 'Nombre': 'Universidad de Barcelona', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad a Distancia de Madrid', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Abat Oliba CEU', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Camilo José Cela', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Cardenal Herrera CEU', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Católica de Ávila', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Católica de Valencia', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad de Deusto', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Francisco de Vitoria', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Internacional de Andalucía', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Internacional de La Rioja', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Internacional de Valencia', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Internacional Menéndez Pelayo', 'Escuelas': 0},
{'Año': 2011, 'Nombre': 'Universidad Oberta de Catalunya', 'Escuelas': 0},
{'Año': 2010, 'Nombre': 'Universidad de Castilla-La Mancha', 'Escuelas': 9},
{'Año': 2010, 'Nombre': 'Universidad del País Vasco', 'Escuelas': 19},
{'Año': 2010, 'Nombre': 'Universidad Politécnica de Madrid', 'Escuelas': 3},
{'Año': 2010, 'Nombre': 'Universidad Politécnica de Cataluña', 'Escuelas': 13},
{'Año': 2010, 'Nombre': 'Universidad de Oviedo', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad Politécnica de Valencia', 'Escuelas': 7},
{'Año': 2010, 'Nombre': 'Universidad de La Coruña', 'Escuelas': 8},
{'Año': 2010, 'Nombre': 'Universidad de Valladolid', 'Escuelas': 18},
{'Año': 2010, 'Nombre': 'Universidad de Vigo', 'Escuelas': 10},
{'Año': 2010, 'Nombre': 'Universidad de La Laguna', 'Escuelas': 8},
{'Año': 2010, 'Nombre': 'Universidad de Granada', 'Escuelas': 16},
{'Año': 2010, 'Nombre': 'Universidad de Alcalá', 'Escuelas': 10},
{'Año': 2010, 'Nombre': 'Universidad de Sevilla', 'Escuelas': 19},



{'Año': 2010, 'Nombre': 'Universidad de Málaga', 'Escuelas': 5},
{'Año': 2010, 'Nombre': 'Universidad de Santiago de Compostela', 'Escuelas': 19},
{'Año': 2010, 'Nombre': 'Universidad de Salamanca', 'Escuelas': 0},
{'Año': 2010, 'Nombre': 'Universidad de Las Palmas de Gran Canaria', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad Politécnica de Cartagena', 'Escuelas': 7},
{'Año': 2010, 'Nombre': 'Universidad de Cádiz', 'Escuelas': 0},
{'Año': 2010, 'Nombre': 'IE Universidad', 'Escuelas': 5},
{'Año': 2010, 'Nombre': 'Universidad de Zaragoza', 'Escuelas': 16},
{'Año': 2010, 'Nombre': 'Universidad Rovira i Virgili', 'Escuelas': 10},
{'Año': 2010, 'Nombre': 'Universidad de Cantabria', 'Escuelas': 15},
{'Año': 2010, 'Nombre': 'Universidad Complutense de Madrid', 'Escuelas': 20},
{'Año': 2010, 'Nombre': 'Universidad de Burgos', 'Escuelas': 5},
{'Año': 2010, 'Nombre': 'Universidad de Córdoba', 'Escuelas': 1},
{'Año': 2010, 'Nombre': 'Universidad de Extremadura', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad de Jaén', 'Escuelas': 20},
{'Año': 2010, 'Nombre': 'Universidad Autónoma de Madrid', 'Escuelas': 4},
{'Año': 2010, 'Nombre': 'Universidad de León', 'Escuelas': 6},
{'Año': 2010, 'Nombre': 'Universidad de Valencia', 'Escuelas': 5},
{'Año': 2010, 'Nombre': 'Universidad Pompeu Fabra', 'Escuelas': 1},
{'Año': 2010, 'Nombre': 'Universidad de Lérida', 'Escuelas': 2},
{'Año': 2010, 'Nombre': 'Universidad Pública de Navarra', 'Escuelas': 4},
{'Año': 2010, 'Nombre': 'Universidad Ramon Llull', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad de Huelva', 'Escuelas': 17},
{'Año': 2010, 'Nombre': 'Universidad de Murcia', 'Escuelas': 14},
{'Año': 2010, 'Nombre': 'Universidad Rey Juan Carlos', 'Escuelas': 15},
{'Año': 2010, 'Nombre': 'Universidad de Gerona', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad de Navarra', 'Escuelas': 9},
{'Año': 2010, 'Nombre': 'Universidad Europea de Madrid', 'Escuelas': 9},
{'Año': 2010, 'Nombre': 'Universidad de Vic', 'Escuelas': 20},
{'Año': 2010, 'Nombre': 'Universidad Nacional de Educación a Distancia', 'Escuelas': 19},
{'Año': 2010, 'Nombre': 'Universidad Pontificia de Salamanca', 'Escuelas': 10},
{'Año': 2010, 'Nombre': 'Universidad de Almería', 'Escuelas': 16},
{'Año': 2010, 'Nombre': 'Universidad San Jorge', 'Escuelas': 17},
{'Año': 2010, 'Nombre': 'Universidad de Alicante', 'Escuelas': 18},

{'Año': 2010, 'Nombre': 'Universidad de las Islas Baleares', 'Escuelas': 2},
{'Año': 2010, 'Nombre': 'Universidad Internacional de Cataluña', 'Escuelas': 16},
{'Año': 2010, 'Nombre': 'Universidad Carlos III de Madrid', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad de Mondragón', 'Escuelas': 4},
{'Año': 2010, 'Nombre': 'Universidad Europea Miguel de Cervantes', 'Escuelas': 18},
{'Año': 2010, 'Nombre': 'Universidad Autónoma de Barcelona', 'Escuelas': 15},
{'Año': 2010, 'Nombre': 'Universidad Jaume I', 'Escuelas': 7},
{'Año': 2010, 'Nombre': 'Universidad Alfonso X El Sabio', 'Escuelas': 12},
{'Año': 2010, 'Nombre': 'Universidad Antonio de Nebrija', 'Escuelas': 16},
{'Año': 2010, 'Nombre': 'Universidad Católica San Antonio de Murcia', 'Escuelas': 6},
{'Año': 2010, 'Nombre': 'Universidad de La Rioja', 'Escuelas': 13},
{'Año': 2010, 'Nombre': 'Universidad Miguel Hernández de Elche', 'Escuelas': 17},
{'Año': 2010, 'Nombre': 'Universidad Pablo de Olavide', 'Escuelas': 10},
{'Año': 2010, 'Nombre': 'Universidad Pontificia de Comillas', 'Escuelas': 15},
{'Año': 2010, 'Nombre': 'Universidad San Pablo CEU', 'Escuelas': 3},
{'Año': 2010, 'Nombre': 'Universidad de Barcelona', 'Escuelas': 17},
{'Año': 2010, 'Nombre': 'Universidad a Distancia de Madrid', 'Escuelas': 4},
{'Año': 2010, 'Nombre': 'Universidad Abat Oliba CEU', 'Escuelas': 2},
{'Año': 2010, 'Nombre': 'Universidad Camilo José Cela', 'Escuelas': 8},
{'Año': 2010, 'Nombre': 'Universidad Cardenal Herrera CEU', 'Escuelas': 8},
{'Año': 2010, 'Nombre': 'Universidad Católica de Ávila', 'Escuelas': 5},
{'Año': 2010, 'Nombre': 'Universidad Católica de Valencia', 'Escuelas': 17},
{'Año': 2010, 'Nombre': 'Universidad de Deusto', 'Escuelas': 16},
{'Año': 2010, 'Nombre': 'Universidad Francisco de Vitoria', 'Escuelas': 6},
{'Año': 2010, 'Nombre': 'Universidad Internacional de Andalucía', 'Escuelas': 11},
{'Año': 2010, 'Nombre': 'Universidad Internacional de La Rioja', 'Escuelas': 13},
{'Año': 2010, 'Nombre': 'Universidad Internacional de Valencia', 'Escuelas': 3},
{'Año': 2010, 'Nombre': 'Universidad Internacional Menéndez Pelayo', 'Escuelas': 4},
{'Año': 2010, 'Nombre': 'Universidad Oberta de Catalunya', 'Escuelas': 8},
{'Año': 2009, 'Nombre': 'Universidad de Castilla-La Mancha', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad del País Vasco', 'Escuelas': 7},
{'Año': 2009, 'Nombre': 'Universidad Politécnica de Madrid', 'Escuelas': 5},
{'Año': 2009, 'Nombre': 'Universidad Politécnica de Cataluña', 'Escuelas': 20},
{'Año': 2009, 'Nombre': 'Universidad de Oviedo', 'Escuelas': 11},

{'Año': 2009, 'Nombre': 'Universidad Politécnica de Valencia', 'Escuelas': 12},
{'Año': 2009, 'Nombre': 'Universidad de La Coruña', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad de Valladolid', 'Escuelas': 18},
{'Año': 2009, 'Nombre': 'Universidad de Vigo', 'Escuelas': 14},
{'Año': 2009, 'Nombre': 'Universidad de La Laguna', 'Escuelas': 13},
{'Año': 2009, 'Nombre': 'Universidad de Granada', 'Escuelas': 15},
{'Año': 2009, 'Nombre': 'Universidad de Alcalá', 'Escuelas': 5},
{'Año': 2009, 'Nombre': 'Universidad de Sevilla', 'Escuelas': 14},
{'Año': 2009, 'Nombre': 'Universidad de Málaga', 'Escuelas': 19},
{'Año': 2009, 'Nombre': 'Universidad de Santiago de Compostela', 'Escuelas': 4},
{'Año': 2009, 'Nombre': 'Universidad de Salamanca', 'Escuelas': 15},
{'Año': 2009, 'Nombre': 'Universidad de Las Palmas de Gran Canaria', 'Escuelas': 8},
{'Año': 2009, 'Nombre': 'Universidad Politécnica de Cartagena', 'Escuelas': 11},
{'Año': 2009, 'Nombre': 'Universidad de Cádiz', 'Escuelas': 15},
{'Año': 2009, 'Nombre': 'IE Universidad', 'Escuelas': 7},
{'Año': 2009, 'Nombre': 'Universidad de Zaragoza', 'Escuelas': 3},
{'Año': 2009, 'Nombre': 'Universidad Rovira i Virgili', 'Escuelas': 13},
{'Año': 2009, 'Nombre': 'Universidad de Cantabria', 'Escuelas': 16},
{'Año': 2009, 'Nombre': 'Universidad Complutense de Madrid', 'Escuelas': 1},
{'Año': 2009, 'Nombre': 'Universidad de Burgos', 'Escuelas': 17},
{'Año': 2009, 'Nombre': 'Universidad de Córdoba', 'Escuelas': 4},
{'Año': 2009, 'Nombre': 'Universidad de Extremadura', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad de Jaén', 'Escuelas': 3},
{'Año': 2009, 'Nombre': 'Universidad Autónoma de Madrid', 'Escuelas': 7},
{'Año': 2009, 'Nombre': 'Universidad de León', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad de Valencia', 'Escuelas': 16},
{'Año': 2009, 'Nombre': 'Universidad Pompeu Fabra', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad de Lérida', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad Pública de Navarra', 'Escuelas': 10},
{'Año': 2009, 'Nombre': 'Universidad Ramon Llull', 'Escuelas': 13},
{'Año': 2009, 'Nombre': 'Universidad de Huelva', 'Escuelas': 20},
{'Año': 2009, 'Nombre': 'Universidad de Murcia', 'Escuelas': 10},
{'Año': 2009, 'Nombre': 'Universidad Rey Juan Carlos', 'Escuelas': 18},
{'Año': 2009, 'Nombre': 'Universidad de Gerona', 'Escuelas': 9},

{'Año': 2009, 'Nombre': 'Universidad de Navarra', 'Escuelas': 3},
{'Año': 2009, 'Nombre': 'Universidad Europea de Madrid', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad de Vic', 'Escuelas': 2},
{'Año': 2009, 'Nombre': 'Universidad Nacional de Educación a Distancia', 'Escuelas': 3},
{'Año': 2009, 'Nombre': 'Universidad Pontificia de Salamanca', 'Escuelas': 16},
{'Año': 2009, 'Nombre': 'Universidad de Almería', 'Escuelas': 3},
{'Año': 2009, 'Nombre': 'Universidad San Jorge', 'Escuelas': 16},
{'Año': 2009, 'Nombre': 'Universidad de Alicante', 'Escuelas': 7},
{'Año': 2009, 'Nombre': 'Universidad de las Islas Baleares', 'Escuelas': 20},
{'Año': 2009, 'Nombre': 'Universidad Internacional de Cataluña', 'Escuelas': 4},
{'Año': 2009, 'Nombre': 'Universidad Carlos III de Madrid', 'Escuelas': 12},
{'Año': 2009, 'Nombre': 'Universidad de Mondragón', 'Escuelas': 9},
{'Año': 2009, 'Nombre': 'Universidad Europea Miguel de Cervantes', 'Escuelas': 13},
{'Año': 2009, 'Nombre': 'Universidad Autónoma de Barcelona', 'Escuelas': 17},
{'Año': 2009, 'Nombre': 'Universidad Jaume I', 'Escuelas': 13},
{'Año': 2009, 'Nombre': 'Universidad Alfonso X El Sabio', 'Escuelas': 14},
{'Año': 2009, 'Nombre': 'Universidad Antonio de Nebrija', 'Escuelas': 11},
{'Año': 2009, 'Nombre': 'Universidad Católica San Antonio de Murcia', 'Escuelas': 15},
{'Año': 2009, 'Nombre': 'Universidad de La Rioja', 'Escuelas': 8},
{'Año': 2009, 'Nombre': 'Universidad Miguel Hernández de Elche', 'Escuelas': 0},
{'Año': 2009, 'Nombre': 'Universidad Pablo de Olavide', 'Escuelas': 16},
{'Año': 2009, 'Nombre': 'Universidad Pontificia de Comillas', 'Escuelas': 0},
{'Año': 2009, 'Nombre': 'Universidad San Pablo CEU', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad de Barcelona', 'Escuelas': 15},
{'Año': 2009, 'Nombre': 'Universidad a Distancia de Madrid', 'Escuelas': 7},
{'Año': 2009, 'Nombre': 'Universidad Abat Oliba CEU', 'Escuelas': 4},
{'Año': 2009, 'Nombre': 'Universidad Camilo José Cela', 'Escuelas': 16},
{'Año': 2009, 'Nombre': 'Universidad Cardenal Herrera CEU', 'Escuelas': 12},
{'Año': 2009, 'Nombre': 'Universidad Católica de Ávila', 'Escuelas': 3},
{'Año': 2009, 'Nombre': 'Universidad Católica de Valencia', 'Escuelas': 7},
{'Año': 2009, 'Nombre': 'Universidad de Deusto', 'Escuelas': 8},
{'Año': 2009, 'Nombre': 'Universidad Francisco de Vitoria', 'Escuelas': 6},
{'Año': 2009, 'Nombre': 'Universidad Internacional de Andalucía', 'Escuelas': 0},
{'Año': 2009, 'Nombre': 'Universidad Internacional de La Rioja', 'Escuelas': 20},

```
{'Año': 2009, 'Nombre': 'Universidad Internacional de Valencia', 'Escuelas': 17},
{'Año': 2009, 'Nombre': 'Universidad Internacional Menéndez Pelayo', 'Escuelas': 15},
{'Año': 2009, 'Nombre': 'Universidad Oberta de Catalunya', 'Escuelas': 17}
]
var visualization = d3plus.viz()
  .container("#viz")
  .data(Datos_Escuelas)
  .type("tree_map")
  .id("Nombre")
  .size("Escuelas")
  .tooltip(["Nombre", "Escuelas", "Año"])
  .time({"value": "Año", "solo": 2011})
  .draw()
</script>
```

Anexo II: Código del gráfico “Network” de la red social

En este ejemplo se ha desarrollado una red social.

La red está compuesta por los enlaces entre las páginas web de nueve de las universidades españolas más importantes.

Las aristas son la cantidad de enlaces y los nodos las páginas web de las universidades, cuya dimensión es el tamaño de sus página web.

Fichero index.html

```
<!doctype html>
<meta charset="utf-8">

<!-- load D3js -->
<script src="http://d3plus.org/js/d3.js"></script>

<!-- load D3plus after D3js -->
<script src="http://d3plus.org/js/d3plus.js"></script>

<!-- create container element for visualization -->
<div id="viz"></div>

<script>
  // create sample dataset
  var Datos_universidades = [
    {"Nombre": "ub.edu", "Paginas": 1900000},
    {"Nombre": "ucm.es", "Paginas": 1870000},
    {"Nombre": "uab.cat", "Paginas": 2030000},
    {"Nombre": "uv.es", "Paginas": 1560000},
    {"Nombre": "ugr.es", "Paginas": 1670000},
    {"Nombre": "upv.es", "Paginas": 1470000},
    {"Nombre": "uam.es", "Paginas": 690000},
    {"Nombre": "us.es", "Paginas": 3050000},
    {"Nombre": "upm.es", "Paginas": 1170000}
  ]
```

```

// create list of node connections
var connections = [
    {source: 'ucm.es', 'target': 'ub.edu', 'strength': 647},
    {source: 'uab.cat', 'target': 'ub.edu', 'strength': 1064},
    {source: 'uv.es', 'target': 'ub.edu', 'strength': 386},
    {source: 'ugr.es', 'target': 'ub.edu', 'strength': 110},
    {source: 'upv.es', 'target': 'ub.edu', 'strength': 304},
    {source: 'uam.es', 'target': 'ub.edu', 'strength': 115},
    {source: 'us.es', 'target': 'ub.edu', 'strength': 68},
    {source: 'upm.es', 'target': 'ub.edu', 'strength': 9},
    {source: 'ub.edu', 'target': 'ucm.es', 'strength': 726},
    {source: 'uab.cat', 'target': 'ucm.es', 'strength': 3384},
    {source: 'uv.es', 'target': 'ucm.es', 'strength': 2231},
    {source: 'ugr.es', 'target': 'ucm.es', 'strength': 300},
    {source: 'upv.es', 'target': 'ucm.es', 'strength': 290},
    {source: 'uam.es', 'target': 'ucm.es', 'strength': 846},
    {source: 'us.es', 'target': 'ucm.es', 'strength': 155},
    {source: 'upm.es', 'target': 'ucm.es', 'strength': 399},
    {source: 'ucm.es', 'target': 'uab.cat', 'strength': 15},
    {source: 'ub.edu', 'target': 'uab.cat', 'strength': 439},
    {source: 'uv.es', 'target': 'uab.cat', 'strength': 54},
    {source: 'ugr.es', 'target': 'uab.cat', 'strength': 44},
    {source: 'upv.es', 'target': 'uab.cat', 'strength': 10},
    {source: 'uam.es', 'target': 'uab.cat', 'strength': 74},
    {source: 'us.es', 'target': 'uab.cat', 'strength': 47},
    {source: 'upm.es', 'target': 'uab.cat', 'strength': 11},
    {source: 'ucm.es', 'target': 'uv.es', 'strength': 91},
    {source: 'uab.cat', 'target': 'uv.es', 'strength': 197},
    {source: 'ub.edu', 'target': 'uv.es', 'strength': 302},
    {source: 'ugr.es', 'target': 'uv.es', 'strength': 226},
    {source: 'upv.es', 'target': 'uv.es', 'strength': 940},
    {source: 'uam.es', 'target': 'uv.es', 'strength': 233},
    {source: 'us.es', 'target': 'uv.es', 'strength': 75},
    {source: 'upm.es', 'target': 'uv.es', 'strength': 58},

```



{source: 'uam.es', 'target': 'uam.es', 'strength': 1000},
{source: 'uab.cat', 'target': 'uab.cat', 'strength': 1000},
{source: 'uv.es', 'target': 'uv.es', 'strength': 1000},
{source: 'ugr.es', 'target': 'ugr.es', 'strength': 1000},
{source: 'upv.es', 'target': 'upv.es', 'strength': 1000},
{source: 'ub.edu', 'target': 'ub.edu', 'strength': 1000},
{source: 'us.es', 'target': 'us.es', 'strength': 1000},
{source: 'upm.es', 'target': 'upm.es', 'strength': 1000},
{source: 'uam.es', 'target': 'uab.cat', 'strength': 176},
{source: 'uab.cat', 'target': 'uam.es', 'strength': 141},
{source: 'uv.es', 'target': 'uam.es', 'strength': 150},
{source: 'ugr.es', 'target': 'uam.es', 'strength': 219},
{source: 'upv.es', 'target': 'uam.es', 'strength': 57},
{source: 'ub.edu', 'target': 'uam.es', 'strength': 131},
{source: 'us.es', 'target': 'uam.es', 'strength': 75},
{source: 'upm.es', 'target': 'uam.es', 'strength': 120},
{source: 'uam.es', 'target': 'upv.es', 'strength': 44},
{source: 'upv.es', 'target': 'uam.es', 'strength': 57},
{source: 'ub.edu', 'target': 'upv.es', 'strength': 54},
{source: 'us.es', 'target': 'upv.es', 'strength': 316},
{source: 'upm.es', 'target': 'upv.es', 'strength': 284},
{source: 'uam.es', 'target': 'upv.es', 'strength': 176},
{source: 'uab.cat', 'target': 'upv.es', 'strength': 61},
{source: 'uv.es', 'target': 'upv.es', 'strength': 1006},
{source: 'ugr.es', 'target': 'upv.es', 'strength': 108},
{source: 'uam.es', 'target': 'upm.es', 'strength': 176},
{source: 'uab.cat', 'target': 'upm.es', 'strength': 32},
{source: 'uv.es', 'target': 'upm.es', 'strength': 145},
{source: 'ugr.es', 'target': 'upm.es', 'strength': 182},
{source: 'upv.es', 'target': 'upm.es', 'strength': 149},
{source: 'ub.edu', 'target': 'upm.es', 'strength': 109},
{source: 'us.es', 'target': 'upm.es', 'strength': 145},
{source: 'uam.es', 'target': 'us.es', 'strength': 126},
{source: 'ub.edu', 'target': 'us.es', 'strength': 109},
{source: 'upv.es', 'target': 'us.es', 'strength': 139},
{source: 'ugr.es', 'target': 'us.es', 'strength': 221},
{source: 'uv.es', 'target': 'us.es', 'strength': 407},
{source: 'uab.cat', 'target': 'us.es', 'strength': 163},
{source: 'uam.es', 'target': 'us.es', 'strength': 126},
{source: 'ub.edu', 'target': 'us.es', 'strength': 109},
{source: 'upm.es', 'target': 'us.es', 'strength': 145},
{source: 'uam.es', 'target': 'uam.es', 'strength': 1000},
{source: 'uab.cat', 'target': 'uab.cat', 'strength': 1000},
{source: 'uv.es', 'target': 'uv.es', 'strength': 1000},
{source: 'ugr.es', 'target': 'ugr.es', 'strength': 1000},
{source: 'upv.es', 'target': 'upv.es', 'strength': 1000},
{source: 'ub.edu', 'target': 'ub.edu', 'strength': 1000},
{source: 'us.es', 'target': 'us.es', 'strength': 1000},
{source: 'upm.es', 'target': 'upm.es', 'strength': 1000}

```

    {source: 'uv.es', 'target': 'upm.es', 'strength': 75},
    {source: 'ugr.es', 'target': 'upm.es', 'strength': 154},
    {source: 'upv.es', 'target': 'upm.es', 'strength': 552},
    {source: 'uam.es', 'target': 'upm.es', 'strength': 274},
    {source: 'us.es', 'target': 'upm.es', 'strength': 108},
    {source: 'ub.edu', 'target': 'upm.es', 'strength': 94}
  ]
  // instantiate d3plus
  var visualization = d3plus.viz()
    .container("#viz")
    .type("network")
    .data(Datos_universidades)
    .edges(connections)
    .size("Paginas")
    .id("Nombre")
    .tooltip(["Nombre","Paginas"])
    .edges({
      "size": "strength",
      // "label": "strength",
      "value": connections,
      "arrows": true
    })

    .draw()
</script>

```



Anexo III: Código del gráfico “Network” de la red social multinivel

En este ejemplo se ha desarrollado una red social multinivel. El primer nivel es compuesto de los enlaces entre las páginas web de nueve de las universidades españolas más importantes, (red social desarrollada en el Anexo II) y es segundo nivel es compuesto por los enlaces entre sus departamentos de matemáticas.

Fichero index.html

```
<!doctype html>
<meta charset="utf-8">

<!-- load D3js -->
<script src="http://d3plus.org/js/d3.js"></script>

<!-- load D3plus after D3js -->
<script src="http://d3plus.org/js/d3plus.js"></script>

<!-- create container element for visualization -->
<div id="viz"></div>

<script>
  // create sample dataset
  var Datos_Universidades = [
    {"Nombre": "ub.edu", "Paginas": 1900000, "Tipo":"Universidad"},
    {"Nombre": "ucm.es", "Paginas": 1870000, "Tipo":"Universidad"},
    {"Nombre": "uab.cat", "Paginas": 2030000, "Tipo":"Universidad"},
    {"Nombre": "uv.es", "Paginas": 1560000, "Tipo":"Universidad"},
    {"Nombre": "ugr.es", "Paginas": 1670000, "Tipo":"Universidad"},
    {"Nombre": "upv.es", "Paginas": 1470000, "Tipo":"Universidad"},
    {"Nombre": "uam.es", "Paginas": 690000, "Tipo":"Universidad"},
    {"Nombre": "us.es", "Paginas": 3050000, "Tipo":"Universidad"},
    {"Nombre": "upm.es", "Paginas": 1170000, "Tipo":"Universidad"},
    {"Nombre": "mat.ub.edu/", "Paginas": 4590, "Tipo":"Matematicas"},
    {"Nombre": "matematicas.ucm.es/", "Paginas": 4650, "Tipo":"Matematicas"},
    {"Nombre": "uv.es/uvweb/matematicas/", "Paginas": 853, "Tipo":"Matematicas"},
    {"Nombre": "analisismatematico.ugr.es/", "Paginas": 374, "Tipo":"Matematicas"},
```

```

{"Nombre": "upv.es/entidades/DMAA/", "Paginas": 441, "Tipo":"Matematicas"},
{"Nombre": "verso.mat.uam.es/", "Paginas": 2550, "Tipo":"Matematicas"},
{"Nombre": "matematicas.us.es/", "Paginas": 2950, "Tipo":"Matematicas"}
]

```

```

var positions = [
{"Nombre": "ub.edu", "x": 10, "y": 15},
{"Nombre": "ucm.es", "x": 12, "y": 24},
{"Nombre": "uab.cat", "x": 16, "y": 18},
{"Nombre": "uv.es", "x": 26, "y": 21},
{"Nombre": "ugr.es", "x": 13, "y": 4},
{"Nombre": "upv.es", "x": 31, "y": 13},
{"Nombre": "uam.es", "x": 19, "y": 8},
{"Nombre": "us.es", "x": 24, "y": 11},
{"Nombre": "upm.es", "x": 20, "y": 5},
{"Nombre": "mat.ub.edu/", "x": 10, "y": 16},
{"Nombre": "matematicas.ucm.es/", "x": 12, "y": 25},
{"Nombre": "uv.es/uvweb/matematicas/", "x": 26, "y": 22},
{"Nombre": "analisismatematico.ugr.es/", "x": 13, "y": 5},
{"Nombre": "upv.es/entidades/DMAA/", "x": 31, "y": 14},
{"Nombre": "verso.mat.uam.es/", "x": 19, "y": 9}
]

```

```

// create list of node connections

```

```

var connections = [
    {source: 'ucm.es', target: 'ub.edu', strength: 647},
    {source: 'uab.cat', target: 'ub.edu', strength: 1064},
    {source: 'uv.es', target: 'ub.edu', strength: 386},
    {source: 'ugr.es', target: 'ub.edu', strength: 110},
    {source: 'upv.es', target: 'ub.edu', strength: 304},
    {source: 'uam.es', target: 'ub.edu', strength: 115},
    {source: 'us.es', target: 'ub.edu', strength: 68},
    {source: 'upm.es', target: 'ub.edu', strength: 9},

```



{source: 'ub.edu', 'target': 'ucm.es', 'strength': 726},
{source: 'uab.cat', 'target': 'ucm.es', 'strength': 3384},
{source: 'uv.es', 'target': 'ucm.es', 'strength': 2231},
{source: 'ugr.es', 'target': 'ucm.es', 'strength': 300},
{source: 'upv.es', 'target': 'ucm.es', 'strength': 290},
{source: 'uam.es', 'target': 'ucm.es', 'strength': 846},
{source: 'us.es', 'target': 'ucm.es', 'strength': 155},
{source: 'upm.es', 'target': 'ucm.es', 'strength': 399},
{source: 'ucm.es', 'target': 'uab.cat', 'strength': 15},
{source: 'ub.edu', 'target': 'uab.cat', 'strength': 439},
{source: 'uv.es', 'target': 'uab.cat', 'strength': 54},
{source: 'ugr.es', 'target': 'uab.cat', 'strength': 44},
{source: 'upv.es', 'target': 'uab.cat', 'strength': 10},
{source: 'uam.es', 'target': 'uab.cat', 'strength': 74},
{source: 'us.es', 'target': 'uab.cat', 'strength': 47},
{source: 'upm.es', 'target': 'uab.cat', 'strength': 11},
{source: 'ucm.es', 'target': 'uv.es', 'strength': 91},
{source: 'uab.cat', 'target': 'uv.es', 'strength': 197},
{source: 'ub.edu', 'target': 'uv.es', 'strength': 302},
{source: 'ugr.es', 'target': 'uv.es', 'strength': 226},
{source: 'upv.es', 'target': 'uv.es', 'strength': 940},
{source: 'uam.es', 'target': 'uv.es', 'strength': 233},
{source: 'us.es', 'target': 'uv.es', 'strength': 75},
{source: 'upm.es', 'target': 'uv.es', 'strength': 58},
{source: 'ucm.es', 'target': 'ugr.es', 'strength': 223},
{source: 'uab.cat', 'target': 'ugr.es', 'strength': 277},
{source: 'uv.es', 'target': 'ugr.es', 'strength': 2215},
{source: 'ub.edu', 'target': 'ugr.es', 'strength': 206},
{source: 'upv.es', 'target': 'ugr.es', 'strength': 179},
{source: 'uam.es', 'target': 'ugr.es', 'strength': 176},
{source: 'us.es', 'target': 'ugr.es', 'strength': 843},
{source: 'upm.es', 'target': 'ugr.es', 'strength': 182},
{source: 'ucm.es', 'target': 'upv.es', 'strength': 180},
{source: 'uab.cat', 'target': 'upv.es', 'strength': 61},



{source: 'uv.es', 'target': 'upv.es', 'strength': 1006},
{source: 'ugr.es', 'target': 'upv.es', 'strength': 108},
{source: 'ub.edu', 'target': 'upv.es', 'strength': 54},
{source: 'uam.es', 'target': 'upv.es', 'strength': 44},
{source: 'us.es', 'target': 'upv.es', 'strength': 316},
{source: 'upm.es', 'target': 'upv.es', 'strength': 284},
{source: 'ucm.es', 'target': 'uam.es', 'strength': 192},
{source: 'uab.cat', 'target': 'uam.es', 'strength': 141},
{source: 'uv.es', 'target': 'uam.es', 'strength': 150},
{source: 'ugr.es', 'target': 'uam.es', 'strength': 219},
{source: 'upv.es', 'target': 'uam.es', 'strength': 57},
{source: 'ub.edu', 'target': 'uam.es', 'strength': 131},
{source: 'us.es', 'target': 'uam.es', 'strength': 75},
{source: 'upm.es', 'target': 'uam.es', 'strength': 120},
{source: 'ucm.es', 'target': 'us.es', 'strength': 78},
{source: 'uab.cat', 'target': 'us.es', 'strength': 163},
{source: 'uv.es', 'target': 'us.es', 'strength': 407},
{source: 'ugr.es', 'target': 'us.es', 'strength': 221},
{source: 'upv.es', 'target': 'us.es', 'strength': 139},
{source: 'uam.es', 'target': 'us.es', 'strength': 126},
{source: 'ub.edu', 'target': 'us.es', 'strength': 109},
{source: 'upm.es', 'target': 'us.es', 'strength': 145},
{source: 'ucm.es', 'target': 'upm.es', 'strength': 149},
{source: 'uab.cat', 'target': 'upm.es', 'strength': 32},
{source: 'uv.es', 'target': 'upm.es', 'strength': 75},
{source: 'ugr.es', 'target': 'upm.es', 'strength': 154},
{source: 'upv.es', 'target': 'upm.es', 'strength': 552},
{source: 'uam.es', 'target': 'upm.es', 'strength': 274},
{source: 'us.es', 'target': 'upm.es', 'strength': 108},
{source: 'ub.edu', 'target': 'upm.es', 'strength': 94},
{source: 'matematicas.ucm.es/', 'target': 'mat.ub.edu/', 'strength': 1445},
{source: 'uv.es/uvweb/matematicas/', 'target': 'mat.ub.edu/', 'strength': 655},
{source: 'analismatematico.ugr.es/', 'target': 'mat.ub.edu/', 'strength': 1253},
{source: 'upv.es/entidades/DMAA/', 'target': 'mat.ub.edu/', 'strength': 1908},

```

    {source: 'verso.mat.uam.es/', 'target': 'mat.ub.edu/', 'strength': 702},
    {source: 'mat.ub.edu/', 'target': 'matematicas.ucm.es/', 'strength': 1598},
    {source: 'uv.es/uvweb/matematicas/', 'target': 'matematicas.ucm.es/', 'strength': 796},
    {source: 'analisismatematico.ugr.es/', 'target': 'matematicas.ucm.es/', 'strength': 602},
    {source: 'upv.es/entidades/DMAA/', 'target': 'matematicas.ucm.es/', 'strength': 2098},
    {source: 'verso.mat.uam.es/', 'target': 'matematicas.ucm.es/', 'strength': 2655},
    {source: 'mat.ub.edu/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 2797},
    {source: 'matematicas.ucm.es/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 1960},
    {source: 'analisismatematico.ugr.es/', 'target': 'uv.es/uvweb/matematicas/', 'strength':
2026},
    {source: 'upv.es/entidades/DMAA/', 'target': 'uv.es/uvweb/matematicas/', 'strength':
2686},
    {source: 'verso.mat.uam.es/', 'target': 'uv.es/uvweb/matematicas/', 'strength': 919},
    {source: 'mat.ub.edu/', 'target': 'verso.mat.uam.es/', 'strength': 742},
    {source: 'uv.es/uvweb/matematicas/', 'target': 'verso.mat.uam.es/', 'strength': 1010},
    {source: 'analisismatematico.ugr.es/', 'target': 'verso.mat.uam.es/', 'strength': 2449},
    {source: 'upv.es/entidades/DMAA/', 'target': 'verso.mat.uam.es/', 'strength': 2151},
    {source: 'matematicas.ucm.es/', 'target': 'verso.mat.uam.es/', 'strength': 936},
    {source: 'mat.ub.edu/', 'target': 'upv.es/entidades/DMAA/', 'strength': 972},
    {source: 'uv.es/uvweb/matematicas/', 'target': 'upv.es/entidades/DMAA/', 'strength':
2493},
    {source: 'analisismatematico.ugr.es/', 'target': 'upv.es/entidades/DMAA/', 'strength': 677},
    {source: 'matematicas.ucm.es/', 'target': 'upv.es/entidades/DMAA/', 'strength': 1909},
    {source: 'verso.mat.uam.es/', 'target': 'upv.es/entidades/DMAA/', 'strength': 1280},
    {source: 'mat.ub.edu/', 'target': 'analisismatematico.ugr.es/', 'strength': 2004},
    {source: 'uv.es/uvweb/matematicas/', 'target': 'analisismatematico.ugr.es/', 'strength':
1529},
    {source: 'upv.es/entidades/DMAA/', 'target': 'analisismatematico.ugr.es/', 'strength': 2920},
    {source: 'matematicas.ucm.es/', 'target': 'analisismatematico.ugr.es/', 'strength': 2612},
    {source: 'verso.mat.uam.es/', 'target': 'analisismatematico.ugr.es/', 'strength': 2334}
]

// instantiate d3plus
var visualization = d3plus.viz()
    .container("#viz")
    .type("network")

```



```
.data(Datos_Universidades)
.edges(connections)
.size("Paginas")
.id("Nombre")
    .nodes(positions)
    .tooltip(["Nombre","Paginas"])
    .edges({
"size": "strength",
    "label": "strength",
"value": connections,
    "arrows": true
})
    .color(function(d){
return d.Tipo == "Universidad" ? "#008800" : "#880000";
})

.draw()
</script>
```

