



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica
Superior d'Enginyeria
Informàtica

Escola Tècnica Superior d'Enginyeria
Informàtica
Universitat Politècnica de València

**Diseño y desarrollo de aplicaciones web
(automatización, consulta a APIs) para el
análisis de datos webométricos**

Trabajo Fin de Grado

Grado en Ingeniería Informática

Autor: Svyatoslav Kuzmychov

Tutor: José Antonio Ontalba Ruipérez

Cotutor: Jorge Ignacio Serrano Cobos

2015 / 2016

Resumen

Muchas de las bodegas de vinos españolas necesitan mejorar su posicionamiento web debido a que no tienen en cuenta el potencial de la web y a la cantidad de personas que pueden llegar. Por ello, mediante ciertos indicadores cibernéticos se realizará una línea temporal en la que se podrán observar las tendencias generalizadas de las prácticas de esas bodegas en cuanto a sus esfuerzos por mejorar su presencia en Internet.

Para lograr este objetivo, en el siguiente trabajo se describirá el proceso de extracción de estos indicadores, el proceso de tratamiento de los mismos y finalmente el proceso de la visualización de los mismos. Una vez en mano todas las representaciones gráficas de los indicadores, se llegarían a las debidas conclusiones para describir la situación de las bodegas españolas a través de los indicadores analizados.

Palabras clave: data science, big data, APIs, herramientas de análisis, cibermetría, webometría



Abstract

Many of the Spanish wineries need to improve his website rankings because they do not take into account the potential of the web and the amount of people they can get. Therefore, using certain webometrics indicators will be performed a timeline in which you can observe widespread trends practices of those wineries in their efforts to improve their Internet presence.

To achieve this goal, in the next job will be described the extraction process of those indicators, the treatment process of the same and finally the visualization process. Once in hand all graphical representations of the webometrics indicators, it would reach the appropriate conclusions to describe the situation of the Spanish wineries through the indicators analysed.

Keywords: data science, big data, APIs, analysis tools, webometrics

Tabla de contenidos

1. Introducción	11
2. Metodología	14
2.1 Extracción de datos	14
2.1.1 Posicionamiento web o SEO	14
2.1.2 Motores de búsqueda	17
2.1.3 Fuentes de datos	18
2.1.4 Tecnologías utilizadas	23
2.2 Manejo de datos	31
2.2.1 ¿Qué lenguaje de programación elegir?.....	31
2.2.2 Entorno y librerías	39
3. Resultados	40
3.1 Tratamiento de datos.....	40
3.1.1 Puesta a punto	40
3.1.2 Obtención de los países que más visitan las bodegas según similarWeb	44
3.1.3 Obtención de los países que más visitan las bodegas según Alexa	48
3.1.4 Obtención de los sitios que hacen referencia y los sitios destinos de las bodegas según similarWeb.....	51
3.1.5 Obtención del porcentaje de uso de las redes sociales más utilizadas según similarWeb	54
3.1.6 Obtención de datos comparativos entre Alexa y similarWeb	58
3.1.7 Obtención de top cinco sitios a partir del trust flow, citation flow, backlinks externos y dominios de referencia según majesticSEO	63
3.2 Visualización de datos	65
3.2.1 Generación gráfica de los países que más visitan las bodegas según similarWeb	70
3.2.2 Generación gráfica de los países que más visitan las bodegas según Alexa	74
3.2.3 Generación gráfica de los sitios que hacen referencia y los sitios destinos según similarWeb	75



3.2.4	Generación gráfica del porcentaje de uso de las redes sociales más utilizadas según similarWeb	77
3.2.5	Generación de gráficas de datos comparativos entre Alexa y similarWeb	79
3.2.6	Generación de tablas con el top de los cinco sitios a partir del trust flow, citation flow, backlinks externos y dominios de referencia según majesticSEO	84
4.	Conclusiones	88
4.1	Conclusiones de los resultados	89
4.2	Conclusiones futuras	92
5.	Referencias.....	94
6.	Anexos	95
6.1	Anexo A.....	95



Índice de figuras

Figura 1. Ciencias de la Información. Fuente: revistalatinacs.org	12
Figura 2. Resultados de búsqueda. Fuente: epymeonline.com	15
Figura 3. Ejemplo JSON de Alexa.....	27
Figura 4. Ejemplo XML de Alexa	28
Figura 5. ¿Qué lenguaje de programación elegir? Fuente: burtchworks.com	32
Figura 6. Preferencias de los analistas a lo largo de los años. Fuente: burtchworks.com	33
Figura 7. Lenguajes de programación en la industria. Fuente: burtchworks.com	34
Figura 8. Lenguajes de programación en la educación. Fuente: burtchworks.com ...	35
Figura 9. Preferencia de lenguaje de programación según años de experiencia. Fuente: burtchworks.com	36
Figura 10. Lenguajes de programación entre los científicos y analistas. Fuente: burtchworks.com	37
Figura 11. Lenguajes de programación entre los científicos y analistas a lo largo de los años. Fuente: burtchworks.com	37
Figura 12. Ejemplos de gráficas que se pueden generar con la biblioteca matplotlib. Fuente: tonysyu.github.io	66
Figura 13. Más ejemplos de gráficas generadas con matplotlib. Fuente: tonysyu.github.io	67
Figura 14. Otro ejemplo de gráfico generado con la biblioteca matplotlib. Fuente: tonysyu.github.io	68
Figura 15. Ventana gráfica que aparece al generar un gráfico. Fuente: stackoverflow.com	69
Figura 16. Gráfica con los países que más visitan las bodegas según similarWeb ...	73
Figura 17. Gráfica con los países que más visitan las bodegas según Alexa	74
Figura 18. Gráfica con los sitios que hacen más referencia a las bodegas según similarWeb	75
Figura 19. Gráfica con los sitios destinos de las bodegas según similarWeb.....	76
Figura 20. Gráfica con el porcentaje de uso de las redes sociales según similarWeb	78
Figura 21. Gráfica que compara los campos Bounce Rate y Search visits de Alexa y similarWeb	81

Figura 22. Gráfica que compara el Time On Site de Alexa y similarWeb	82
Figura 23. Gráfica que compara el Page Views de Alexa y similarWeb.....	83
Tabla 1. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de marzo	85
Tabla 2. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de abril.....	86
Tabla 3. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de mayo	86
Tabla 4. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de junio.....	87
Tabla 5. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de julio.....	87
Tabla 6. Sitios web de las bodegas	95

1. Introducción

La tendencia actual y no tan actual de llevar el negocio a un nuevo nivel para dar a conocer a la empresa a un ámbito mayor de personas está evolucionando a pasos agigantados. Este nuevo nivel es el de la web donde todo está cada vez más interconectado y acceder a la información es tan sencillo como tener un ordenador con conexión a Internet.

Sin embargo, aunque muchas empresas tienen su propia web informativa para que el mundo las conozca, esto en muchos casos no basta e incluso se desaprovecha esta herramienta, teniendo la capacidad de ampliar horizontes.

Cuando se habla de que se desaprovecha la web, se quiere decir que ya no basta con tener una web con información de la empresa a modo de presentarla y mostrar las direcciones de contacto y los teléfonos, sino que hace falta tener más contacto con las personas, anunciarse de otros modos. Como el ejemplo más claro están las redes sociales donde a primera mano se puede obtener información de lo que opina la gente sobre la empresa o de algún determinado producto.

Pero para ello, hay que estar en esas redes sociales y muchas empresas por desgracia no lo están, con lo cual permanecen en una niebla donde no saben qué se dice de ellos si es que ocurre. A parte de potenciar el ser conocido por todos sirve muchas veces para dar la oportunidad de abrirse al mundo y darse a conocer desde cero, para lo que como ya se ha dicho es importante estar presente en las principales redes sociales como Facebook, Twitter, YouTube, Instagram, etc.

Gracias al auge de las redes sociales está creciendo la profesión de “community manager” ^[1] donde entre sus principales funciones están las de creador de contenido atractivo y de calidad no sólo en las redes sociales sino también en el blog corporativo de la empresa, además tiene que saber cuándo es el momento ideal para publicar cierto contenido, es decir, qué día de la semana y a qué hora debería publicarse alguna entrada o noticia.

Por otro lado, hay otras formas más técnicas de saber cómo va la empresa en el mundo digital, estas formas son mediante indicadores métricos, que al ser analizados permiten comprender, controlar y predecir hasta cierto punto un fenómeno en particular como puede ser la presentación de un nuevo producto o el cambio del logo de la empresa.

¹ ¿Qué es un community manager? <http://comunidad.iebschool.com/iebs/general/ques-un-community-manager/>



La cibermetría ^[2] se puede definir, según Ali Ashgar Shiri, como la medición, el estudio o el análisis de cualquier información en el mundo digital que donde se emplean técnicas bibliométricas, cienciométricas e infométricas.

Sin embargo, en este trabajo del que se hará uso y se profundizará es el de la webometría, que estaría integrado en el campo de las métricas de las Ciencias de la Información, tal como se muestra en la siguiente figura.

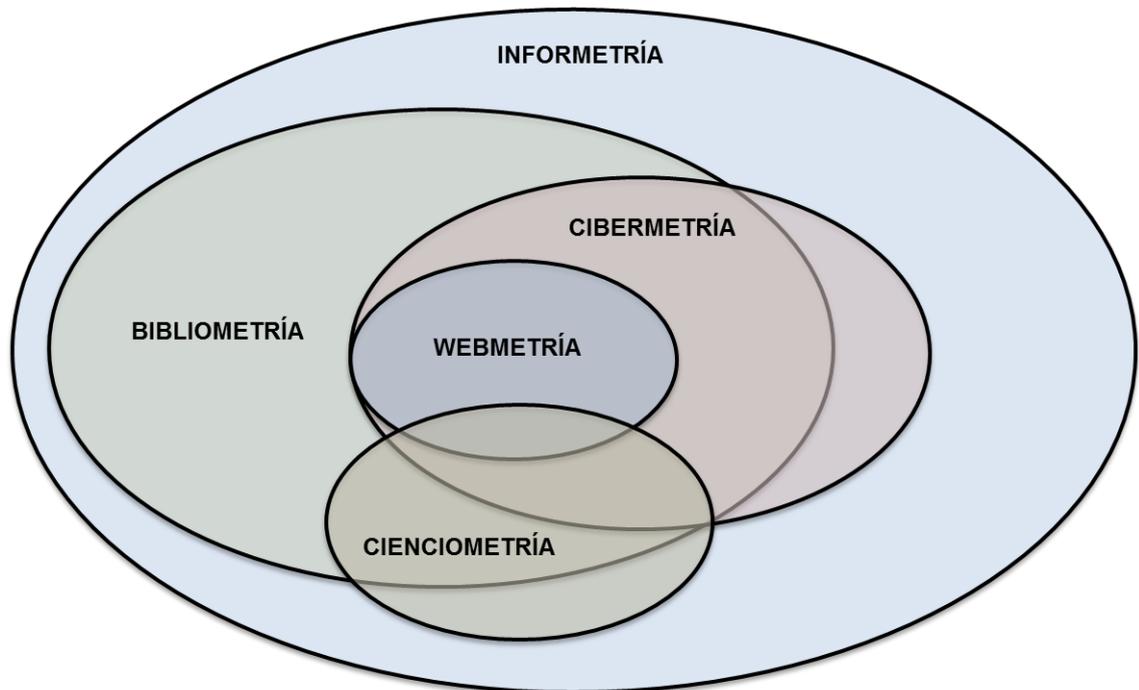


Figura 1. Ciencias de la Información. Fuente: *revistalatinacs.org*

Este campo de la cibermetría, la definió BjÖrneborn e Ingwersen (2004), como:

“El estudio de los aspectos cuantitativos de la construcción y uso de los recursos de información, estructuras y tecnologías de una parte concreta de Internet, por regla general a una web o portal, desde perspectivas bibliométricas e infométricas.”

Con ello, llevar a cabo estudios webométricos puede llegar a ser de gran utilidad para las empresas ya que es posible obtener datos como:

- Sitios y páginas web más enlazados.
- Países que más visitan determinados sitios web.
- Quién visita los sitios web, ya sea el género de la persona, si es un adolescente, adulto o personas mayores.
- Cantidad total de visitas

² Cibermetría. Artículo de Wikipedia. Consultado en agosto de 2016. <https://es.wikipedia.org/wiki/Cibermetr%C3%ADa>

Con todo esto, se pretende en este trabajo mostrar, describir y analizar el estado de la situación de las bodegas españolas en Internet mediante los indicadores cibernéticos analizados. A su vez, esta información puede servir como apoyo a las bodegas para compararse con la mayoría de las otras bodegas españolas y así decidir qué estrategias deberían seguir, si las mismas de la mayoría de esas bodegas, o prefieren utilizar técnicas distintas para mejorar en los diversos indicadores estudiados.

Para lograr este objetivo, se tendrán que obtener los indicadores métricos de cada sitio web a partir de otros sitios web que los proporcionen. Estos sitios web serán Alexa, similarWeb y majesticSEO. Más tarde, una vez con los datos, se pasaría a la fase del tratamiento de datos, donde se desecharía aquellos que fueran irrelevantes para el estudio y se harían los respectivos cálculos.

Finalmente, estos datos estarían listos para ser representados gráficamente ya sea por gráficas o tablas de distintos tipos para sacar y analizar las debidas conclusiones a las que se ha llegado durante todo el proceso.

Uno de los sectores empresariales donde la webometría se podría aplicar con éxito es la del sector del vino. Aunque hay marcas que están avanzadas en esta materia, son una ínfima parte de la cantidad de empresas que están atrasadas en este sector en cuanto a que sólo tienen el sitio web a modo de información, llegando incluso en algunos casos a no tener algo primordial en estos días que es tener presencia en las redes sociales más importantes. Por no decir que, la calidad de los sitios webs en muchos casos es anticuada, es decir, que se hizo hace varios años y no se modificó ni actualizó el diseño.

2. Metodología

En este capítulo se aborda todo el proceso de extracción de datos, donde se incluyen las fuentes desde donde se han obtenido los datos, las tecnologías utilizadas (tanto las definitivas como las que se probaron y luego no dieron el resultado esperado) para la extracción de datos y las dificultades y los problemas encontrados.

A su vez, se abordará el proceso previo del tratamiento de datos, donde se analizará qué lenguaje de programación se ha elegido, las opciones que había y finalmente por cual se ha optado.

En el mundo web actual, existen múltiples indicadores que ayudan a los sitios web a conocer el nivel de tráfico o visitas que se realizan en su sitio web. Si bien cada sitio web tiene sus propios métodos (si es que tienen) para obtener estadísticas, hay otros indicadores que ofrecen distintas webs, gestionadas por expertos especializados en el posicionamiento web, que pueden certificar la capacidad de éxito de un sitio web.

2.1 Extracción de datos

2.1.1 Posicionamiento web o SEO

El posicionamiento web ^[3] o *Search Engine Optimization* (SEO), es una labor de optimización para los motores de búsqueda más populares como Google o Bing.

³ Posicionamiento en buscadores. Artículo de Wikipedia. Consultado en agosto de 2016. https://es.wikipedia.org/wiki/Posicionamiento_en_buscadores

En la siguiente figura se puede observar los resultados que se ofrecen tanto de pago como orgánicos.

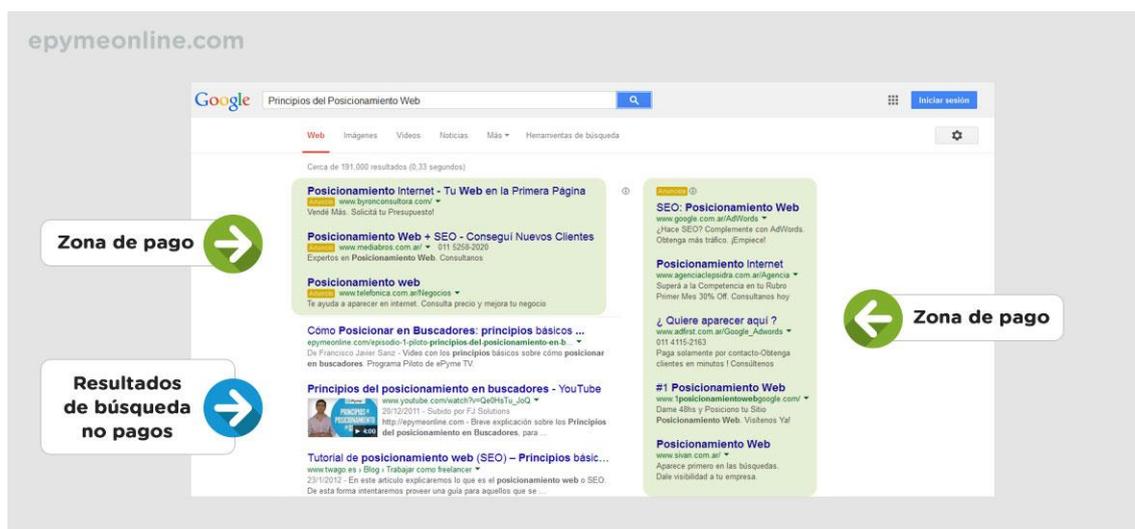


Figura 2. Resultados de búsqueda. Fuente: epymeonline.com

Al realizar una búsqueda en Google, como se puede observar en la figura 2, se obtienen diferentes resultados, donde parte de los resultados son de pago y el resto orgánicos. Por ello, si una empresa quiere que su sitio web aparezca entre los primeros resultados de distintas búsquedas, tiene dos opciones: pagar al buscador para que de esa forma aparezca en las zonas de pago o realizando un trabajo de SEO o posicionamiento web.

Este trabajo de posicionamiento web, puede llevar un período de cinco a ocho meses dependiendo de la competencia existente. Sin embargo, una vez alcanzada la primera o primeras posiciones, esta durará un tiempo largo dependiendo nuevamente de la competencia.

Para lograr este posicionamiento web, hay que tener en cuenta los factores de cada buscador que inciden o no a la hora de posicionar una página web. En el caso del buscador de Google, éste posee más de 200 factores además de realizar más de 500 ajustes o cambios al año en su proceso.

Por todo esto, se puede decir que los factores de mayor importancia y que se están teniendo en cuenta en mayor medida para un posicionamiento web a largo plazo son:

- **Autoridad:** es la popularidad de una web, cuanto más popular sea una página o una web más valiosa es la información que contiene. Un motor de búsqueda como Google suele tener en más en cuenta este factor ya que se basa en la propia experiencia del usuario. Cuanto más se comparta un contenido es que a más usuarios les ha parecido útil.



- **Relevancia:** es la relación que tiene una página frente a una búsqueda dada, esto no es simplemente que un sitio contenga repetido muchas veces el término buscado (al principio funcionaba así) sino que un motor de búsqueda se basa en cientos de factores *on-site* para determinar esto.
- **Rapidez:** es el hecho de que la página de un sitio web cargue lo más rápido posible tanto en dispositivos móviles como en ordenadores de escritorio. Este factor es más reciente comparado con los demás factores ya que ha sido impulsado por la navegación desde móviles y *tablets* debido a que la navegación por estos dispositivos está creciendo a un ritmo muy alto, además de que es lo que quiere y necesita el usuario, es decir, resultados rápidos.

Por otro lado, el posicionamiento web se puede dividir en dos grandes grupos: el *on-site* y el *off-site*.

- **On-site:** este grupo se preocupa por la relevancia, asegurándose de que el sitio web está optimizado para que el motor de búsqueda pueda leer y “entender” el contenido principal el sitio web correctamente. Dentro de este grupo se incluye la optimización de *keywords* (palabras clave), tiempo de carga, experiencia de usuario, optimización de código y formato de las URLs (*Uniform Resource Locator*).
- **Off-site:** este grupo se centra en factores externos al sitio web en el que se trabaja. Los factores más importantes en los que se centra este grupo son el número y la calidad de los enlaces, presencia en las redes sociales, menciones en medios locales, autoridad de la marca y rendimiento en los resultados de búsqueda, es decir, la cantidad de CTR (proporción de clics o ratio de cliqueo) que tengan los resultados en un motor de búsqueda.

Por otra parte, el SEO también se puede diferenciar en si se sigue o no las recomendaciones del motor de búsqueda: *Black Hat SEO* o *White Hat SEO*.

- **Black Hat SEO:** se denomina así al intento de mejorar el posicionamiento en buscadores de un sitio web mediante técnicas poco éticas o que contradicen las directrices del motor de búsqueda. Algunos ejemplos de este tipo de SEO son el *Cloaking*, *Spinning*, *SPAM* en foros y comentarios de blogs o el *Keyword Stuffing*. El *Black Hat* puede proporcionar beneficios a corto plazo, pero generalmente es una estrategia arriesgada, sin continuidad a largo plazo y que no aporta ningún valor.
- **White Hat SEO:** Al contrario que el anterior este tipo de SEO, consiste en todas aquellas acciones éticamente correctas y que cumplen las directrices de los motores de búsqueda para posicionar



un sitio web en los resultados de búsqueda. Dado que los buscadores dan una mayor importancia a los sitios que mejor responden a una búsqueda del usuario, el *White Hat* comprende las técnicas que buscan hacer más relevante un sitio para los buscadores a través de aportar valor para sus usuarios.

Por todo esto, el SEO es importante debido a que hace más útil el sitio web de una empresa tanto para los usuarios como para los motores de búsqueda, aunque estos últimos cada día son más sofisticados, pero aún no son capaces de ver una página web como lo hace un humano. El SEO es necesario para ayudar a los motores de búsqueda a entender sobre qué trata cada página y si es o no útil para los usuarios.

2.1.2 Motores de búsqueda

Un motor búsqueda ^[4] es una herramienta de hardware y software que indexa páginas web para que se pueda buscar contenido relacionado con las palabras clave que se introducen en el formulario. Cada motor de búsqueda utiliza robots o *spiders* que recorren las páginas web indexando su contenido dentro de enormes bases de datos que se pueden consultar, además de que cada motor utiliza sus propios algoritmos para que las búsquedas estén actualizadas según el contenido de cada página web.

El funcionamiento de un motor de búsqueda se puede resumir en dos pasos: rastreo e indexación.

Un motor de búsqueda recorre la web rastreando con los *spiders*, éstos recorren todas las páginas a través de los enlaces (de ahí la importancia de una buena estructura de enlaces) al igual que haría cualquier usuario al navegar por el contenido de la web, pasan de un enlace a otro y recopilan datos sobre esas páginas web que proporcionan a sus servidores.

El proceso de rastreo empieza con una lista de direcciones web de rastreos anteriores y de *sitemaps* proporcionada por otras páginas web. Una vez acceden a estas webs, los *spiders* buscan enlaces a otras páginas para visitarlas. Los *spiders* detectan especialmente a los sitios nuevos y a los cambios en las webs existentes.

Son los propios *spiders* los que deciden qué páginas visitar, con qué frecuencia y cuánto tiempo van a rastrear esa web, por eso es importante tener un tiempo de carga óptimo y un contenido actualizado.

⁴ Motor de búsqueda. Artículo de Wikipedia. Consultado en agosto de 2016. https://es.wikipedia.org/wiki/Motor_de_b%C3%BAsqueda



Una vez que un *spider* ha rastreado una página web y ha recopilado la información necesaria, estas páginas se incluyen en un índice donde se ordenan según su contenido, su autoridad y su relevancia. De este modo, cuando se haga una consulta al motor de búsqueda le resultará mucho más fácil mostrarnos los resultados que están más relacionados con nuestra consulta.

Al principio los motores de búsqueda se basaban en el número de veces que se repetía una palabra en una página, al hacer una búsqueda rastreaban en su índice esos términos para encontrar qué páginas los tenían en sus textos, posicionando mejor la que más veces lo tenía repetido. Actualmente, los motores de búsqueda son más sofisticados, y basan sus índices en cientos de aspectos diferentes como la fecha de publicación, si contienen imágenes, vídeos o animaciones, micro formatos, etc. Ahora dan más prioridad a la calidad del contenido.

Finalmente, una vez que las páginas son rastreadas e indexadas, llega el momento en el que actúa el algoritmo de cada motor de búsqueda: los algoritmos son los procesos informáticos que deciden qué páginas aparecen antes o después en los resultados de búsqueda. Una vez realizada la búsqueda, en cuestión de milisegundos, los algoritmos son capaces de buscar en los índices y sabes cuáles son las páginas más relevantes teniendo en cuenta los cientos de factores de posicionamiento.

2.1.3 Fuentes de datos

Para la obtención de los indicadores y más tarde su tratamiento y visualización se ha partido de tres sitios que ofrecen datos estadísticos, relevancia en Internet y posicionamiento web sobre el sitio web que se consulta. Estos tres sitios son los siguientes:

- Alexa ^[5]
- SimilarWeb ^[6]
- MajesticSEO ^[7]

Una vez indicados cuáles son las fuentes, se procederá a describir las características de cada una e indicar qué indicadores se han obtenido de cada una.

⁵ Alexa. <http://www.alexa.com/>

⁶ SimilarWeb. <https://www.similarweb.com/>

⁷ MajesticSEO. <https://es.majestic.com/>

La primera fuente que se ha tenido en cuenta es Alexa, una compañía subsidiaria de Amazon que fue fundada en 1996 y que recibe su nombre por la famosa biblioteca de Alejandría donde se almacenaba la mayor colección de conocimiento de la antigüedad, haciendo el símil con la actual Internet. Lleva elaborando un ranking de posicionamiento de sitios desde su fundación y es un punto de referencia a tener en cuenta por todos los que desean tener una idea de qué posición ocupan en el mundo con su sitio web en función de la cantidad de visitas que recibe.

Para evaluar su ranking, Alexa hace uso de un *plugin* (anteriormente una barra de herramientas) que se instala en los navegadores web y ofrece información relativa del sitio web en el que se esté navegando en el momento. A su vez, gracias a este *plugin* obtiene la información necesaria acerca de las visitas que se están produciendo en un determinado sitio web.

Una vez analizadas las visitas de los *plugins* de los millones de usuarios que la tengan, Alex almacena y compara entre sí para generar medidas estadísticas que posicionan a cada sitio web dentro de su ranking.

Al depender de los usuarios que instalan este *plugin*, su grado de acierto no es totalmente objetivo. Las visitas de los usuarios concedores de dicha herramienta suelen centrarse siempre en los mismos sitios, por tanto no es un medidor objetivo ya que atiende más a las costumbres de navegación de sus usuarios que al uso global que se puede realizar de un mismo sitio web en sí.

Se trata, pues, más de una muestra que de un todo, desde luego la muestra es representativa del todo, pero no es objetivamente el valor global o definitivo. Considerado como ayuda, estimación o muestra representativa es desde luego un valor útil de facto y reconocida por la mayor parte de los profesionales del medio.

El ranking de Alexa varía constantemente, pero la publicación de los indicadores en su web se actualiza para cada sitio una vez al día. Para las webs relativamente jóvenes o que nunca han sido indexadas por sus sistemas de análisis de información, una forma de que comience a visualizarse en el ranking es consultarla en Alexa o instalarse el *plugin* en el navegador para que la tenga en cuenta para futuras actualizaciones de rankings.

Finalmente, de los 2744 sitios de bodegas de vinos españoles, se obtienen mediante consulta a Alexa, los siguientes indicadores (si es que no son nulos):

- Posición en el ranking global.
- Posición en el ranking del país de origen, en este caso España.
- Bounce rate: tiempo medio en el que el usuario permanece en la página web hasta que se va.
- Daily pageviews per visitor: cantidad de páginas visitadas en el sitio web por el usuario diariamente.

- Daily time on site: tiempo diario que el usuario navega por el sitio web.
- Search visits: Porcentaje de usuarios que entran al sitio a través de motores de búsqueda.
- Indicadores demográficos según género, educación y lugar desde el cual han entrado (desde casa, trabajo o colegio) los usuarios al sitio.

En cuanto a la fuente de datos similarWeb, es posible obtener gracias a esta herramienta los indicadores relacionados con las visitas, origen de las mismas y aspectos sociales de cualquier web. Otra característica es que permite realizar una comparativa con un competidor para realizar un análisis de los indicadores de ambos sitios web.

Esta herramienta, permite ver la evolución de las visitas recibidas durante el último año, así como el tiempo medio de permanencia, el porcentaje de rebote y el ranking dentro de diversas regiones en todo el mundo. A su vez, es posible averiguar desde qué lugar entran los usuarios al sitio web, si desde un buscador, acceso directo, redes sociales o enlaces de otros sitios web. Esto último, muy importante a la hora de definir la estrategia de divulgación de una web además de por temas de marketing.

Asimismo, es posible obtener información de la localización geográfica de las visitas que recibe un sitio web, donde se indica la distribución de los visitantes por países mostrando la región con mayor cantidad de visitantes y el porcentaje en diversas regiones del planeta.

Aunque hoy en día las redes sociales suelen liderar las referencias a sitios web, es importante conocer qué páginas enlazan a nuestra web y desde dónde llega más tráfico. Sin embargo, los buscadores son, y seguirán siendo, el objetivo de la mayoría de los profesionales dedicados a la divulgación de sitios web. Aunque las técnicas SEO sean cada vez más complejas, es importante conocer los motivos por los cuales alguien ha encontrado un sitio en Google o en Bing.

Finalmente, de los 2744 sitios de bodegas de vinos españoles, se obtienen mediante consulta a similarWeb, los siguientes indicadores (si es que no son nulos):

- Posición en el ranking global.
- País de origen del sitio web.
- Posición en el ranking del país de origen.
- Temática del sitio web.
- Posición en el ranking según la temática.
- Visitas totales de la página.
- Duración media de las visitas.
- Número de páginas que se acceden por visita al sitio web.
- Porcentaje de aquellas visitas en las que sólo se visita una página del sitio web para luego pasar a otro sitio web.

- Países y su porcentaje desde donde se visita el sitio web.
- Lugar desde donde entran los usuarios al sitio web. Desde acceso directo, desde otros sitios web, desde buscadores, desde redes sociales o desde un mensaje de correo electrónico.
- Los cinco sitios desde donde se accede mayoritariamente al sitio web en cuestión.
- Los cinco sitios que se acceden una vez visitado el sitio web.
- Porcentaje indicando las búsquedas orgánicas y las de pago.
- Las cinco redes sociales, desde las cuales se accede al sitio web.

Por último, la fuente de datos majesticSEO, es una herramienta de posicionamiento web que inspecciona páginas web y mapea internet con el fin de ofrecer la mayor base de datos comercial de enlaces en el mundo.

La herramienta permite tiene la posibilidad de construir enlaces (*link building*). Esta técnica consiste en conseguir que otras páginas web enlacen a la página que interesa que los buscadores consideren relevante y la posicionen mejor en sus rankings. La técnica puede hacerse de manera natural, cuando otros sitios webs enlazan sin previo acuerdo por algún hecho o dicho. O bien de manera artificial, cuando se simula que los enlaces se han conseguido de manera natural.

Definitivamente, esta técnica favorece la posibilidad de medir la demanda y cantidad de personas que están buscando a través de una palabra clave, consigue efectividad del posicionamiento y posiciona la marca de la empresa o la empresa en sí.

Hay distintos tipos de *link building* ^[8], estos son:

- Comentarios en blogs.
- Firmas de perfil en foros.
- Conseguir un enlace desde otro blog.
- Intercambio de enlaces entre webmasters.
- Compra de enlaces.

⁸ Link building. Artículo de Wikipedia. Consultado en agosto de 2016. https://es.wikipedia.org/wiki/Link_building

Por otro lado, majesticSEO también permite la administración de reputación del sitio web y por tanto de la empresa y el desarrollo de tráfico. Esta herramienta se estructura en once secciones distintas:

- **Site explorer:** permite explorar un dominio o una URL con todo detalle. Aporta información sobre los subdominios, enlaces, imágenes, redireccionamientos, etc.
- **Backlink history:** permite averiguar el número de *backlinks* (enlaces externos) detectados por los robots de la herramienta. Es ideal para investigar cómo se comporta la competencia, se pueden comparar hasta cinco dominios simultáneamente.
- **Informes de mapa de enlaces:** Se pueden crear informes para comparar un sitio con la competencia añadiendo hasta 20 URLs.
- **Keyword checker:** permite obtener el volumen de búsqueda de cada palabra clave, además se puede buscar en el índice una palabra clave y determina cuántas veces aparece en un sitio. Para analizar estas palabras clave majesticSEO utiliza datos de búsquedas orgánicas.
- **Link profile fight:** permite comparar de forma gráfica los perfiles de *backlinks* de dos sitios.
- **Bulk backlink checker:** permite contabilizar los *backlinks* para muchos dominios.
- **Clique hunter:** permite encontrar un listado de *clicks* que enlazan con una lista de dominios.
- **Majestic million:** un ranking de un millón de páginas ordenados por índices de la propia herramienta y actualizados con frecuencia.
- **Comparator:** permite comparar hasta cinco dominios.
- **Neighbourhood checker:** presenta una lista de sitios alojados en una misma IP (*Internet Protocol*). Esta herramienta puede servir para conocer mejor la empresa de *hosting* contratada y qué tipos de páginas alojan bajo una misma IP.

Finalmente, de los 2744 sitios de bodegas de vinos españoles, se obtienen mediante consulta a majesticSEO, los siguientes indicadores (si es que no son nulos):

- **Trust flow:** indica el número de *clicks* desde un conjunto de sitios de confianza para acceder a una URL o un dominio determinados.
- **Citation flow:** indica el número de menciones hechas de una URL o un dominio determinados.
- Número de *backlinks* externos.
- Número de dominios de referencia.
- Número de IPs de referencia.

- Número de subnets de referencia.
- Número de URLs indexadas

2.1.4 Tecnologías utilizadas

Durante el proceso de extracción de datos, se han utilizado distintas tecnologías, según la fuente de datos, ya que cada una de ellas restringe de una manera u otra. A su vez, para almacenar los datos obtenidos también se han utilizado diferentes. También, algunas de las tecnologías aplicadas no se acabaron utilizando finalmente debido a que no ofrecían los resultados esperados y por ello se acabó utilizando otras.

Las tecnologías de extracción de datos utilizadas en el proceso, tienen una cosa en común y es que los datos de interés se obtienen a partir del HTML ^[9] (*HyperText Markup Language*) de la página web. Para obtener estos datos de las páginas web se ha utilizado XPath (*XML (eXtensible Markup Language) Path Language*).

XPath ^[10] es un lenguaje que permite recuperar información de un documento XML ^[11] o HTML, haciendo un recorrido a través de los elementos que los compone. Es decir, evaluar una expresión XPath consiste en buscar si hay elementos que satisfagan el recorrido indicado.

Un documento XML o HTML, es procesado por un analizador (o *parser*) construyendo un árbol de nodos. Este árbol comienza con un elemento raíz, que se diversifica a lo largo de los elementos que cuelgan de él y acaba en nodos hoja, que contienen sólo texto, comentarios, instrucciones de proceso o incluso que están vacíos y sólo tienen atributos.

⁹ HTML. Artículo de Wikipedia. Consultado en agosto de 2016. <https://es.wikipedia.org/wiki/Html>

¹⁰ XPath. Artículo de Wikipedia. Consultado en agosto de 2016. <https://es.wikipedia.org/wiki/XPath>

¹¹ XML. Artículo de Wikipedia. Consultado en agosto de 2016. https://es.wikipedia.org/wiki/Extensible_Markup_Language



Existen distintos tipos de nodos en un árbol a partir de un documento XML o HTML, a saber: raíz, elemento, atributo, texto, comentario e instrucción de procesamiento. El nodo raíz se identifica por /, sin embargo, no es el elemento raíz. El nodo elemento es cualquier elemento del documento, todos tienen padre y tienen hijos, los cuales serían otros nodos elementos, nodos texto, nodos comentario o nodos de instrucciones de proceso. Los nodos texto hacen referencia a todos los caracteres del documento que no están marcados con alguna etiqueta, estos nodos no tienen hijos. Los nodos atributo constan de un nombre, un valor (siempre una cadena) y un posible espacio de nombres. Y los nodos comentario y de instrucciones de proceso, ambos se explican por su nombre.

La extracción de datos de este trabajo consiste en obtener datos durante varios meses de las 2744 bodegas de vino españolas a partir de las fuentes de datos Alexa, similarWeb y majesticSEO.

Para llevar a cabo lo anterior se partió del programa *import.io* ^[12] la cual es una herramienta que permite extraer datos de varias webs de manera estructurada. Es un programa totalmente gratuito y que ofrece gran cantidad de posibilidades a la hora de obtener los datos necesarios.

Su funcionamiento es el siguiente:

- 1) Se elige la herramienta *Extractor*.
- 2) Se selecciona la opción de *Single URL* para extraer al principio sólo de un sitio web, se introduce una URL de por ejemplo Alexa que sería algo así y se ejecuta la consulta:

<code>http://www.alex.com/siteinfo/upv.es</code>
--

- 3) Se abrirá la página en cuestión como si de un navegador web se tratara, una vez cargada la página se activaría la opción de seleccionar los datos que se quieren extraer.
- 4) Para seleccionar un dato a extraer bastaría con pulsar sobre cualquier texto, tras lo cual éste se ilumina mostrando de esta forma que se ha seleccionado. Para seleccionar varios datos para su posterior extracción habría que crear más columnas.
- 5) Una vez seleccionados todos los datos necesarios, se guarda el *extractor*. De esta forma, esa primera página cargada actuaría como plantilla para las demás ya que todos los datos a extraer estarían localizados en el mismo lugar de la página web sólo que con distinto valor, por ello la página plantilla debería contener todos los datos existentes y que no fueran nulos o vacíos.

¹² Import.io. <https://www.import.io/>

- 6) Con la plantilla se iría a las opciones del *extractor* y se cambiaría la opción de *single URL* por *Bulk Extract*, esto quiere decir que ahora se pasaría a extraer los datos según la plantilla de varias webs automáticamente y no de una.
- 7) En este caso, se pegarían las 2744 URLs de bodegas de vino españolas aplicadas a la fuente de datos Alexa y a similarWeb, cuyas URLs serían algo así:

```
http://www.alex.com/siteinfo/ + sitio web
```

```
https://www.similarweb.com/website/ + sitio web
```

- 8) Se ejecutaría el *extractor* y al ser 2744 los sitios webs que tiene que analizar, tardaría pues entorno a la hora y media.
- 9) Finalmente, una vez acabado el proceso, existe la posibilidad de exportar los datos en diferentes formatos, entre los que se encuentran: JSON, CSV, hoja de Excel y en hojas de Google. En este caso se eligió el formato JSON por comodidad y al ser un formato más familiar.

Para el caso de la extracción de datos de similarWeb, hubo un problema y es que aquellos sitios donde no tenía ningún dato el programa de import.io daba error con lo cual ese sitio web no aparecía en los resultados de la extracción. Para solucionar esto, en el archivo final con todos los datos se indicó los que habían fallado.

El formato de texto JSON ^[13] (*JavaScript Object Notation*) se caracteriza por ser ligero para el intercambio de datos. Describe los datos con una sintaxis dedicada que se usa para identificar y gestionar datos.

¹³ JSON. Artículo de Wikipedia. Consultado en agosto de 2016. <https://es.wikipedia.org/wiki/JSON>

En la siguiente figura se puede observar la sintaxis de JSON, además de ser un ejemplo de los datos sacados de un sitio web con el programa import.io:

```
{
  "pageurl": "http://www.alexa.com/siteinfo/7magnifics.com",
  "results": [
    {
      "audience_demographics_male": [
        "0",
        "0"
      ],
      "audience_demographics_nocollege": [
        "0",
        "0"
      ],
      "audience_demographics_college": [
        "0",
        "0"
      ],
      "audience_geography": "No data available for this site",
      "bounce_rate": "-",
      "audience_demographics_work": [
        "0",
        "0"
      ],
      "rank_in_spain": "--",
      "audience_demographics_somcollege": [
        "0",
        "0"
      ],
      "audience_demographics_home": [
        "0",
        "0"
      ],
      "daily_time_on_site": "-",
      "search_visits": "-",
      "audience_demographics_school": [
        "0",
        "0"
      ],
      "audience_demographics_graduateschool": [
        "0",
        "0"
      ],
      "daily_pageviews_per_visitor": "1.40",
      "audience_demographics_female": [
        "0",
        "0"
      ],
      "global_rank": "5,858,161"
    }
  ]
}
```

```

    }
  ]
}

```

Figura 3. Ejemplo JSON de Alexa

Una vez obtenidos todos los datos de cada uno de los sitios web en un archivo de tipo JSON, se procedería a pasar estos datos a una hoja de Excel para almacenar todos los meses extraídos en un mismo archivo y que la comparación entre ellos sea más sencilla. Para ello, se haría una conversión del archivo JSON obtenido anteriormente a un documento XML debido a que los métodos de lectura y reconocimiento son mejores para este formato que para el JSON en el lenguaje VBA script de Excel.

A su vez, XML se plantea como un lenguaje estándar para el intercambio de información entre diferentes programas de una manera segura, fiable y libre, ya que no pertenece a ninguna compañía. La tecnología XML, busca dar solución al problema de expresar información estructurada de la manera más abstracta y reutilizable posible. Que la información sea estructurada quiere decir que se compone de partes bien definidas, y que esas partes se componen a su vez de otras partes, entonces se tiene un árbol de trozos de información. Estas partes se llaman elementos y se las señala mediante etiquetas.

Una etiqueta consiste en una marca hecha en el documento, que señala una porción de éste como un elemento. Un pedazo de información con un sentido claro y definido.

```

<pages>
<pageurl>http://www.alexa.com/siteinfo/torres.es</pageurl>
<results>
<audience_demographics_male>35</audience_demographics_male>
<audience_demographics_male>0</audience_demographics_male>
<audience_demographics_nocollege>0</audience_demographics_nocollege>
<audience_demographics_nocollege>0</audience_demographics_nocollege>
<audience_demographics_college>100</audience_demographics_college>
<audience_demographics_college>98</audience_demographics_college>
<audience_geography>No data available for this site</audience_geography>
<bounce_rate>37.50%</bounce_rate>
<audience_demographics_work>45</audience_demographics_work>
<audience_demographics_work>0</audience_demographics_work>
<rank_in_spain>44,047</rank_in_spain>
<audience_demographics_somcollege>0</audience_demographics_somcollege>
<audience_demographics_somcollege>0</audience_demographics_somcollege>
<audience_demographics_home>100</audience_demographics_home>
<audience_demographics_home>56</audience_demographics_home>
<daily_time_on_site>2:47</daily_time_on_site>
<search_visits>26.10%</search_visits>
<audience_demographics_school>0</audience_demographics_school>
<audience_demographics_school>0</audience_demographics_school>

```

```
<audience_demographics_graduateschool>100</audience_demographics_graduate
school>
<audience_demographics_graduateschool>4</audience_demographics_graduatesch
ool>
<daily_pageviews_per_visitor>4.20</daily_pageviews_per_visitor>
<audience_demographics_female>100</audience_demographics_female>
<audience_demographics_female>98</audience_demographics_female>
<global_rank>577,361</global_rank>
</results>
</pages>
```

Figura 4. Ejemplo XML de Alexa

Para la conversión de un archivo JSON a un documento XML, se han utilizado la siguiente página, donde bastaría con copiar el contenido del archivo JSON y directamente realizaría la conversión a XML y también existe la posibilidad de realizar el proceso inverso, es decir, de un documento XML a un archivo JSON:

<http://www.utilities-online.info/xmltojson/>

Con el documento XML en mano, el siguiente paso consistiría en pasar los datos del XML a la hoja de Excel. Para realizar este paso, se hará uso del lenguaje VBA (*Visual Basic for Applications*) que leerá del documento XML e irá insertando los datos de un sitio web en cada fila mediante un algoritmo programado expresamente para ello.

Este método de extracción fue rápido a la hora de obtener los datos y sencillo de poner en marcha. Con este método se pudo conseguir los datos de febrero, marzo, abril y mayo para las fuentes de datos Alexa y similarWeb. Sin embargo, en junio el programa *import.io* empezó a fallar y llegó al punto de no poder obtener ningún dato. En el caso de Alexa por un fallo en el extractor, a lo cual se creó otro y tampoco funcionaba. Y en el caso de similarWeb el propio sitio web restringió la posibilidad de extraer mediante *scraping* información de su sitio web.

Para solucionar ambos problemas se hizo uso del lenguaje de programación Java, creando un programa orientado a objetos. Este programa realizaba las siguientes funciones:

- 1) Leer sitio web a partir de un fichero de texto en el que en cada línea se encontraba una dirección de Alexa para los datos de la bodega de vino correspondiente.
- 2) Obtenía los datos requeridos mediante expresiones XPath al documento HTML de la página web.
- 3) Finalmente, guardaba todos esos datos en un archivo XML, con la estructura vista en la *Figura 4*.

Una vez con el archivo XML, se importaba los datos de éste de la misma forma indicada anteriormente, es decir, con un algoritmo del lenguaje de programación de *Microsoft Office* que insertaba una línea en Excel para cada sitio web. La fuente de datos Alexa, cada aproximadamente 1000 direcciones obtenidas, restringía el poder continuar con la extracción, pero esta restricción duraba solamente un día, con lo cual en tres días se podía obtener los datos de las 2744 bodegas de vinos españolas.

Para el caso de similarWeb, se programó un programa similar en Java, al principio sacaba datos sin ningún inconveniente, pero al extraer unos pocos datos limitaba el proceso de extracción dejándolo parado. Aunque se dejase uno o dos días seguía de la misma manera, con lo que se llegó a la conclusión que similarWeb no dejaba obtener el documento HTML para su análisis. La misma situación se tuvo cuando se utilizó el lenguaje de programación PHP, al principio obtenía correctamente los datos, pero tras muy poco tiempo dejaba de obtenerlo. Con lo que se llegó a la conclusión que similarWeb no dejaba de ninguna manera realizar *scraping* sobre su sitio web. Debido a este problema no se pudieron obtener los datos del mes de junio de similarWeb.

Para solucionar este problema se optó por utilizar el mismo método que con majesticSEO, es decir, utilizar la *plugin* para el navegador web Firefox de nombre *iMacros*. *iMacros* ^[14] es una herramienta que graba acciones realizadas sobre el navegador Firefox y las convierte en código que posteriormente es posible usarlo para repetirlo de forma periódica.

Una vez creado el algoritmo, el funcionamiento de *iMacros* consiste en acceder visualmente a la página en cuestión, obtener mediante expresiones XPath los datos necesarios y almacenar esos datos en un archivo CSV (Valores Separados por Comas) para luego mediante Excel poder separar cada uno de los valores en su columna correspondiente. Desde un principio se utilizó este procedimiento para majesticSEO porque esta fuente de datos requería tener iniciada una sesión, y ninguno de las anteriores maneras daba la posibilidad de iniciar sesión ya que hacía falta realizar esto gráficamente y no por código.

Una vez con el archivo CSV, gracias al programa Excel sólo era necesario separar cada valor en una columna según la coma. Debido a la eficacia de este procedimiento se desarrolló un algoritmo similar pero aplicado a similarWeb, terminando en un resultado satisfactorio a la hora de extraer datos de esta fuente de datos.

¹⁴ iMacros. <http://imacros.net/>



Este método, aunque eficaz, el gran inconveniente que tiene es que tarda un 70% más que los anteriores debido a que tiene que esperar a que cada página que carga el navegador se termine por completo. Hablando de tiempos, si los otros métodos tardaban una media de veinte minutos en obtener datos de mil sitios web, el programa *iMacros* tarda alrededor de una hora y media en obtener los mismos datos debido, como se ha indicado antes, que se tiene que cargar por completo cada página web.

2.2 Manejo de datos

En este apartado, se explicará las diferentes alternativas que se tenía para elegir el lenguaje de programación apropiado para el manejo de datos, además de diferentes comparaciones entre ellos. Por su parte, una vez seleccionado el lenguaje se detallará qué entornos se han utilizado y las diferentes librerías que se han utilizado a lo largo del proceso del tratamiento de datos.

2.2.1 ¿Qué lenguaje de programación elegir?

A lo largo de los años han ido apareciendo diferentes lenguajes de programación orientados al tratamiento, análisis y visualización de datos de carácter científico. Por ello, se pueden indicar los siguientes lenguajes: Python, R, MATLAB, SAS, Stata y otros. Sin embargo, debido a la popularidad, a la extensa comunidad y que son de código abierto se realizará una comparación entre Python ^[15] y R ^[16].

La diferencia principal entre ambos es que Python es un lenguaje de programación de carácter general, mientras que R está concebido especialmente para el análisis estadístico de métricas financieras, aunque es perfectamente aplicable a la minería de datos, la investigación biomédica, tratamiento de datos climáticos, entre otros. Sin embargo, en Python se han ido desarrollando diferentes librerías que pueden solucionar problemas muy específicos a la hora de analizar datos científicos.

En cuanto al uso se refiere, Python lo utilizan aquellos programadores que quieren profundizar en el análisis de datos o aplicar técnicas estadísticas además de desarrolladores que se pasan a la ciencia de los datos. Mientras que el lenguaje R, es usado mayoritariamente por investigadores y académicos.

Ambos tienen una fuerte comunidad detrás, donde se puede destacar el sitio *Stackoverflow* ^[17], la cual es una fuente para resolver dudas y solucionar errores de todo tipo. A su vez, existe una gran cantidad de usuarios que aporta código y documentación en todos los ámbitos de los lenguajes.

¹⁵ Python. <https://www.python.org/>

¹⁶ Lenguaje de programación R. <https://www.r-project.org/>

¹⁷ Foro de preguntas y respuesta, Stackoverflow. <https://stackoverflow.com/>



Python tiene una sintaxis clara al estilo parecido al idioma inglés que hace que la depuración y la comprensión del código sea más sencilla, mientras que R tiene una sintaxis poco convencional que puede resultar más complicado de entender, especialmente si conoce otros lenguajes de programación. Por ello la curva de aprendizaje en R puede ser mayor dependiendo de los conocimientos sobre la ciencia de los datos que se tenga, en cambio para Python es menor tanto si se viene o no de otros lenguajes de programación.

Finalmente, en las siguientes gráficas ^[18] se puede observar el uso de las diferentes herramientas de análisis que utilizan los profesionales de distintas industrias, además de cuáles prefieren ellos.

En la siguiente gráfica se pueden observar, las tecnologías que prefieren los analistas profesionales, si SAS, Python o R.

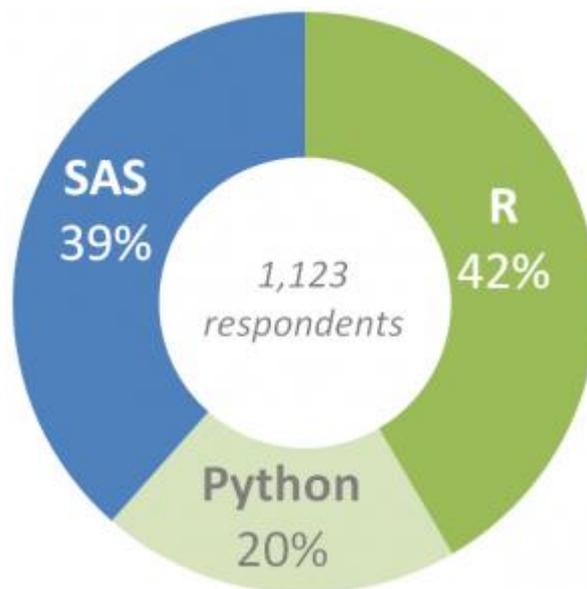


Figura 5. ¿Qué lenguaje de programación elegir? Fuente: burtchworks.com

¹⁸ Encuesta a profesionales sobre qué lenguaje utilizan para analizar datos. www.burtchworks.com/2016/07/13/sas-r-python-survey-2016-tool-analytics-pros-prefer/

A continuación, se puede observar, a lo largo de tres años cómo ha ido evolucionando la opinión de los analistas y se puede ver cómo las herramientas de código abierto se están imponiendo poco a poco a la de pago (SAS). También comentar que en los años anteriores no se tuvo en cuenta la herramienta Python, sólo a partir de este año.

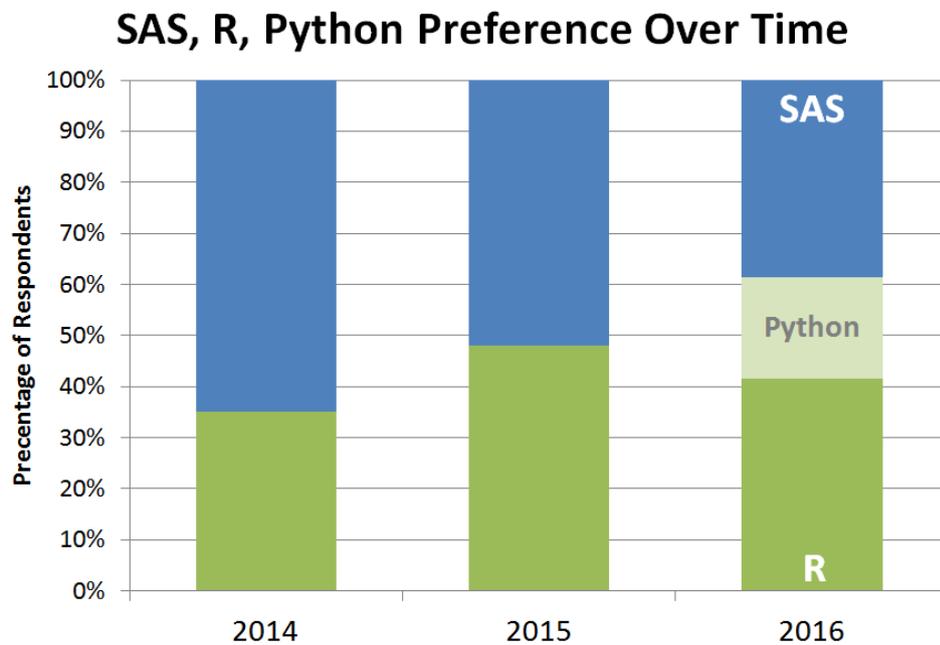


Figura 6. Preferencias de los analistas a lo largo de los años. Fuente: *burchworks.com*

Por otra parte, en cuanto a qué herramientas se utilizan en la industria, se puede ver como en la industria informática se opta más por herramientas de código abierto y en otras sin embargo se opta por la herramienta de pago SAS.

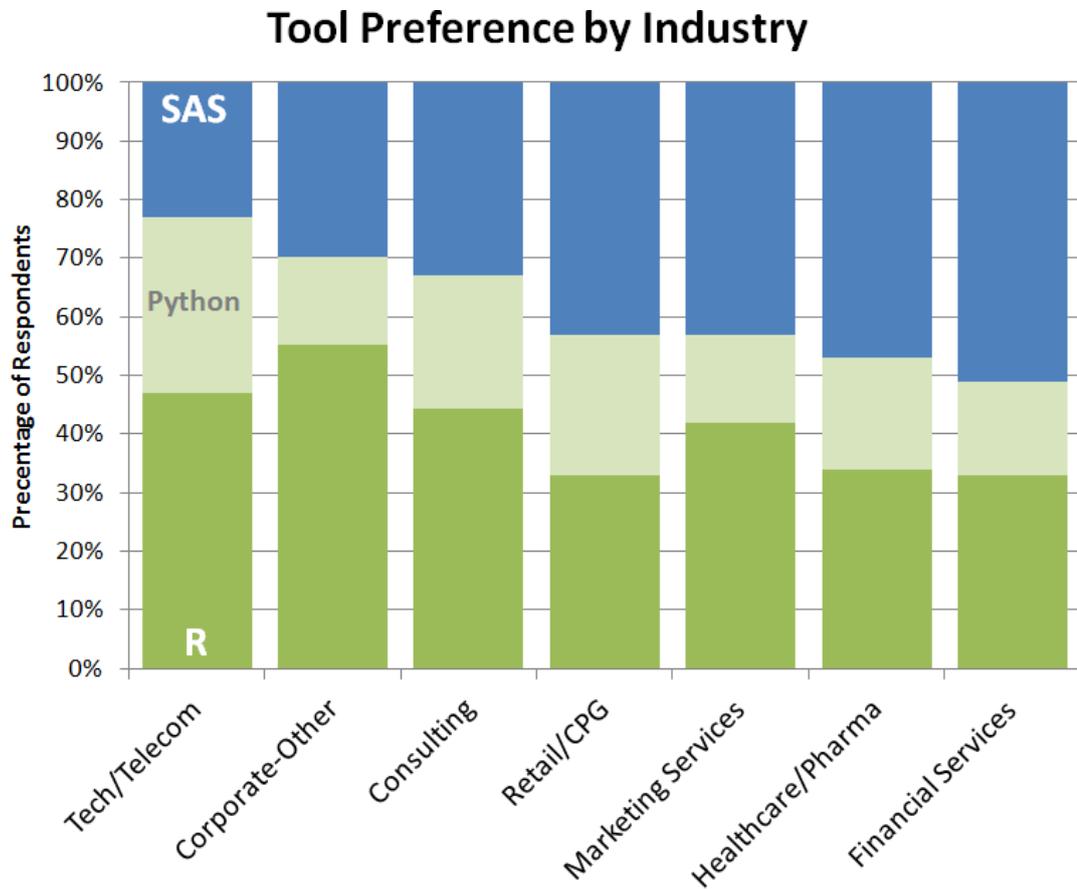


Figura 7. Lenguajes de programación en la industria. Fuente: *burtchworks.com*

En cuanto a qué herramientas se usan en el ámbito de la educación, se puede ver que los doctores prefieren el uso de R, mientras que los graduados optan más por SAS

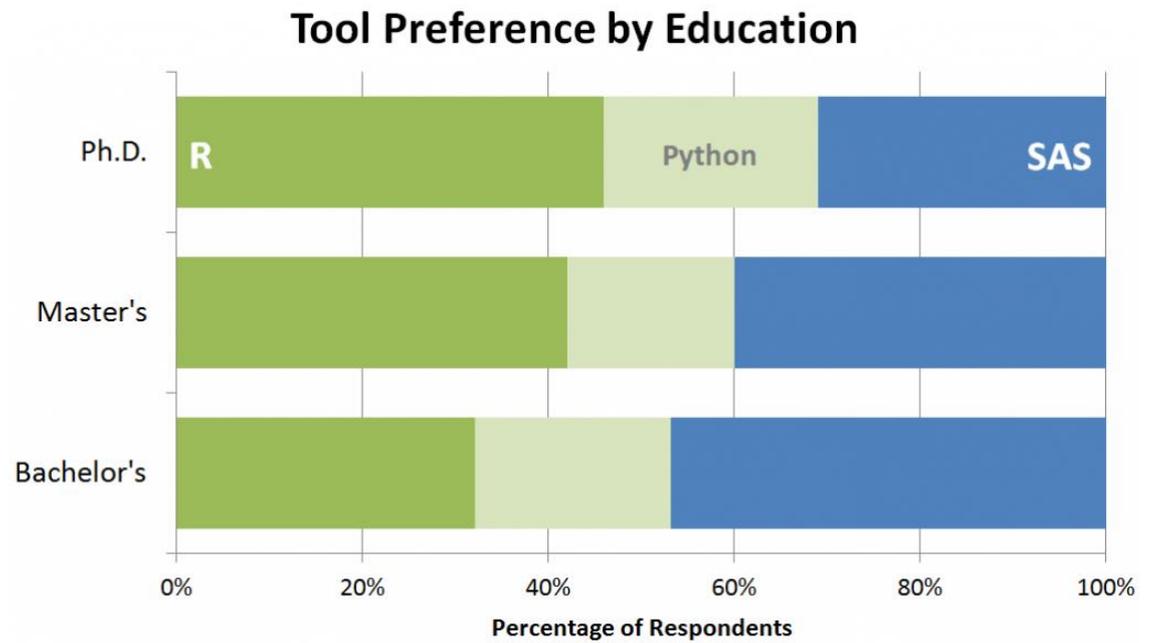


Figura 8. Lenguajes de programación en la educación. Fuente: burtchworks.com

Por otro lado, se puede observar como en los últimos años las universidades lanzan más programas empleando herramientas de código abierto, en este caso R y Python, en vez de las de pago. Sin embargo, los analistas que tienen más de 16 años de experiencia prefieren las de pago, en este caso SAS.

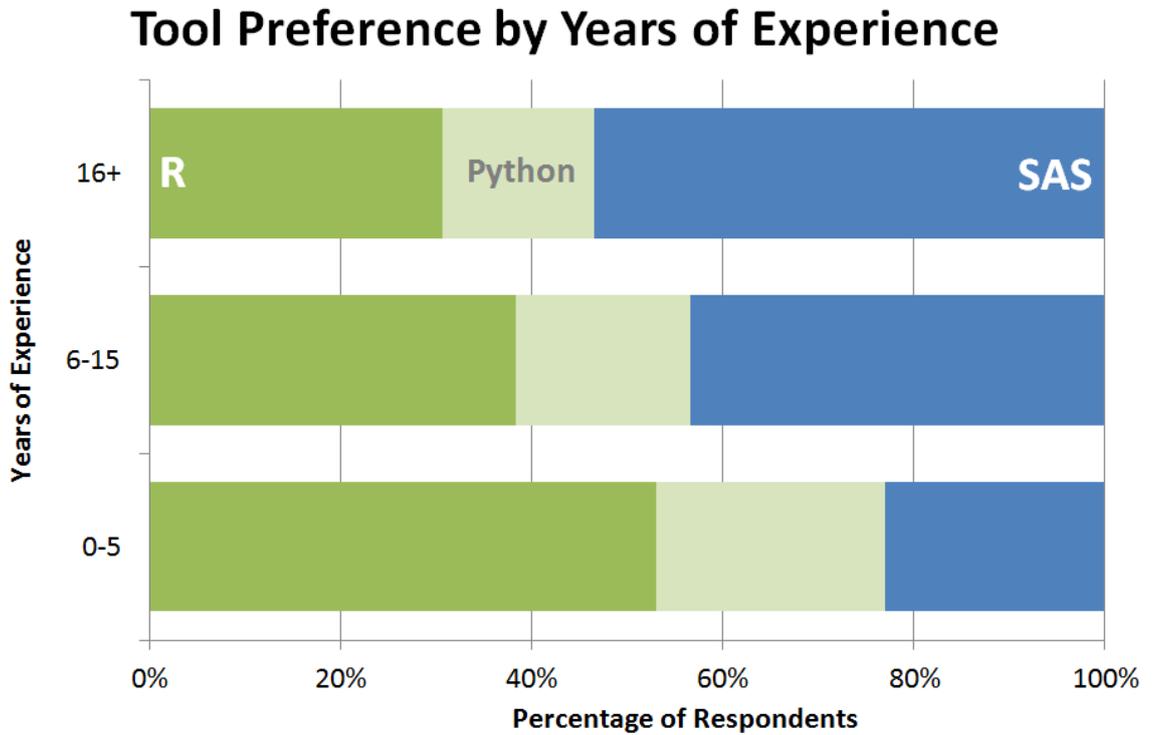


Figura 9. Preferencia de lenguaje de programación según años de experiencia. Fuente: burtchworks.com

En la siguiente figura, se puede ver como Python posee la mayoría de apoyo entre los científicos de datos y como SAS cuenta con apenas un 3%. Esto puede ser debido a sus limitaciones en la construcción de herramientas personalizadas que muchos especialistas usan para gestionar datos no estructurados. Entre los analistas predictivos, el uso de R y SAS se mantiene relativamente uniforme.

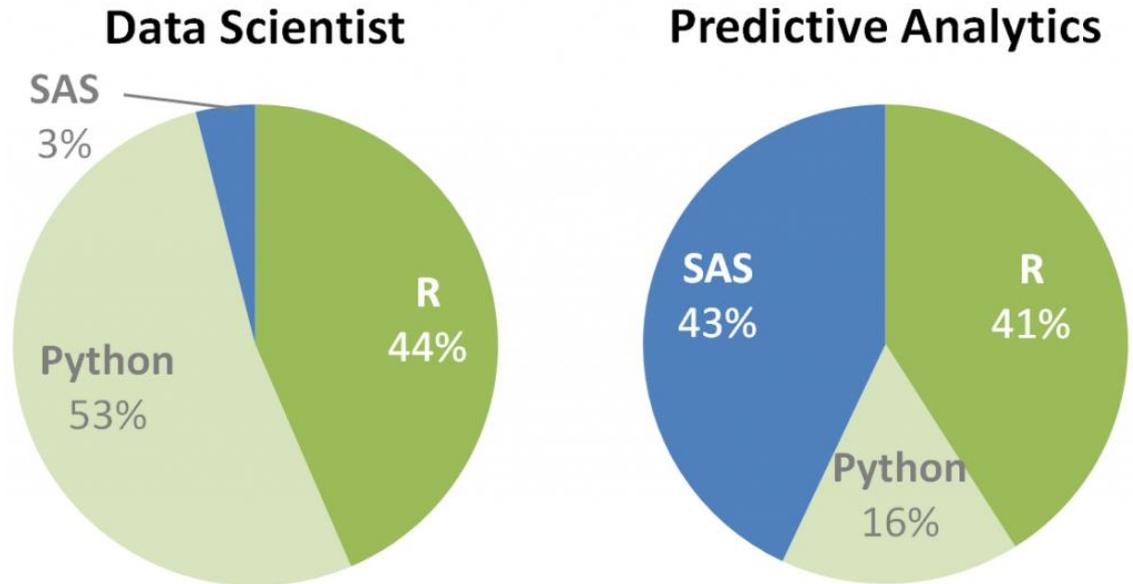


Figura 10. Lenguajes de programación entre los científicos y analistas.
Fuente: burtchworks.com

Finalmente, en esta gráfica se pueden observar los datos de la anterior figura, pero indicando además los datos de 2015.

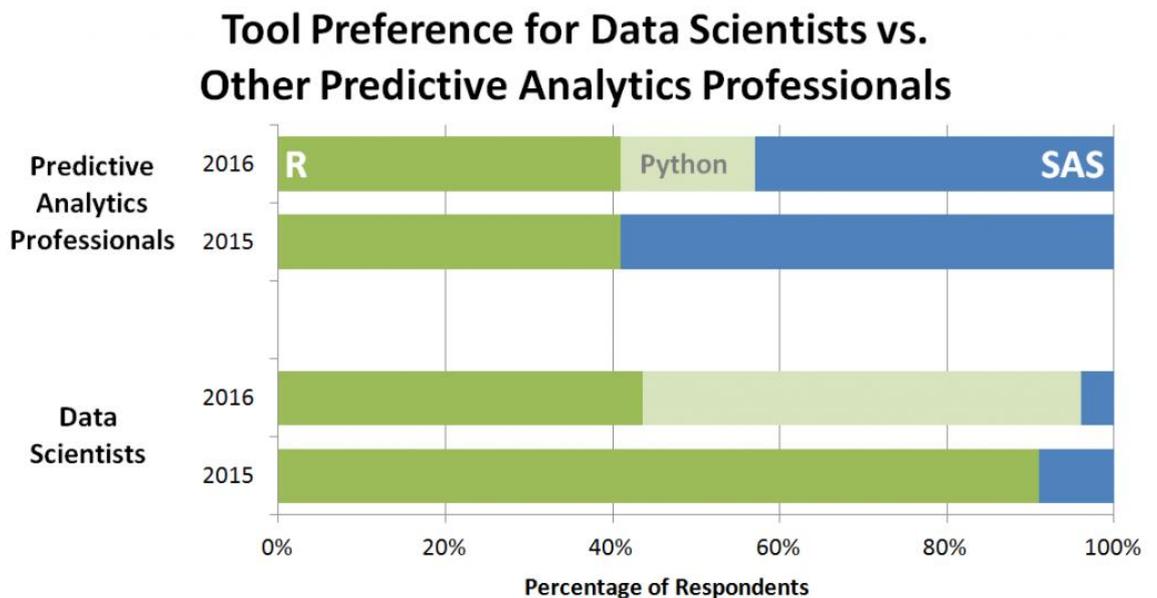


Figura 11. Lenguajes de programación entre los científicos y analistas a lo largo de los años. Fuente: burtchworks.com

Concluyendo, el lenguaje de programación que se usará para el análisis de datos y la visualización de los mismos, entre R y Python, será Python. Los motivos para esta elección son los siguientes:

- Sintaxis clara y sencilla de entender.
- Curva de aprendizaje relativamente baja viniendo de lenguajes de programación como Java, javascript, PHP, VB script, C#
- Gran comunidad para la resolución de dudas y problemas, aportaciones de usuarios y documentación completa.

2.2.2 Entorno y librerías

El entorno Enthought Canopy ^[19] es el elegido debido a que es un entorno de análisis de Python integral que proporciona una fácil instalación de más de 450 paquetes de análisis científico para Python, creando con ello una plataforma robusta donde se puede explorar, desarrollar y visualizar todo tipo de datos. Además, este entorno incluye la versión 2.7.11 de Python.

En cuanto a las librerías, las más populares para el análisis de datos científicos y que ofrecen posibilidades de todo tipo son las siguientes: NumPy, pandas, IPython.

NumPy ^[20], abreviatura de *Numerical Python*, es una librería que agrega mayor soporte para vectores y matrices, constituyendo un conjunto de funciones matemáticas para operar con esos vectores y matrices. La librería provee entre otra funcionalidad:

- Funciones para la realización de cálculos de elementos entre matrices.
- Herramientas para la lectura y escritura de conjuntos de datos basados en matrices.
- Operaciones de álgebra lineal, la transformada de Fourier y la generación de números aleatorios.
- Herramientas para la integración de código de C, C++ y Fortran a Python.

pandas ^[21], es otra librería para Python que ofrece estructuras y operaciones de datos para la manipulación de tablas numéricas. A su vez, combina las características de computación de NumPy con las capacidades de manipulación de datos flexibles de las hojas de cálculo y las bases de datos relacionales (como SQL). Proporciona, además, funcionalidad de indexación para que sea más sencillo remodelar, dividir, realizar agregaciones y seleccionar subconjuntos de datos.

IPython ^[22] es un componente en el conjunto de herramientas de Python que sirve de unión entre distintas librerías de Python. Es una consola de comandos mejorada de Python diseñada para acelerar la escritura, las pruebas y la depuración del código Python. Es especialmente útil para trabajar con datos de forma interactiva y visualizarlos con la librería para la generación de gráficos matplotlib.

¹⁹ Enthought Canopy. <https://www.enthought.com/products/canopy/>

²⁰ NumPy. <http://www.numpy.org/>

²¹ pandas. <http://pandas.pydata.org/>

²² IPython. <https://ipython.org/>



3. Resultados

Una vez obtenidos todos los datos, en este capítulo se pasaría a la parte del tratamiento de los mismos para analizar los datos relevantes y desechar los prescindibles con el fin de llegar a la parte de visualización con los datos preparados para ser mostrados gráficamente y así poder sacar conclusiones.

Por su parte, una vez tratados estos datos se realizará una representación de los mismos para que mediante gráficas y tablas sea más fácil y cómodo sacar conclusiones sobre ellos.

3.1 Tratamiento de datos

Para tener los datos listos a la hora de representarlos, primero hay que realizar diversos cálculos entre ellos, desechar aquellos que sean irrelevantes y tener en cuenta sólo aquellos que pueden ser de utilidad. Por tanto, primero se describe cómo empezar con el entorno para tenerlo todo preparado a la hora de manejar los datos. Seguidamente, se pasaría a detallar y explicar los diferentes algoritmos implementados para tratar estos datos.

3.1.1 Puesta a punto

Una vez descrito qué herramientas se van a utilizar durante el proceso de análisis y tratamiento de datos, se pasaría a describir en este apartado desde qué datos se parte, su formato y demás características.

Primero, indicar que el sistema operativo en el que se ejecutará todo el proceso de análisis y visualización será Windows 8.1 de 64 bits. Con la versión de Python 2.7.11 bajo el entorno Enthought Canopy antes explicado. El directorio de trabajo contendrá los datos de los diferentes meses en formato JSON y los diferentes scripts de Python necesarios para el tratamiento de los datos.

Tal como se terminó en el capítulo de extracción de datos, éstos se habían guardado para su visualización de forma interactiva en Excel, donde cada hoja representaba un mes de extracción de datos. Para pasarlos al formato JSON, mediante un script desarrollado en VB Script, se leían los datos correspondientes de las diferentes bodegas y se guardan en un archivo externo con extensión JSON. Además de lo anterior, los datos de cada mes están en un archivo externo, es decir, que cada mes tiene su archivo con los datos no estando todos los datos de todos los meses en un archivo.

Un ejemplo del formato seguido para indicar los datos de las bodegas es el siguiente:

```
{
  "site": "torres.es",
  "global_rank_alexa": 526629,
  "rank_spain_alexa": 19281,
  "bounce_rate_alexa": 33.7,
  "daily_pageviews_per_visitor_alexa": 5.2,
  "daily_time_on_site_alexa": 3.27,
  "search_visits_alexa": 25,
  "audience_geography_alexa": [{
    "country": "Spain",
    "percent": "78.3",
    "rank": "19281"
  }, {
    "country": "United States",
    "percent": "6.0",
    "rank": "1026401"
  }],
  "ad_male": 20,
  "ad_female": 99,
  "ad_no_college": 0,
  "ad_some_college": 0,
  "ad_graduate_school": 57.5,
  "ad_college": 97.5,
  "ad_home": 61.5,
  "ad_school": 0,
  "ad_work": 30,
  "global_rank_sw": 856569,
  "country_rank_country_sw": "Spain",
  "country_rank_rank_sw": 50721,
  "category_rank_cat_sw": "Food and Drink > Beverages",
  "category_rank_rank_sw": 753,
  "total_visits_sw": 204000,
  "time_on_site_sw": 2.11,
  "page_views_sw": 3.63,
  "bounce_rate_sw": 37.98,
  "traffic_by_countries_sw": [{
    "country": "Spain",
    "percentage": "49.96"
  }, {
    "country": "Finland",
    "percentage": "16.71"
  }, {
    "country": "United States",
    "percentage": "4.24"
  }, {
    "country": "Mexico",
    "percentage": "3.95"
  }
}
```

```
}, {
  "country": "Lithuania",
  "percentage": "2.91"
}],
"tf_direct_sw": 31.56,
"tf_referrals_sw": 15.06,
"tf_search_sw": 50.96,
"tf_social_sw": 1.27,
"tf_mail_sw": 1.16,
"tf_display_sw": 0,
"top_referring_sites_sw": [{
  "site": "Delfi.lt"
}, {
  "site": "Winestyle.ru"
}, {
  "site": "Mishanita.ru"
}, {
  "site": "Appcatalogo.ifema.es"
}, {
  "site": "Jacquelinevanpaassen.com"
}],
"top_destination_sites_sw": [{
  "site": "Clubtorres.com"
}, {
  "site": "Torresearch.com"
}, {
  "site": "Marimarestate.com"
}, {
  "site": "Shop.clubtorres.com"
}, {
  "site": "Torres.isoco.com"
}],
"organic_search_sw": 99.86,
"paid_search_sw": 0.14,
"social_sw": [{
  "social_network": "Facebook",
  "percentage": "85.85"
}, {
  "social_network": "Youtube",
  "percentage": "7.09"
}, {
  "social_network": "Linkedin",
  "percentage": "7.06"
}],
"trust_flow": 0,
"citation_flow": 0,
"backlinks_externos": 0,
"dominios_referencia": 0,
"ip_referencia": 0,
"subnets_referencia": 0,
```

```
"url_indexadas": 0
}
```

Antes que nada, hay que hacer un breve apunte y es que hay meses que no se pudieron obtener los datos de Alexa, similarWeb y MajesticSEO, debido a que se tuvieron que rehacer los programas de extracción por restricciones de los sitios web. Por ello, se dispone de los siguientes datos de cada sitio:

- Alexa: todos los meses (de febrero a julio).
- SimilarWeb: todos los meses menos mayo.
- MajesticSEO: todos los meses menos febrero.

Por otro lado, hay varios indicadores que pueden no aparecer en muchas bodegas como son “traffic_by_countries_sw” o “audience_geography_alexas”, entre otros, debido a que no había datos referentes a esos indicadores.

Finalmente, una vez instalado el entorno Enthought Canopy, bastaría con ejecutar el *Canoy command prompt* que es una consola de comandos y desde donde se ejecutarían todos los scripts a lo largo del trabajo para el tratamiento de datos. Tras lo cual, se tendría que situar en carpeta de trabajo correspondiente e iniciar el IPython, explicado anteriormente, de la siguiente manera:

```
(Canopy 64bit) C:\TFG>ipython

Python 2.7.11 | 64-bit | (default, Jun 11 2016, 11:33:47) [MSC v.1500 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.

IPython 4.1.2 -- An enhanced Interactive Python.
?      -> Introduction and overview of IPython's features.
%quickref -> Quick reference.
help    -> Python's own help system.
object? -> Details about 'object', use 'object??' for extra details.

In [1]: a = 5

In [2]: a
Out[2]: 5
```

Como se puede observar, al ejecutar el comando “ipython” se entraría en el entorno de forma rápida y sencilla. Como prueba, se ha definido y declarado una variable de nombre “a” con valor igual a 5, para más tarde con sólo poner el nombre de la variable se obtendría su valor actual. A su vez, como diferencia entre los comandos que entran en la consola y los que salen de ella, se pondrán las palabras “In” para indicar que se va a insertar un comando y “Out” para indicar la salida de un comando, aunque depende del comando que se ejecute se mostrará o no una salida.



3.1.2 Obtención de los países que más visitan las bodegas según similarWeb

Una vez todo listo y aclarado, se pasaría a analizar los primeros datos. Éstos son los países que más han visitado las bodegas según los datos obtenidos de similarWeb. A partir del formato JSON, el campo será “traffic_by_countries_sw”, un ejemplo de este campo, puede ser el siguiente:

```
"traffic_by_countries_sw": [{
  "country": "Spain",
  "percentage": "59.74"
}, {
  "country": "Poland",
  "percentage": "4.72"
}, {
  "country": "Japan",
  "percentage": "3.67"
}, {
  "country": "Canada",
  "percentage": "3.47"
}, {
  "country": "Mexico",
  "percentage": "2.95"
}]
```

En este caso, se puede observar que para esta bodega los países más visitados son España, Polonia, Japón, Canadá y México. Se haría hincapié en obtener este campo, si es que existe en la bodega, y obtener los nombres de todos los países. Una vez obtenidos se almacenarían en una lista de tuplas con el conteo de veces que ha aparecido ese país.

Para realizar lo anterior, se seguirían los siguientes pasos:

Primero, se tendrían que cargar los datos en memoria de un mes, por ejemplo, de febrero:

```
In [1]: import json

In [2]: path = "Febrero.json"

In [3]: with open(path) as data_file:
...:     data = json.load(data_file)
...:

In [4]: data["bodegas"][0]["site"]
Out[4]: u'10dabril.com'
```

En la primera instrucción se importaría la librería para leer el archivo JSON externo localizado en el directorio de trabajo. Luego se indicaría el nombre del archivo en concreto y se le asociaría la variable “path”, tras lo cual se abriría el archivo de nombre el valor de la variable “path”, declarado anteriormente. Más tarde se asociarían los datos del archivo a la variable “data” y finalmente como muestra de cómo se accedería a algún dato según la estructura del archivo se tendría la última sentencia donde se obtiene el campo “site” que corresponde al nombre del sitio de la primera bodega.

Una vez cargados los datos del mes, se procedería a realizar el conteo de los países que más visitan las bodegas según similarWeb. Para ello se introducirían las siguientes sentencias:

```
[1] counts = {}
[2] for i in range(len(data["bodegas"])):
[3]     if "traffic_by_countries_sw" in data["bodegas"][i]:
[4]         for j in range(len(data["bodegas"][i]["traffic_by_countries_sw"])):
[5]             if data["bodegas"][i]["traffic_by_countries_sw"][j]["country"] in counts:
[6]                 counts[data["bodegas"][i]["traffic_by_countries_sw"][j]["country"]] += 1
[7]             else:
[8]                 counts[data["bodegas"][i]["traffic_by_countries_sw"][j]["country"]] = 1
```

En la primera sentencia se crea una lista de tuplas vacía que contendrá el nombre del país y el número de veces que aparece en todas las bodegas a lo largo del mes. Luego, con el bucle se recorrería todas las bodegas para comprobar si existe el campo “traffic_by_countries_sw”. En el caso de que existiese, se procedería a recorrer todos los países que este contuviera, estando el nombre del país alojado en el campo “country”. Finalmente, se comprobaría si el país existe en nuestra lista de tuplas, en caso afirmativo se sumaría en uno la cantidad que hubiera anteriormente indicada, en caso de que el país no estuviera, se añadiría además de que el conteo pasaría a ser uno.

El último paso que faltaría por hacer sería obtener los cinco países más visitados. Para ello se tendría que ordenar la lista de tuplas generada anteriormente según el número de conteo de cada país. Se introducirían pues las siguientes sentencias:

```
[0] import operator
[1] counts = sorted(counts.items(), key=lambda x: x[1], reverse=True)
[2] counts_dict = dict(counts[:5])
[3] siglas = {"United States": "US", "United Kingdom": "UK", "Spain": "ES"}
[4] for k,v in counts_dict.items():
[5]     counts_dict[siglas[k]] = counts_dict.pop(k)
```

Primero se ordenaría según el conteo de cada país, como por defecto esa sentencia ordena ascendentemente, mediante el parámetro “reverse=True” se le indica que realice el ordenamiento descendientemente. En la siguiente sentencia, se pasaría a



crear un diccionario de datos clave-valor, donde la clave sería el nombre del país y el valor su conteo, también se indica que sólo se cree el diccionario con los cinco primeros elementos de nuestra lista de tuplas “counts”. Por último, se declararía otro diccionario clave-valor donde el nombre de cada país sería la clave y el valor sus siglas en dos letras, este diccionario se ha acertado de lo que realmente ocupa por temas de espacio. Se ha hecho así para que en la gráfica final los datos se puedan leer con mayor comodidad. En el bucle de la sentencia 4, se realizaría una sustitución del nombre del país por su correspondiente sigla.

```
In [6]: counts_dict
Out[6]: {'ES': 39, 'FR': 13, 'MX': 15, 'UK': 16, 'US': 25}
```

Al mostrar el contenido del diccionario, se puede ver cómo se han sustituido los nombres de sus países y además el conteo que tiene cada uno. Con estos datos listos, ya se podría pasar a la fase de visualización que se verá en el capítulo siguiente.

Realizar todo este proceso para cada mes puede ser un poco pesado, por ello lo que se ha hecho es realizar una función para obtener los datos de cada mes y otra función para calcular el conteo de los países, ambas funciones estarían en un archivo externo en el directorio de trabajo.

La lógica por tanto sería la siguiente:

```
[1] from data_process import get_data
[2] from data_process import calculate_counts
[3] meses = ["Febrero", "Marzo", "Abril", "Junio", "Julio"]
[4] for i, val in enumerate(meses):
[5]     globals()['data_' + val.lower()] = get_data(val)
[6]     data = globals()['data_' + val.lower()]
[7]     globals()['paises_' + val.lower()] = calculate_counts(data, 1)
```

```
In [33]: paises_febrero
Out[33]: {'ES': 39, 'FR': 13, 'MX': 15, 'UK': 16, 'US': 25}
```

```
In [34]: paises_julio
Out[34]: {'ES': 44, 'MX': 12, 'RU': 8, 'UK': 14, 'US': 26}
```

En las dos primeras instrucciones se importan las funciones “get_data” y “calculate_counts” desde el archivo “data_process.py”. Más tarde se crearía una lista de meses correspondientes a los datos que hay disponibles de similarWeb y se recorrería. Para cada iteración se crearían variables globales del tipo “data_nombremes” que contendrán los datos de cada mes a partir de la función “get_data”. Finalmente, también se crean variables que serán diccionarios clave-valor que contendrán las siglas del país y su conteo, se llamarán de la siguiente manera: “paises_nombremes”.

Todo el proceso anteriormente explicado de cómo se obtendría el conteo estaría en la función “calculate_counts” donde por parámetro se le pasarían los datos del mes. Además, en las dos últimas sentencias se puede ver cómo quedarían representados los datos.

3.1.3 Obtención de los países que más visitan las bodegas según Alexa

En este apartado se hará hincapié en cómo obtener los datos del campo “audience_geography_alex”, cuyo ejemplo de estructura puede ser la siguiente:

```
"audience_geography_alex": [{
  "country": "Mexico",
  "percent": "23.4",
  "rank": "914"
}, {
  "country": "Spain",
  "percent": "17.2",
  "rank": "1569"
}, {
  "country": "Argentina",
  "percent": "10.7",
  "rank": "1006"
}, {
  "country": "Venezuela",
  "percent": "8.8",
  "rank": "494"
}, {
  "country": "Colombia",
  "percent": "7.1",
  "rank": "856"
}]
```

Para obtener los países que más visitan las bodegas según Alexa, el proceso sería prácticamente el mismo, la única diferencia radicaría en que el campo que contiene todos los países y el que habría que recorrer y obtener es “audience_geography_alex”. Se podría, por tanto, reusar la función creada para calcular el conteo de cada país “calculate_counts”, pero añadiéndole un parámetro para que, en función de él, se accediese al campo “traffic_by_countrie_sw” o el que se requiera en este apartado que es “audience_geography_alex”. Por ello, se han agregado las siguientes instrucciones:

```
[1] def calculate_counts(data, campo):
[2]     if campo == 1:
[3]         nom_campo = "traffic_by_countries_sw"
[4]     else:
[5]         nom_campo = "audience_geography_alex"
```

Como se puede observar, la primera sentencia corresponde a la definición de la función además de los parámetros que contendrá. Aquí es donde se ha agregado el parámetro “campo” que será un número y dependiendo de él, el nombre del campo al

que se accederá será uno u otro. En caso de que sea uno se accederá a “traffic_by_countries_sw” en cualquier otro caso será “audience_geography_alexas”.

La lógica definitiva para esta función sería la siguiente:

```
def calculate_counts (data, campo):
    import operator
    if campo == 1:
        nom_campo = "traffic_by_countries_sw"
    else:
        nom_campo = "audience_geography_alexas"
    counts = {}
    for i in range(len(data["bodegas"])):
        if nom_campo in data["bodegas"][i]:
            for j in range(len(data["bodegas"][i][nom_campo])):
                if data["bodegas"][i][nom_campo][j]["country"] in counts:
                    counts[data["bodegas"][i][nom_campo][j]["country"]] += 1
                else:
                    counts[data["bodegas"][i][nom_campo][j]["country"]] = 1

    counts = sorted(counts.items(), key=lambda x: x[1], reverse=True)
    counts_dict = dict(counts[:5])
    siglas = {"United States": "US", "United Kingdom": "UK", "Spain": "ES"}
    for k,v in counts_dict.items():
        counts_dict[siglas[k]] = counts_dict.pop(k)
    return counts_dict
```

Después de todo el proceso, la función devolverá un diccionario que contendrá las siglas de los cinco países más visitados y su correspondiente conteo. El proceso, por tanto, sería:

1. Cargar los datos de todos los meses mediante la función “get_data”
2. Aplicar la función “calculate_counts” con el segundo parámetro igual a un número distinto de uno y el primero siendo los datos del mes que toque almacenados en una variable global definida anteriormente.



Con lo cual la lógica a seguir sería la que siguiente:

```
from data_process import get_data
from data_process import calculate_counts
meses_alexa = ["Febrero", "Mayo", "Junio", "Julio"]
for i, val in enumerate(meses):
    globals()['data_' + val.lower()] = get_data(val)
    data = globals()['data_' + val.lower()]
    globals()['alexa_paises_' + val.lower()] = calculate_counts (data, 2)
```

```
In [28]: alexa_paises_mayo
```

```
Out[28]: {'AR': 3, 'ES': 104, 'MX': 3, 'US': 7, 'VE': 2}
```

```
In [29]: alexa_paises_febrero
```

```
Out[29]: {'ES': 71, 'IN': 2, 'MX': 2, 'US': 8, 'VE': 2}
```

```
In [30]: alexa_paises_junio
```

```
Out[30]: {'CO': 2, 'ES': 93, 'MX': 3, 'US': 3, 'VE': 2}
```

Sólo se han introducido esos meses debido a que en los demás no hay datos del tráfico de países. Finalmente, se pueden observar los valores de conteo de los diferentes países representados en siglas.

3.1.4 Obtención de los sitios que hacen referencia y los sitios destinos de las bodegas según similarWeb

En este apartado se obtendrán aquellos cinco sitios que hacen más referencias a las bodegas, es decir, aquellos sitios desde donde llegan a las bodegas en cuestión a través de enlaces de todo tipo. Y aquellos sitios que los usuarios visitan después de entrar en la bodega. Estos datos estarán en los campos “top_referring_sites_sw” y “top_destination_sites_sw”, cuyo ejemplo de estructura puede ser la siguiente:

```
"top_referring_sites_sw": [{  
    "site": "Sherry.org"  
}]  
  
"top_destination_sites_sw": [{  
    "site": "Pinterest.com"  
}, {  
    "site": "Facebook.com"  
}]
```

Siguiendo con la función “calculate_counts” debido a que ya no sólo se van a obtener los cinco países que más visitan las bodegas sino también los cinco sitios que hacen más referencia a las bodegas y los cinco sitios destinos que se visitan después de las bodegas.

Por consiguiente, se añadiría la siguiente lógica a la función “calculate_counts”:

```
def calculate_counts(data, campo):  
    if campo == 1:  
        nom_campo = "traffic_by_countries_sw"  
        nom_item = "country"  
    elif campo == 2:  
        nom_campo = "audience_geography_alexa"  
        nom_item = "country"  
    elif campo == 3:  
        nom_campo = "top_referring_sites_sw"  
        nom_item = "site"  
    else campo == 4:  
        nom_campo = "top_destination_sites_sw"  
        nom_item = "site"
```

En función del valor del parámetro “campo”, se accederá a unos campos o a otros dependiendo de los indicadores que se quiera obtener. A diferencia de como estaba la función antes, se ha añadido otra variable de nombre “nom_item” que corresponde al campo hijo dentro del padre que es el valor del campo “nom_campo”. Esto quiere decir que cuando se quieran obtener los países que más visitan las bodegas,



se necesitará el campo “traffic_by_countries_sw” y además el “country” en el cual está el nombre de un país. Lo correspondiente se aplicaría para cuando se quiera obtener los sitios.

También se ha agregado otra comprobación del valor del parámetro “campo”, para que cuando sea uno o dos se sustituya el país por sus siglas. Esta sustitución no lo hará cuando se decida obtener por ejemplo los cinco sitios que hacen más referencias a las bodegas ya que no hay unas siglas universales para cada sitio web debido a la gran cantidad que hay de éstos.

Por lo tanto, la lógica de la función quedaría de la siguiente manera:

```
def calculate_counts(data, campo):
    import operator
    if campo == 1:
        nom_campo = "traffic_by_countries_sw"
        nom_item = "country"
    elif campo == 2:
        nom_campo = "audience_geography_alexa"
        nom_item = "country"
    elif campo == 3:
        nom_campo = "top_referring_sites_sw"
        nom_item = "site"
    else:
        nom_campo = "top_destination_sites_sw"
        nom_item = "site"

    counts = {}
    for i in range(len(data["bodegas"])):
        if nom_campo in data["bodegas"][i]:
            for j in range(len(data["bodegas"][i][nom_campo])):
                if data["bodegas"][i][nom_campo][j][nom_item] in counts:
                    counts[data["bodegas"][i][nom_campo][j][nom_item]] += 1
                else:
                    counts[data["bodegas"][i][nom_campo][j][nom_item]] = 1

    counts = sorted(counts.items(), key=lambda x: x[1], reverse=True)
    counts_dict = dict(counts[:5])
    if campo == 1 or campo == 2:
        siglas = {"United States": "US", "United Kingdom": "UK", "Spain": "ES"}
        for k,v in counts_dict.items():
            counts_dict[siglas[k]] = counts_dict.pop(k)

    return counts_dict
```

Por último, se calcularía el conteo de los sitios que hacen referencia y los destinos de la siguiente manera:

```
from data_process import get_data
from data_process import calculate_counts
meses = ["Febrero", "Marzo", "Abril", "Junio", "Julio"]
for i, val in enumerate(meses):
    globals()['data_' + val.lower()] = get_data(val)
    data = globals()['data_' + val.lower()]
    globals()['refer_' + val.lower()] = calculate_counts(data, 3)
    globals()['dest_' + val.lower()] = calculate_counts(data, 4)
```

In [19]: refer_febrero

Out[19]:

```
{u'Avinicolacatalana.cat': 16,
u'Denominacion-origen-almansa.com': 8,
u'Lamanchawines.com': 12,
u'Orujodegalicia.org': 17,
u'Ribeirasacra.org': 9}
```

In [20]: dest_febrero

Out[20]:

```
{u'Alimentasonrisas.es': 2,
u'Facebook.com': 5,
u'Grupoartevino.com': 1,
u'Paypal.com': 3,
u'Wineinmoderation.eu': 3}
```

Se realizarían los mismos pasos que en las dos obtenciones anteriores. Primero se cargarían los datos de todos los meses, en caso de tenerlos cargados anteriormente no haría falta volver a hacerlo. Y luego se llamaría a la función “calculate_counts” pasándole como parámetro tres significa que se requieren los sitios que hacen referencia y con el parámetro cuatro los sitios destinos. Se almacenarían los diccionarios obtenidos en variables globales y con ello ya se tendrían listos los datos para visualizarse en el capítulo siguiente.

3.1.5 Obtención del porcentaje de uso de las redes sociales más utilizadas según similarWeb

En la siguiente obtención, se analizarán qué redes sociales utilizan más las bodegas a lo largo de todos los meses. Esta obtención es diferente a las anteriores ya que en vez de mostrar los datos por meses, se mostrarán de forma global. El campo con el que se trabajara en esta obtención es “social_sw” y un ejemplo de la estructura de éste puede ser el siguiente:

```
"social_sw": [{
  "social_network": "Facebook",
  "percentage": "85.85"
}, {
  "social_network": "Youtube",
  "percentage": "7.09"
}, {
  "social_network": "Linkedin",
  "percentage": "7.06"
}]
```

El proceso sería parecido a los de las demás obtenciones, pero con la diferencia de que al final se creará otra función que haga más cálculos sobre los datos tratados. Con esto dicho, lo primero que se tendría que efectuar es ajustar la función “calculate_counts” para que haga las mismas operaciones que hasta ahora pero también con el campo “social_sw”, con lo cual los cambios serían los siguientes:

```
if campo == 1:
    nom_campo = "traffic_by_countries_sw"
    nom_item = "country"
elif campo == 2:
    nom_campo = "audience_geography_alexa"
    nom_item = "country"
elif campo == 3:
    nom_campo = "top_referring_sites_sw"
    nom_item = "site"
elif campo == 4:
    nom_campo = "top_destination_sites_sw"
    nom_item = "site"
else:
    nom_campo = "social_sw"
    nom_item = "social_network"
```

Como se recordará, la variable “campo” es el segundo parámetro que se le pasa a esta función y a partir de ese parámetro se obtendrán unos datos u otros. Para la

obtención en la que se aplica con indicar un valor mayor o igual a cinco bastaría para que se obtengan los datos que se necesitan.

También recordar, que esta función contaba las veces que se repetía cierto dato en todo un mes, en este caso actual se obtendrán las veces que se ha utilizado cierta red social por mes. Sin embargo, se necesitan saber el porcentaje global de las redes sociales más utilizadas. Para ello se aplicará la lógica siguiente, mediante la creación de una nueva función:

```
def per_social(febrero, marzo, abril, junio):
    mes_total = {}
    for k, v in febrero.items():
        if k in mes_total:
            mes_total[k] += v
        else:
            mes_total[k] = v

    for k, v in marzo.items():
        if k in mes_total:
            mes_total[k] += v
        else:
            mes_total[k] = v

    for k, v in abril.items():
        if k in mes_total:
            mes_total[k] += v
        else:
            mes_total[k] = v

    for k, v in junio.items():
        if k in mes_total:
            mes_total[k] += v
        else:
            mes_total[k] = v

    mes_total = dict(mes_total)
    total = 0

    for k, v in mes_total.items():
        total += v

    for k, v in mes_total.items():
        mes_total[k] = round((v / float(total)) * 100)

    return mes_total
```



Debido a que no hay datos respecto a este campo en otros meses, se hará hincapié en los meses de febrero, marzo, abril y junio. Esta función recibe como parámetros los datos con el conteo de las redes sociales más utilizadas en ese mes, lo que se ha hecho es juntar todos estos datos para luego obtener el porcentaje. Los pasos describiendo la lógica han sido los siguientes:

1. Crear una lista de tuplas vacía “mes_total” que contendrá las redes sociales más utilizadas además del conteo total de todos los meses.
2. Recorrer cada uno de los cuatro meses e ir sumando el valor de cada red social en la lista creada anteriormente.
3. Convertir la lista de tuplas en diccionario para mejor manejo de los datos.
4. Crear una variable “total” que será la que almacene la suma de todos los conteos de todas las redes sociales del diccionario “mes_total” para así poder obtener el correspondiente porcentaje del total. Se recorrería pues el diccionario “mes_total” para obtener el valor de la variable “total”
5. Finalmente, se volvería a recorrer el diccionario “mes_total” pero esta vez se sustituiría el valor actual (conteo) de cada red social por su correspondiente porcentaje, este cálculo se haría de la siguiente manera:

```
round((v / float(total)) * 100)
```

Primero se dividiría el conteo actual de la red social entre el total hallando anteriormente, para luego pasarlo a formato decimal multiplicándolo por cien y finalmente redondearlo para que no haya ningún decimal.

Por último, la ejecución de estas funciones sería la siguiente y en este orden:

```
from data_process import *
meses = ['Febrero', 'Marzo', 'Abril', 'Junio', 'Julio']
for i, val in enumerate(meses):
    globals()['data_' + val.lower()] = get_data(val)
    data = globals()['data_' + val.lower()]
    globals()['social_' + val.lower()] = calculate_counts(data, 5)

In [50]: porce = per_social(social_febrero, social_marzo, social_abril, social_junio)

In [51]: porce
Out[51]:
{u'Facebook': 68.0,
 u'Linkedin': 9.0,
 u'Pinterest': 2.0,
 u'Twitter': 11.0,
 u'Youtube': 10.0}
```

Como se puede observar, el proceso a seguir es el mismo con la diferencia de que se aplicaría la función “per_social” creada para esta obtención para obtener los porcentajes correspondientes de todas las redes sociales. Finalmente se puede ver como quedarían los porcentajes listos para ser visualizados gráficamente, para comprobar que son correctos se puede verificar como la suma de todos da como resultado cien.



3.1.6 Obtención de datos comparativos entre Alexa y similarWeb

En este apartado se describirá el proceso para la obtención de datos comparativos entre Alexa y similarWeb. Como se ha visto en el capítulo de extracción de datos, se indicó que indicadores se extraían de Alexa y de similarWeb, ambos tienen en común que obtienen las mismas métricas, pero a su manera. De esta manera, se explicará cómo se han tratado estos datos para luego representarlos gráficamente.

Los indicadores que tienen en común Alexa y similarWeb son los siguientes:

- Bounce rate
- Time on site
- Search visits
- Page views

Tanto los indicadores “bounce rate” y “search visits” se medirán en porcentajes, el “time on site” se medirá en minutos y el “page views” en número de páginas que se han visto. Por ello, los dos primeros se representarán en la misma gráfica y los dos últimos en una gráfica cada uno, debido a que la unidad de medida es distinta.

La lógica que se ha implementado para la siguiente obtención se ha dividido en dos partes. En la primera función se obtendrán todos los indicadores antes mencionados divididos por cada mes y en la segunda función se unirán todos los datos obteniendo los datos globales de todos los meses que más tarde se representarán gráficamente.

Por lo tanto, la primera función donde se obtendrían los indicadores para cada mes sería la que sigue:

```
def averages(data):  
  
    # Primer paso  
    search_visits_alex = []  
    search_visits_sw = []  
    bounce_rate_alex = []  
    bounce_rate_sw = []  
    time_on_site_alex = []  
    time_on_site_sw = []  
    page_views_alex = []  
    page_views_sw = []  
  
    medias = {}
```



Segundo paso

```
for i in range(len(data["bodegas"])):
    if data["bodegas"][i]["search_visits_alexa"] > 0.0:
        search_visits_alexa.append(data["bodegas"][i]["search_visits_alexa"])

    if data["bodegas"][i]["tf_search_sw"] > 0.0:
        search_visits_sw.append(data["bodegas"][i]["tf_search_sw"])
    if data["bodegas"][i]["bounce_rate_alexa"] > 0.0:
        bounce_rate_alexa.append(data["bodegas"][i]["bounce_rate_alexa"])
    if data["bodegas"][i]["bounce_rate_sw"] > 0.0:
        bounce_rate_sw.append(data["bodegas"][i]["bounce_rate_sw"])
    if data["bodegas"][i]["daily_time_on_site_alexa"] > 0.0:
        time_on_site_alexa.append(data["bodegas"][i]["daily_time_on_site_alexa"])
    if data["bodegas"][i]["time_on_site_sw"] > 0.0:
        time_on_site_sw.append(data["bodegas"][i]["time_on_site_sw"])
    if data["bodegas"][i]["daily_pageviews_per_visitor_alexa"] > 0.0:
        page_views_alexa.append(data["bodegas"][i]["daily_pageviews_per_visitor_alexa"])
    if data["bodegas"][i]["page_views_sw"] > 0.0:
        page_views_sw.append(data["bodegas"][i]["page_views_sw"])
```

Tercer paso

```
medias["search_visits_alexa"] = round(sum(search_visits_alexa) /
float(len(search_visits_alexa)), 2)
if not search_visits_sw:
    medias["search_visits_sw"] = 0.0
else:
    medias["search_visits_sw"] = round(sum(search_visits_sw) /
float(len(search_visits_sw)), 2)

medias["bounce_rate_alexa"] = round(sum(bounce_rate_alexa) /
float(len(bounce_rate_alexa)), 2)
if not bounce_rate_sw:
    medias["bounce_rate_sw"] = 0.0
else:
    medias["bounce_rate_sw"] = round(sum(bounce_rate_sw) /
float(len(bounce_rate_sw)), 2)

medias["time_on_site_alexa"] = round(sum(time_on_site_alexa) /
float(len(time_on_site_alexa)), 2)
if not time_on_site_sw:
    medias["time_on_site_sw"] = 0.0
else:
    medias["time_on_site_sw"] = round(sum(time_on_site_sw) /
float(len(time_on_site_sw)), 2)

medias["page_views_alexa"] = round(sum(page_views_alexa) /
```



```
float(len(page_views_alexa), 2)
    if not page_views_sw:
        medias["page_views_sw"] = 0.0
    else:
        medias["page_views_sw"] = round(sum(page_views_sw) /
float(len(page_views_sw)), 2)
```

En esta función se le pasan los datos globales de cada mes por el único parámetro que tiene. A continuación, como primer paso se declaran las listas de los diferentes indicadores tanto para Alexa como para similarWeb y un diccionario que contendrá la media de cada indicador por mes.

El segundo paso consiste en recorrer todos los datos del mes en busca de los indicadores, como estos indicadores son numéricos, se realiza una comprobación de sólo obtener aquellos que sean mayores que cero debido a que en estos datos no habrá números negativos. Tras la comprobación de que sean mayores que cero, se insertan en su lista correspondiente.

El tercer y último paso consistiría en calcular la media de todos los indicadores de todas las bodegas en el mes que se esté tratando. Debido a problemas de extracción no hay datos de los correspondientes indicadores en el mes de mayo de similarWeb, por ello si al obtener los datos del mes las correspondientes listas de indicadores pertenecientes a similarWeb están vacías la media pasará a ser cero.

Para calcular la media, primero se suman todos los valores de la lista para dividirlos entre el número de elementos de la lista. Finalmente, una vez obtenida la media se redondea el valor a dos decimales para que no hayan números período ni tampoco demasiado largos en cuanto a decimales se refiere.

La siguiente parte de este apartado se basaría en volver a realizar la media, pero esta vez sería la media de todos los meses antes calculada. Por consiguiente, la lógica que se seguiría sería la siguiente:

```
def total_average(*args):
    from data_process import averages

    # Primer paso

    media_total = {'bounce_rate_alexa': 0.0, 'bounce_rate_sw': 0.0,
'page_views_alexa': 0.0, 'page_views_sw': 0.0, 'search_visits_alexa': 0.0,
'search_visits_sw': 0.0, 'time_on_site_alexa': 0.0, 'time_on_site_sw': 0.0}

    # Segundo paso

    for arg in args:
        media_mes = averages(arg)
        media_total["search_visits_alexa"] +=
```

```

media_mes["search_visits_alexa"]
    if media_mes["search_visits_sw"] > 0.0:
        media_total["search_visits_sw"] +=
media_mes["search_visits_sw"]
        media_total["bounce_rate_alexa"] += media_mes["bounce_rate_alexa"]
        if media_mes["bounce_rate_sw"] > 0.0:
            media_total["bounce_rate_sw"] +=
media_mes["bounce_rate_sw"]
            media_total["time_on_site_alexa"] +=
media_mes["time_on_site_alexa"]
            if media_mes["time_on_site_sw"] > 0.0:
                media_total["time_on_site_sw"] +=
media_mes["time_on_site_sw"]
                media_total["page_views_alexa"] += media_mes["page_views_alexa"]
                if media_mes["bounce_rate_sw"] > 0.0:
                    media_total["page_views_sw"] +=
media_mes["page_views_sw"]

    # Tercer paso

    media_total["search_visits_alexa"] = round(media_total["search_visits_alexa"]
/ len(args), 2)
    media_total["search_visits_sw"] = round(media_total["search_visits_sw"] /
(len(args) - 1), 2)
    media_total["bounce_rate_alexa"] = round(media_total["bounce_rate_alexa"] /
len(args), 2)
    media_total["bounce_rate_sw"] = round(media_total["bounce_rate_sw"] /
(len(args) - 1), 2)
    media_total["time_on_site_alexa"] = round(media_total["time_on_site_alexa"]
/ len(args), 2)
    media_total["time_on_site_sw"] = round(media_total["time_on_site_sw"] /
(len(args) - 1), 2)
    media_total["page_views_alexa"] = round(media_total["page_views_alexa"] /
len(args), 2)
    media_total["page_views_sw"] = round(media_total["page_views_sw"] /
(len(args) - 1), 2)

    return media_total

```

En los parámetros de esta función se pasan los datos globales de todos los meses para de esta manera hacer el cálculo de la media global. Esta lógica también la se dividiría en tres pasos. En el primero de ellos se crea un diccionario que contendrá la media global de cada indicador, pero por ahora sólo se han inicializado a cero para que los datos existan.



En el segundo paso, se pasaría a recorrer cada uno de los parámetros que serían los datos globales de cada mes. En cada iteración se obtendría la media del correspondiente mes, y el valor de la media de cada indicador por mes se insertaría en el diccionario de datos que se ha creado en el anterior paso. Para obtener la media de cada país se hace uso de la función que se ha explicado anteriormente, es decir, de la “averages(data)”. A su vez, durante la inserción de la media de cada indicador por mes, se comprueba que ésta sea mayor que cero debido a como se ha comentado antes, similarWeb no dispone del mes de mayo.

Por último, en el tercer paso se procede a calcular la media global a partir de todos los meses disponibles. La manera de calcularlo es la misma que en la función antes explicada de este apartado. El proceso sería realizar la suma de todos los valores de las medias, para luego dividirlo entre el número de elementos de la lista y finalmente redondear el valor obtenido a dos decimales.

Por último, para llamar correctamente a esta última función se debería introducir lo siguiente en el intérprete de comandos:

```
media_total = total_average(data_febrero, data_marzo, data_abril, data_mayo, data_junio, data_julio)
```

Donde, las variables “data_NOMBREmes” serían los datos cargados en memoria de cada mes, realizados en los apartados anteriores. El resultado de esta función se le asocia a la variable “media_total” que contendrá los datos listos para ser representados gráficamente en el siguiente capítulo en el correspondiente apartado.

3.1.7 Obtención de top cinco sitios a partir del trust flow, citation flow, backlinks externos y dominios de referencia según majesticSEO

En este apartado, se obtendrán los cinco mejores sitios para cada uno de los siguientes indicadores: trust flow, citation flow, el número de backlinks externos y el número de dominios de referencia. Es decir, para cada uno de ellos se obtendrá una tabla por mes donde se indicará el sitio y su correspondiente valor dependiendo del indicador. Con esto, se pretende ver el impacto que tiene cada uno de estos indicadores sobre los sitios web. Comprobar a su vez, que si teniendo un gran trust flow implica tener un gran número de backlinks externos y viceversa.

Para obtener estos datos, se ha creado una función donde a partir de los datos del mes se ordena según el campo correspondiente y se almacena todo en un diccionario de listas de tuplas. Por otra parte, los meses para los que se pueden obtener los datos son marzo, abril, mayo, junio y julio, no se tiene febrero debido a que el algoritmo para extraer datos de majesticSEO no estaba en listo en ese mes.

La lógica de la función, por consiguiente, es la que sigue:

```
def top_majestic(data):
    import pandas as pd

    tops = {}

    # Primer paso

    data_frame = pd.DataFrame(data["bodegas"])
    ordenado_trust = data_frame.sort_values(by=["trust_flow"],
ascending=False)
    ordenado_citation = data_frame.sort_values(by=["citation_flow"],
ascending=False)
    ordenado_backlinks = data_frame.sort_values(by=["backlinks_externos"],
ascending=False)
    ordenado_dominios = data_frame.sort_values(by=["dominios_referencia"],
ascending=False)

    # Segundo paso

    tops["trust_flow"] = zip(ordenado_trust["site"].iloc[:5],
ordenado_trust["trust_flow"].iloc[:5])
    tops["citation_flow"] = zip(ordenado_citation["site"].iloc[:5],
ordenado_citation["citation_flow"].iloc[:5])
    tops["backlinks_externos"] = zip(ordenado_backlinks["site"].iloc[:5],
ordenado_backlinks["backlinks_externos"].iloc[:5])
    tops["dominios_referencia"] = zip(ordenado_dominios["site"].iloc[:5],
ordenado_dominios["dominios_referencia"].iloc[:5])
```



```
return tops
```

Esta función recibiría como parámetro el mes correspondiente desde marzo hasta julio. Antes que nada, se importaría una librería que se ha explicado en el apartado de puesta a punto y que servirá para crear un fragmento de datos, además se crea un diccionario vacío que es el que contendrá todos los datos ya ordenados.

En el primer paso, se crea el fragmento de datos a partir de los datos globales del mes que se haya pasado como parámetro. A continuación, se realizan cuatro ordenaciones y se guardan en cuatro variables distintas, cada una de las cuales contendrá todos los datos ordenados de mayor a menor según el campo que se haya indicado.

Por último en el tercer paso, se procede a insertar los cinco primeros sitios con los valores correspondientes de los distintos marcadores en el diccionario que se ha creado al principio de la función. El resultado de esta función para el mes de marzo, es el siguiente:

```
In [41]: top_marzo = top_majestic(data_marzo)

In [42]: top_marzo
Out[42]:
{'backlinks_externos': [(u'tobelos.com', 2221601),
 (u'vinosjuliana.com', 1470450),
 (u'principedeviana.com', 1262348),
 (u'alfredomaestro.com', 255091),
 (u'valdemar.es', 224362)],
 'citation_flow': [(u'tobelos.com', 64),
 (u'vidyespiga.es', 50),
 (u'habarcelo.es', 44),
 (u'grupomatarromera.com', 43),
 (u'ontanon.es', 42)],
 'dominios_referencia': [(u'vinosjuliana.com', 8571),
 (u'tobelos.com', 5595),
 (u'alvareznolting.com', 4195),
 (u'bodegasosca.com', 4122),
 (u'santacatalina.es', 3686)],
 'trust_flow': [(u'torres.es', 52),
 (u'codorniu.com', 50),
 (u'freixenet.es', 50),
 (u'lopezdeheredia.com', 49),
 (u'marquesdemurrieta.com', 48)]}
```

Una vez con estos datos, ya es posible representarlos en una tabla como se verá en la sección de visualización de datos, en el apartado correspondiente.



3.2 Visualización de datos

En este capítulo se haría hincapié en la visualización de datos, con esto se refiere a mostrar los datos gráficamente de manera que sea más fácil y rápido entender lo que significan y sacar conclusiones al respecto. Para la visualización de estos datos se seguirá utilizando el entorno Enthought Canopy, el cual provee de la librería *matplolib* que se utiliza para dibujar todo tipo de gráficas.

`matplotlib` ^[23] es la biblioteca de Python más popular para la generación de gráficas y otras visualizaciones de datos en dos dimensiones. Fue concebida inicialmente por John D. Hunter, pero en la actualidad es mantenida por un gran equipo de desarrolladores. Se integra a la perfección con la consola de comandos mejorada de Python, IPython (como se verá más adelante), lo que proporciona un entorno interactivo para el trazado y exploración de los datos. Las figuras generadas también son interactivas ya que existe la posibilidad de hacer zoom en una sección de la gráfica y desplazarse por ella usando la barra de herramientas en la ventana gráfica.

²³ matplotlib. <http://matplotlib.org/>



Como ejemplos de las distintas gráficas que se pueden generar gracias a esta librería, se tienen las siguientes:

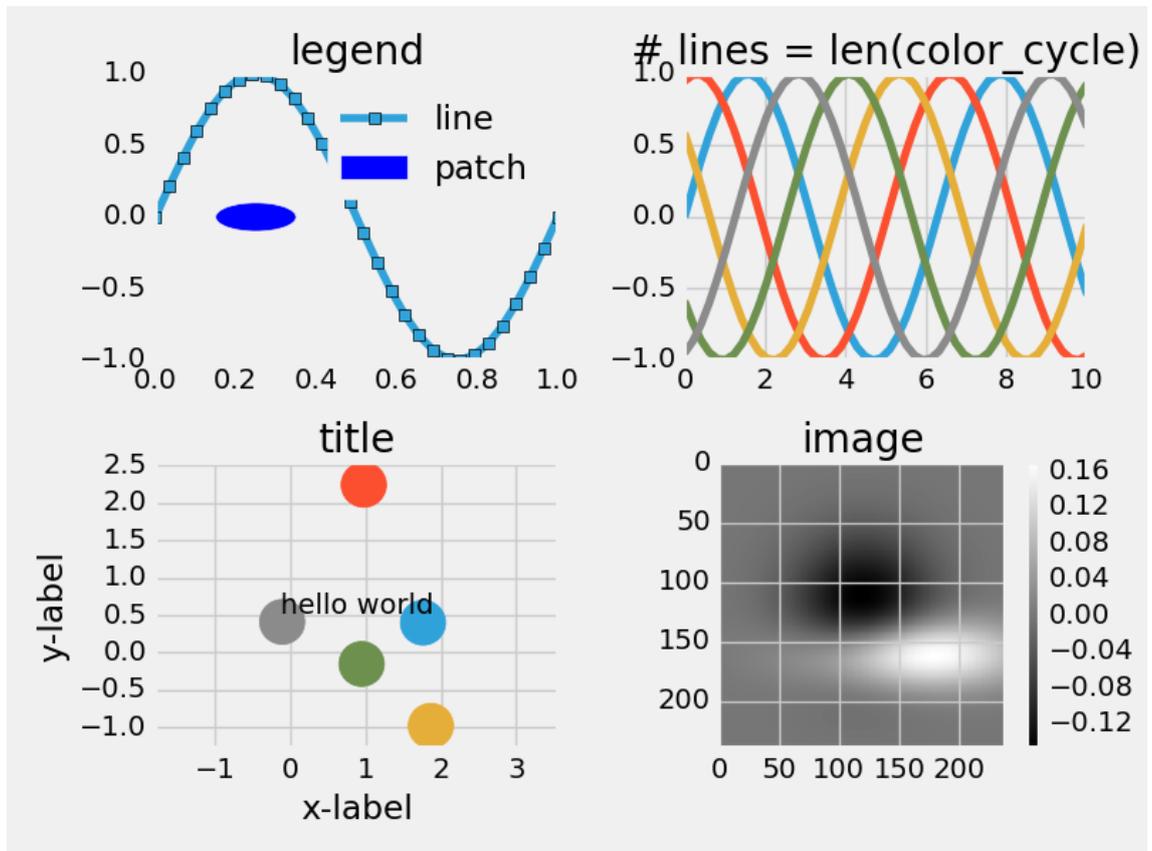


Figura 12. Ejemplos de gráficas que se pueden generar con la biblioteca matplotlib. Fuente: [tonysyu.github.io](https://github.com/tonysyu)

Más ejemplos de gráficos que se pueden gracias a la librería matplotlib:

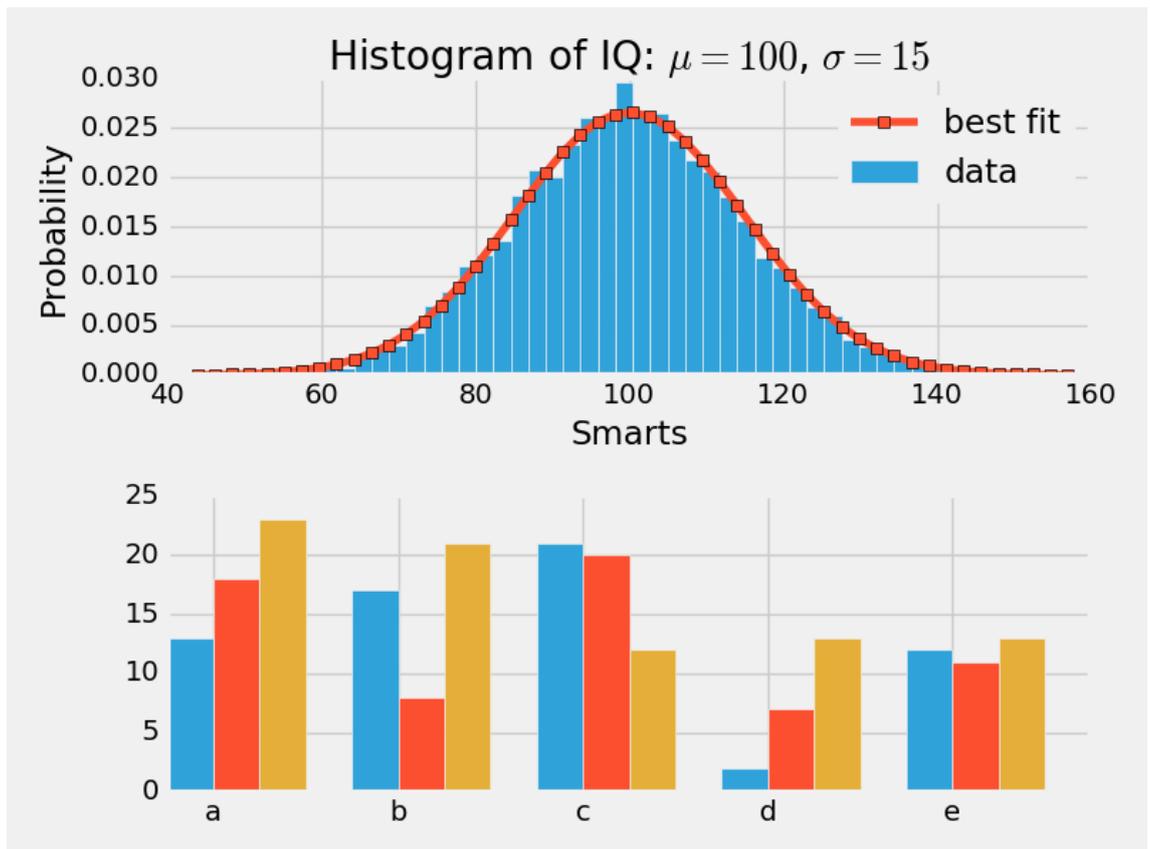
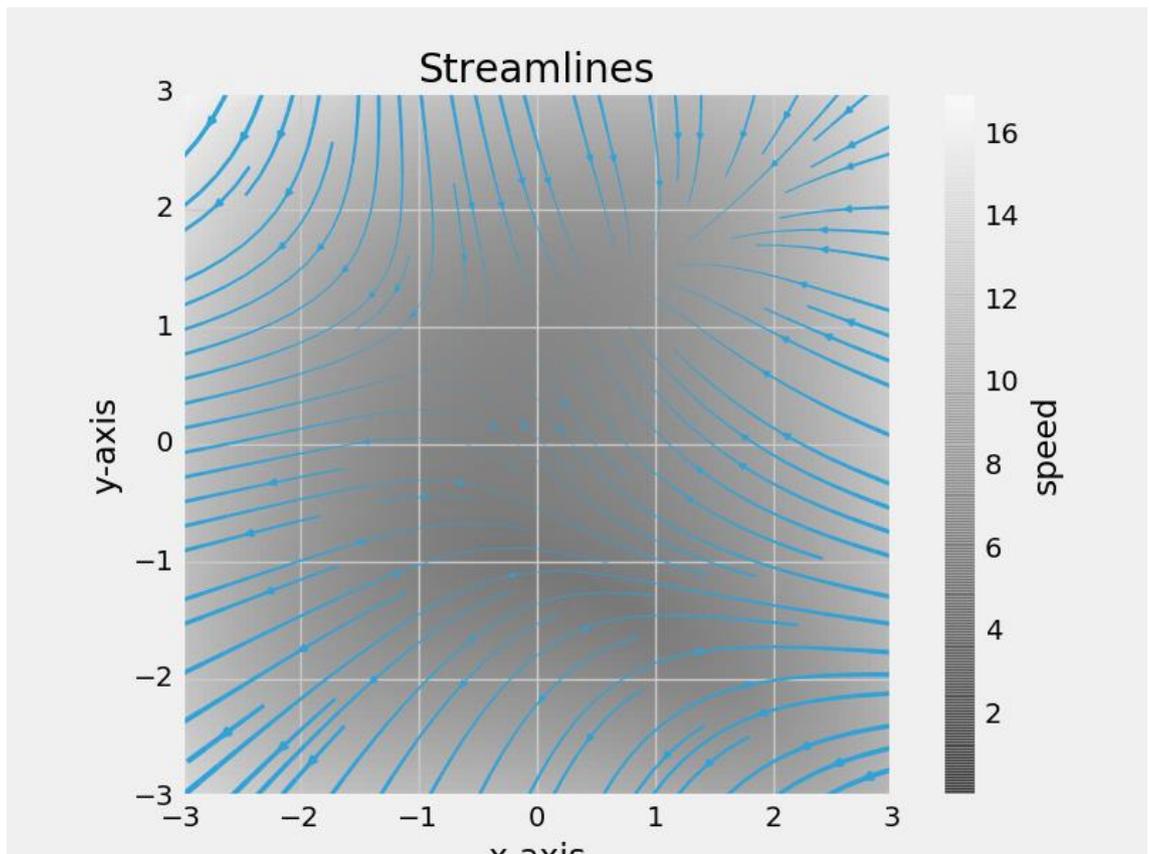


Figura 13. Más ejemplos de gráficas generadas con matplotlib. Fuente: tonysyu.github.io

Otro ejemplo de gráfica que es posible generar con esta biblioteca:



*Figura 14. Otro ejemplo de gráfico generado con la biblioteca matplotlib.
Fuente [tonysyu.github.io](https://github.com/tonysyu)*

A la hora de generar un gráfico hay que indicar ciertos aspectos como son los datos que se van a visualizar, los textos de los mismos, los diferentes colores que se van a usar, algún título, alguna etiqueta, etc. Para cada una de las gráficas se ha creado su propia función y todas estas funciones como con el análisis de datos están en un archivo externo, en este caso “plot_data.py” que contendrá las todas las funciones para la generación de gráficos. Un ejemplo de la ventana que aparecerá al generar cualquier gráfico, podría ser la siguiente:

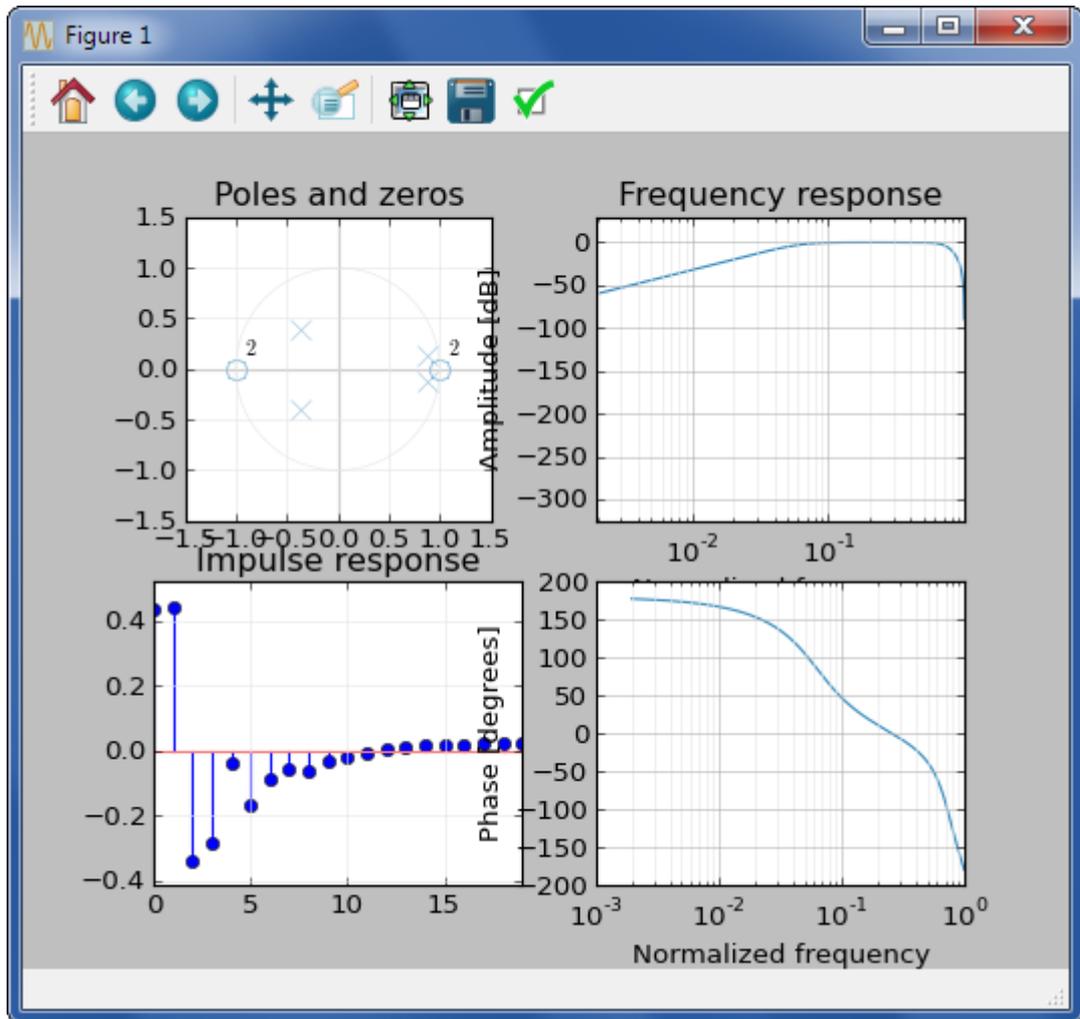


Figura 15. Ventana gráfica que aparece al generar un gráfico. Fuente: stackoverflow.com

3.2.1 Generación gráfica de los países que más visitan las bodegas según similarWeb

Como ha ocurrido en el tratamiento de datos donde bajo una misma función se han podido obtener los datos requeridos, para el caso de las gráficas es parecido ya que el diseño de las mismas es igual, pero con ligeras variaciones. Las diferencias que hay son en que algunos datos sólo están disponibles en determinados meses y por tanto los textos tiene que ser acordes a éstos.

Para esta primera gráfica el objetivo es generar para cada país una línea evolutiva que representarían los países que más visitan las bodegas según similarWeb. Además, cada línea será de un color diferente para distinguirlas más claramente, además de que se mostrará la correspondiente leyenda. A su vez los meses que se mostrarán serán febrero, marzo, abril, junio y julio.

La lógica para la generación de esta gráfica es la que sigue:

```
def plot_general(*args):
    import matplotlib.pyplot as plt
    import numpy as np
    from collections import defaultdict

    plt.style.use('fivethirtyeight')

    fig = plt.figure()
    ax = fig.add_subplot(111)

    ind = np.array(range(len(args)))

    colors = ["#b594af", "#e1d146", "#00a0b0", "#f64b4a", "#8ab44b"]

    data = defaultdict(list)
```

Primero, se comenzaría definiendo la función en la que se indican los parámetros opcionales estos contendrán los datos que habrá que mostrar en el gráfico, debido a que los datos de los gráficos están en un número de meses diferente. Para la gráfica que se pretenden generar en este apartado los meses son febrero, marzo, abril, junio y julio

Lo siguiente y muy importante es importar las librerías necesarias para el muestro de los datos, además de librerías complementarias donde se almacenarán los datos a representar. La primera para generar la gráfica, la segunda porque se necesitaría crear un array donde se posicionarán los distintos grupos de barras según los meses y gracias a esta librería se puede hacer fácilmente y la tercera para crear un diccionario de listas.

Se proseguiría con la siguiente sentencia y en esta se le dará un estilo “fivethirtyeight” para que la figura generada no sea sosa, además de que la representación sea más agradable visualmente. Más tarde, se crearía una figura para el gráfico y en la siguiente una cuadrícula de 1x1 y primera sub-traza.

Por último, se crea un array que contendrá las posiciones de los grupos en el eje de las X, una lista de colores en formato hexadecimal para cada uno de los países y un diccionario de listas vacío, un ejemplo del cual sería el siguiente:

```
[(u'United States', [25, 27, 26, 24, 26]), (u'United Kingdom', [16, 18, 16, None, 14]),  
(u'Mexico', [15, 15, 13, 14, 12]), (u'Spain', [39, 43, 46, 45, 44]), (u'France', [13, 14, 17,  
None])]
```

Después, se continuaría con la siguiente lógica:

```
k = 0  
  
for arg in args:  
    k += 1  
    for i in range(len(arg)):  
        if len(data) == 5 and arg.keys()[i] in data:  
            if (k - 1) > len(data[arg.keys()[i]]):  
                data[arg.keys()[i]].append(None)  
                data[arg.keys()[i]].append(arg.values()[i])  
            if len(data) < 5:  
                data[arg.keys()[i]].append(arg.values()[i])  
  
    for j in range(len(data)):  
        if len(data.values()[j]) < (j + 1):  
            data[data.keys()[j]].append(None)  
            ax.plot(data.values()[j], color=colors[j], linestyle='solid', marker='o')
```

Esta lógica es la más importante de toda la función. En ella se leerán los datos que vienen como parámetro y se introducen en un diccionario de listas de nombre “data”. Por su parte, se leerán solamente los cinco primeros países. A continuación se recorre el diccionario de listas y mediante la sentencia “ax.plot(...)” se dibujan las líneas, se les asocia un color, un estilo y el tipo de marcador.



Finalmente, las últimas sentencias para que se genere el gráfico, serían las que siguen:

```
ax.margins(0.4)

ax.legend([data.keys()[0], data.keys()[1], data.keys()[2], data.keys()[3],
data.keys()[4]], loc='upper right')

if len(args) == 4:
    ax.set_xticklabels(('Febrero', 'Mayo', 'Junio', 'Julio'))
else:
    ax.set_xticklabels(('Febrero', 'Marzo', 'Abril', 'Junio', 'Julio'))
ax.set_xticks(ind)

plt.show()
```

En estas últimas instrucciones, primero se indica un margen para que las líneas que se van a representar no estén pegadas a los bordes y con ello se dificulte su lectura. Seguidamente, se crea la leyenda del gráfico y se indica la ubicación donde deberá aparecer, en este caso en la parte superior derecha.

Más tarde, se comprueba la cantidad de parámetros que le llegan a la función. Si son cuatro quiere decir que se van a dibujar los países más visitados según Alexa y con ello se indican las etiquetas correspondientes que se mostrarán en el eje de las X. Por el contrario, si son cinco quiere decir que se van a representar datos de similarWeb, ya sean de países que más visitan las bodegas o los sitios de referencia o destinos de las bodegas que se verán en los próximos apartados.

Una vez indicado todo, la posición de cada texto, los valores de la tabla, el diseño, etc., se pasaría a mostrar la ventana del gráfico desde donde se podrá hacer zoom, guardar el gráfico en distintos formatos de imagen, entre otras opciones. La última sentencia pues, es la que se encargaría de mostrar esta ventana.

Para que se pueda lanzar esta ventana gráfica, se tendrían que haber realizado los pasos correspondientes en el capítulo de análisis de datos y llamar a la función que se acaba de crear importando todas las funciones almacenadas en el archivo “plot_data.py” del directorio de trabajo, la lógica por tanto sería la siguiente:

```
from plot_data import *
plot_general(paises_febrero, paises_marzo, paises_abril, paises_junio, paises_julio)
```

Después de realizar todo el proceso de extracción, tratamiento y la preparación para la visualización, esta sería la gráfica que representaría los países que más visitan las bodegas según similarWeb:

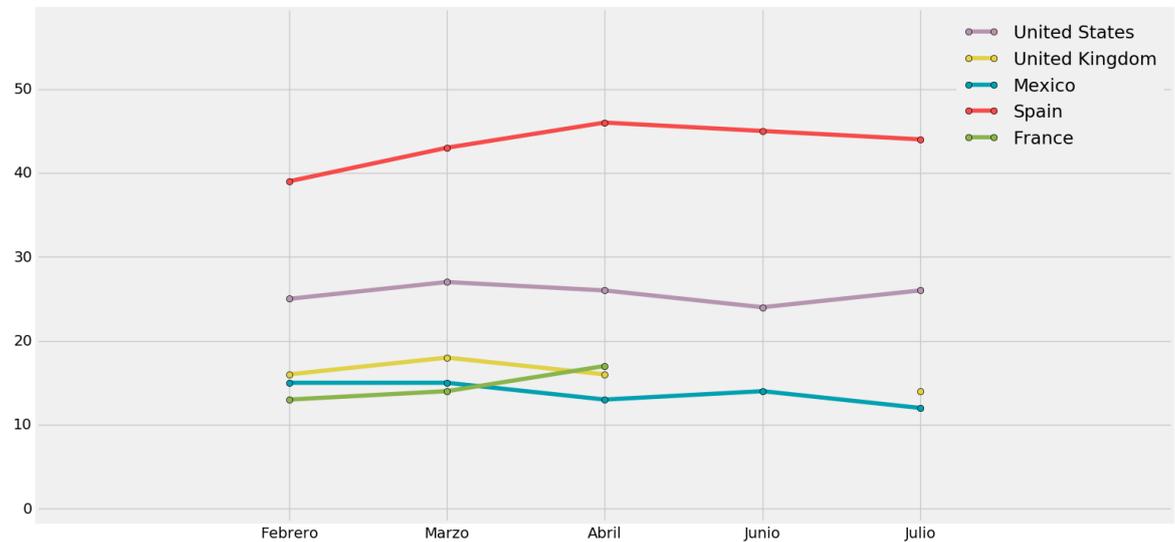


Figura 16. Gráfica con los países que más visitan las bodegas según similarWeb

En cuanto a los datos que muestra la figura en sí, se obtendrán las debidas conclusiones en el siguiente capítulo, en este sólo se ha explicado el proceso de generación de las distintas gráficas.

3.2.2 Generación gráfica de los países que más visitan las bodegas según Alexa

En este apartado se procederá a generar la gráfica de los países que más visitan las bodegas según Alexa, para ello tal y como se comentó en el apartado anterior la lógica sería prácticamente la misma con la diferencia de que los meses a visualizar serán distintos. Esto se debe a que como se ha comentado anteriormente, hubo distintos problemas a la hora de extraer los datos.

La función que generará la figura es la misma que para la gráfica anterior, sólo que se tendría que realizar la llamada a esta función de manera apropiada y pasarle los datos que corresponden a esta gráfica. En el capítulo anterior se guardaban los datos correspondientes a esta gráfica en las variables “alexapaíses_NOMBREMes”.

Por consiguiente, la llamada a la función anterior debería realizarse de la siguiente manera en el intérprete de comandos:

```
from plot_data import *  
plot_general(alexapaíses_febrero, alexapaíses_mayo, alexapaíses_junio,  
alexapaíses_julio)
```

Donde primero se importa la función para que el intérprete la reconozca y finalmente, se le indica los datos de los meses que se van a representar.

Tras ejecutar lo anterior, la gráfica que se generaría sería la siguiente:

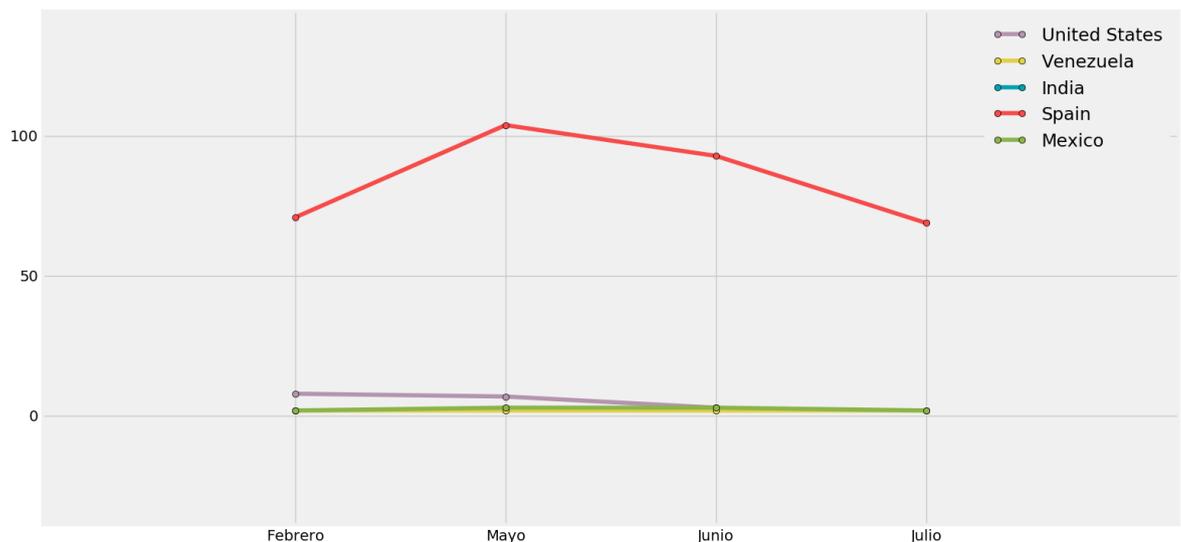


Figura 17. Gráfica con los países que más visitan las bodegas según Alexa

3.2.3 Generación gráfica de los sitios que hacen referencia y los sitios destinos según similarWeb

En este apartado se procederá a generar una gráfica para los sitios que hacen más referencia a las bodegas y otra gráfica para los sitios destinos de las bodegas. El diseño de estas dos gráficas es el mismo con la diferencia de que en la leyenda aparecerán los sitios que más referencia realizan a las bodegas o los sitios destinos, dependiendo de la gráfica.

Para la generación de estas dos gráficas se utilizará la misma función que en las dos gráficas anteriores, pero indicando los datos correspondientes en los parámetros de la llamada de la función “plot_general”.

Las variables que contendrán los datos, generados en el capítulo anterior en el correspondiente apartado, para los sitios que más hacen referencia a las bodegas serían del tipo: “refer_NOMBREmes” y por tanto la llamada a la función sería:

```
from plot_data import *  
plot_general(refer_febrero, refer_marzo, refer_abril, refer_junio, refer_julio)
```

Tras lanzar la función, la gráfica que se generaría sería la siguiente:

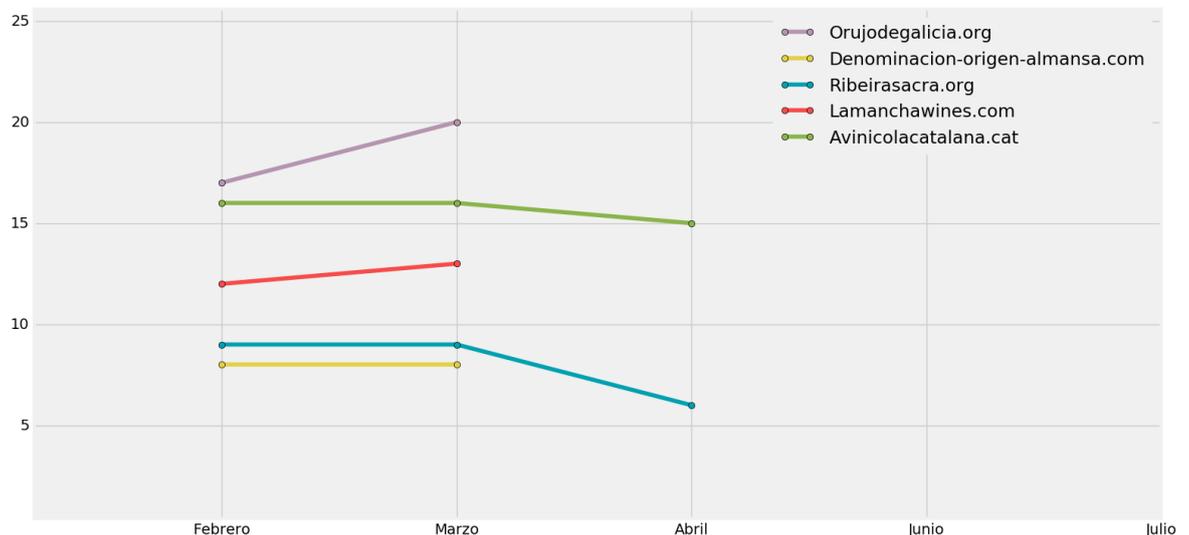


Figura 18. Gráfica con los sitios que hacen más referencia a las bodegas según similarWeb

A su vez, para representar la figura que contendrá los sitios que más visitan los usuarios después de las bodegas divididas por mes, se deberá tener en cuenta que las variables donde se almacenan los datos de cada mes según lo calculado en el apartado correspondiente del capítulo anterior serán del tipo: “dest_NOMBREmes”. Sabiendo lo anterior, la llamada a la función sería:

```
from plot_data import *  
plot_general(dest_febrero, dest_marzo, dest_abril, dest_junio, dest_julio)
```

Donde al igual que en la anterior gráfica de este apartado, se indican los datos a representar para cada mes. Por tanto, al ejecutar la lógica anterior en el intérprete de comandos, se generaría la siguiente gráfica:

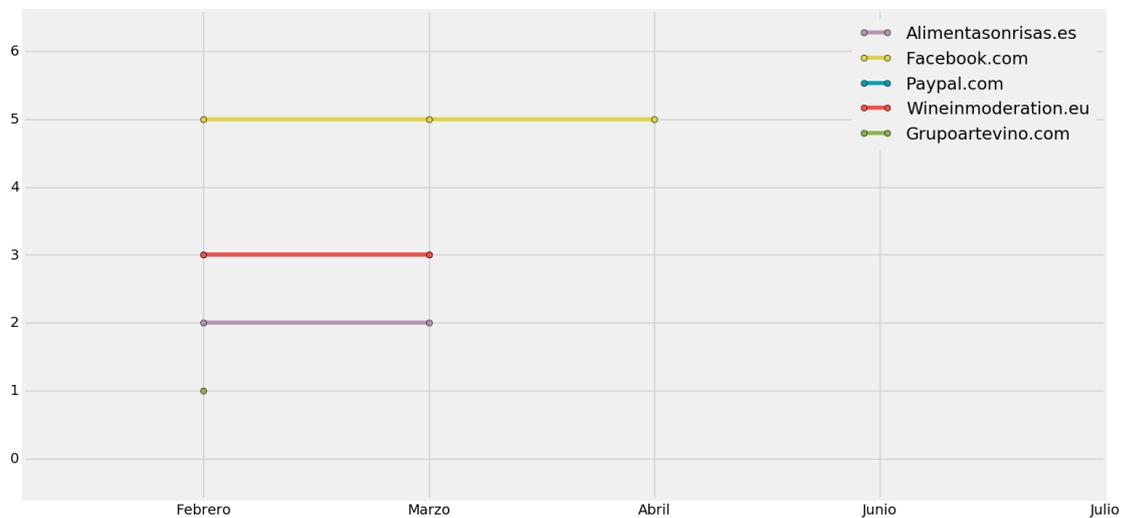


Figura 19. Gráfica con los sitios destinos de las bodegas según similarWeb

Las conclusiones a las que se han llegado una vez generadas las gráficas, se describirán en el siguiente capítulo.

3.2.4 Generación gráfica del porcentaje de uso de las redes sociales más utilizadas según similarWeb

En este apartado se explicará el proceso de generación de la representación del porcentaje de uso de las redes sociales más utilizadas según similarWeb. Para ello se parte de los datos calculados en el correspondiente apartado del capítulo anterior. Los datos que se representarán se encuentran en el diccionario clave-valor que se ha declarado como resultado de la función “per_social” y cuyo nombre es: “porce”. Este diccionario contendrá el nombre de las cinco redes sociales más utilizadas y sus respectivos porcentajes.

La representación gráfica de estos datos no será mediante barras y ejes X e Y, sino como un gráfico donut, es decir, circular pero con un hueco por el centro del mismo. Se ha decantado por este tipo de gráfico debido a que comúnmente se usa este tipo de gráficos para representar porcentajes de forma que de un vistazo se pueda apreciar cuáles son las redes sociales más utilizadas y cuáles no.

En cuanto a la lógica se refiere, se ha tenido que desarrollar una función a las demás para poder generar esta gráfica, y es la que sigue:

```
def donut_chart(datos):
    import matplotlib.pyplot as plt

    colors = ['#b594af', '#e1d146', '#00a0b0', '#f64b4a', '#8ab44b']
    explode = (0, 0, 0, 0, 0)

    plt.pie(datos.values(), explode=explode, labels=datos.keys(), colors=colors,
            autopct='%1.1f%%', shadow=True)

    centre_circle = plt.Circle((0,0),0.75,color='black', fc='white',linewidth=1.25)
    fig = plt.gcf()
    fig.gca().add_artist(centre_circle)

    plt.axis('equal')
    plt.show()
```

Como se puede observar, en la declaración de la función, el único parámetro que hace falta son los datos que se van a representar en el gráfico. Lo primero que se realizaría como siempre es importar la librería matplotlib que será la encargada de dibujar la gráfica y mostrarla. Más tarde, se declara una lista de cinco colores, debido a que se representarían cinco redes sociales, en formato hexadecimal y serían los mismos que se han ido utilizando hasta ahora para seguir con la tónica. La segunda sentencia consiste en indicar si es posible separar algún trozo del donut de los demás como manera de destacarlo. En este caso se dejarán todos los trozos unidos.



A continuación, se crea un círculo que representará los porcentajes y los nombres de las redes sociales. En esta sentencia se indican los valores que serán los porcentajes, la lista de los trozos que se van a destacar, los nombres de las redes sociales, la lista de colores, el parámetro “autopct” indica cómo se quiere mostrar el porcentaje como signo y finalmente se indica una pequeña sombra.

Lo que se acaba de indicar es que se dibuje un círculo sin el centro hueco. Para mostrar este hueco se crearía un círculo que se dibujará en el centro de fondo blanco para que sea acorde al fondo blanco en general y tenga ese aspecto de donut. Seguidamente, se crea la figura y se le asocia el círculo que aparecerá en el centro. Y finalmente, se indica que la relación de aspecto sea igual para que de esa manera el gráfico se represente como un círculo, y en la última instrucción se indica que se quiere mostrar el gráfico, el cual aparecería de la siguiente manera:

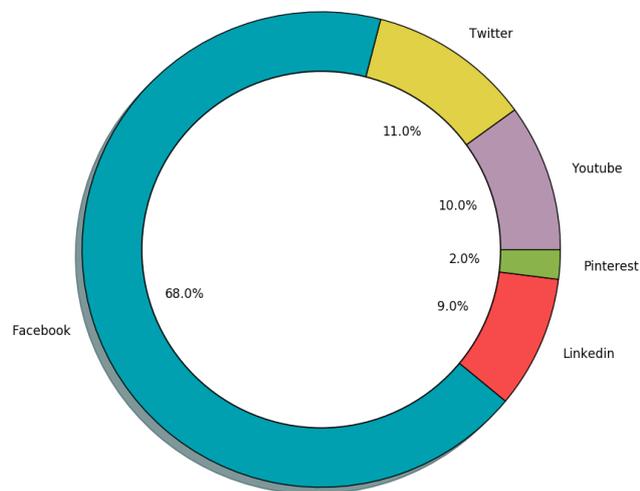


Figura 20. Gráfica con el porcentaje de uso de las redes sociales según similarWeb

Se puede ver como el nombre de las distintas redes sociales aparecen por fuera del donut y por dentro sus respectivos porcentajes. Finalmente, las conclusiones a las que se llegarán a partir de este gráfico, se describirán en el capítulo siguiente.

3.2.5 Generación de gráficas de datos comparativos entre Alexa y similarWeb

En este apartado se va a hacer uso de los datos obtenidos en el apartado correspondiente del tratamiento de datos, es decir, se van a representar visualmente la comparación entre el mismo indicador de Alexa y similarWeb. Para representar estas comparaciones, se hará uso del gráfico de tipo barras y se representarán tres gráficos. Una para analizar el “bounce rate” y los “search_visits”, otra para analizar el “time on site” y la última para analizar el “page views”. Se representan tres gráficas distintas debido a que cada uno de ellas tiene una unidad de media diferente. Para la primera serían los porcentajes, para la segunda los minutos y la tercera la cantidad de páginas por visita en decimal.

Siguiendo la temática hasta ahora, se procederá a analizar primeramente la lógica seguida para la visualización de los indicadores y después mostrar las gráficas obtenidas a partir de esta lógica.

Por tanto, se empezaría con la siguiente lógica:

```
def plot_averages(medias, tipo):

    import numpy as np
    import matplotlib.pyplot as plt

    fig = plt.figure()
    ax = fig.add_subplot(111)

    if tipo == 1: # Porcentajes
        alexa = [medias["bounce_rate_alex"], medias["search_visits_alex"]]
        similarweb = [medias["bounce_rate_sw"], medias["search_visits_sw"]]
        ax.set_ylabel('Porcentajes')
        ax.set_xticklabels(('Bounce rate', 'Search visits'))
    elif tipo == 2: # Minutos
        alexa = [medias["time_on_site_alex"]]
        similarweb = [medias["time_on_site_sw"]]
        ax.set_ylabel('Minutos')
        ax.set_xticklabels(('Time on site', ''))
    else:
        alexa = [medias["page_views_alex"]]
        similarweb = [medias["page_views_sw"]]
        ax.set_ylabel('Cantidad de paginas por visita')
        ax.set_xticklabels(('Page views', ''))

    if tipo == 1:
        ind = np.array([1.0,1.3])
    else:
        ind = np.array([1.0])
```

```
width = 0.10
```

Donde, como siempre, primero se declararía la función cuyos parámetros serían: el primero perteneciente a la información a representar y el segundo al tipo de gráfico que se va a dibujar. Si el tipo es igual a uno significa que se representará el gráfico de porcentajes donde aparecerán los indicadores “bounce rate” y “search visits”, si el tipo es igual a dos se dibujaría la gráfica con el indicador “time on site” de Alexa y similarWeb y finalmente si es tres u otro se representaría el indicador “page views”.

Se divide así la función porque las etiquetas de los ejes X e Y tendrán que ser los correspondientes a los datos a mostrar además de que se cargarán a su vez los indicadores pertinentes. Más adelante, se indican las posiciones de las barras en el eje de las X y se crea un variable que contendrá la anchura máxima de las barras.

La última parte de esta función para dibujar los datos comparativos sería la siguiente:

```
rects1 = ax.bar(ind, alexa, width, color='#00a0b0')
rects2 = ax.bar(ind + width, similarweb, width, color='#e1d146')

ax.set_xticks(ind + width)
ax.legend((rects1[0], rects2[0]), ('Alexa', 'SimilarWeb'))

def autolabel(rects, ax):
    (y_bottom, y_top) = ax.get_ylim()
    y_height = y_top - y_bottom

    for rect in rects:
        height = rect.get_height()

        p_height = (height / y_height)

        if p_height > 0.95:
            label_position = height - (y_height * 0.05)
        else:
            label_position = height + (y_height * 0.01)

        ax.text(rect.get_x() + rect.get_width()/2., label_position,
                '%d' % int(height),
                ha='center', va='bottom')

autolabel(rects1, ax)
autolabel(rects2, ax)
plt.show()
```

En esta última parte de la lógica de la función, primeramente se indica el valor que contendrán las barras cargado en función del tipo de gráfica que se haya escogido para representar, además de su posición y un color en formato hexadecimal. Seguidamente, se indica la posición de las etiquetas en los ejes de las X y la leyenda que contendrá el gráfico, en este caso “Alexa” y “SimilarWeb”.

Más tarde, se define una función “autolabel” que se basará en indicar el valor de la barra y mostrarlo encima de ella o dentro si está en los límites superiores de la gráfica. De esta forma, será más sencillo apreciar el valor de cada barra de un vistazo.

Finalmente, se llama a esta función dos veces debido a que se representarán dos barras por cada indicador, donde como parámetro se le pasan la configuración de las barras y la figura en general. Tras ello, la última sentencia consiste en mostrar el gráfico en una ventana nueva donde es posible hacer zoom en cualquier área del gráfico y donde se exportaría a un archivo externo.

Una vez explicada la lógica seguida para representar los indicadores, se procede a mostrar el resultado visual final de cada una de las tres gráficas comentadas hasta el momento.

Para representar la gráfica con los indicadores de “bounce rate” y “search visits” se tendría que realizar la siguiente llamada a la función:

```
from plot data import *  
plot_averages(media_total, 1)
```

Consecuentemente, la gráfica que se mostrará será la siguiente:

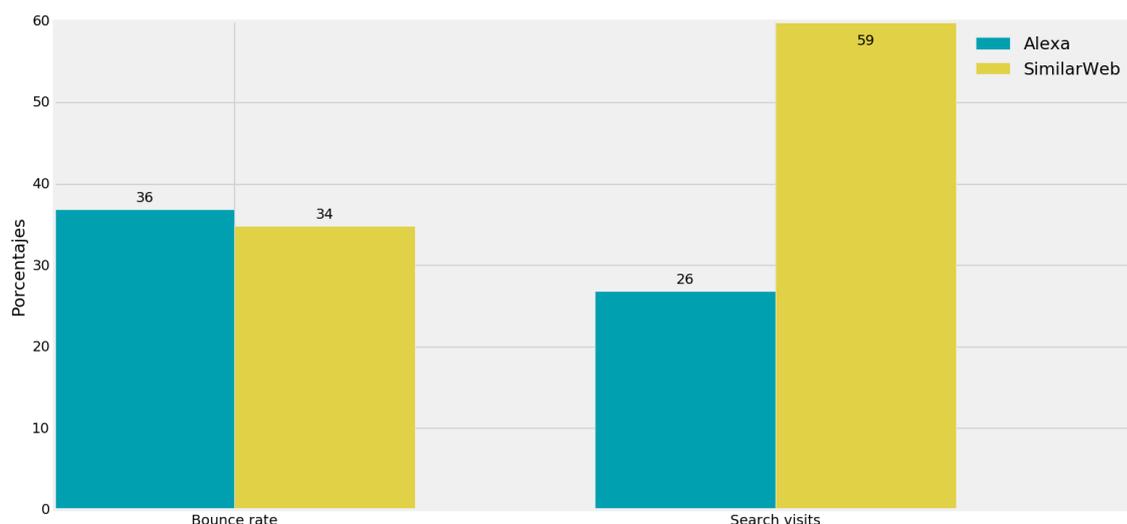


Figura 21. Gráfica que compara los campos Bounce Rate y Search visits de Alexa y similarWeb

Por otro lado, para dibujar la gráfica con el indicador de “time on site”, se tendría que realizar la siguiente llamada a la función:

```
from plot data import *  
plot_averages(media_total, 2)
```

Con ello, la gráfica que se representaría sería la que sigue:

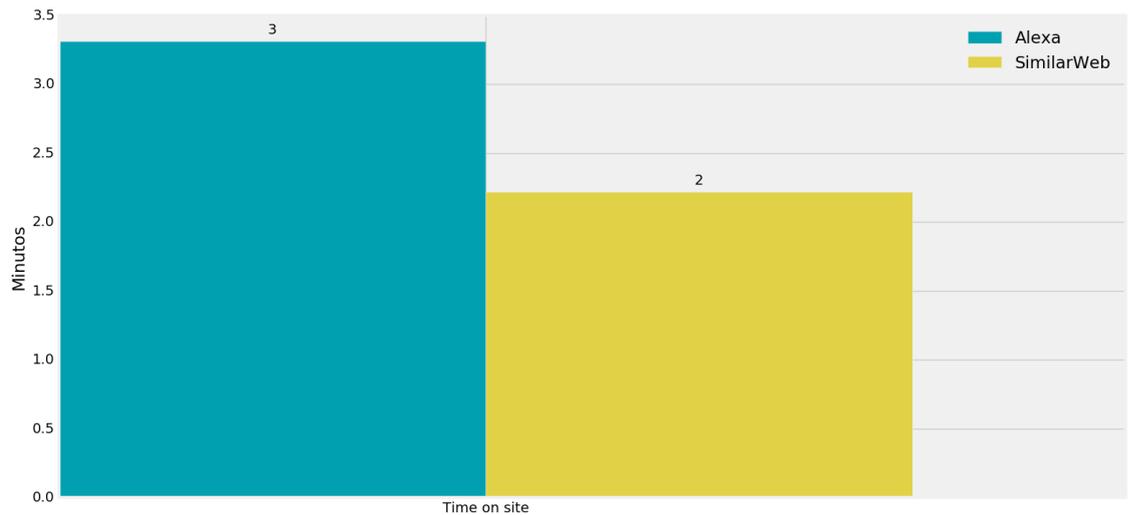


Figura 22. Gráfica que compara el Time On Site de Alexa y similarWeb

Por último, la gráfica para representar el indicador “page views”, se procedería a llamar a la función de representación de la siguiente manera:

```
from plot data import *  
plot_averages(media_total, 3)
```

Tras lo cual, la gráfica que se representaría sería la siguiente:

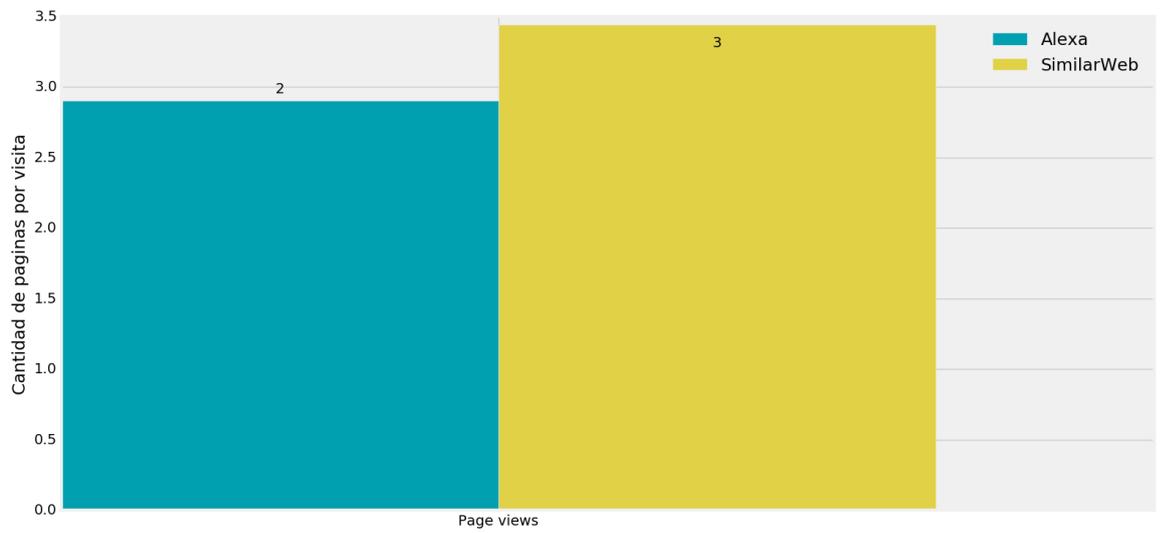


Figura 23. Gráfica que compara el Page Views de Alexa y similarWeb

3.2.6 Generación de tablas con el top de los cinco sitios a partir del trust flow, citation flow, backlinks externos y dominios de referencia según majesticSEO

En este apartado de la visualización de datos, se procederá a representar mediante tablas el top de los cinco sitios a partir del trust flow, citation flow, el número de backlinks externos y el número de dominios de referencia por mes. Es decir, por cada mes habrá una tabla con todos estos datos. Para más tarde en el capítulo de las conclusiones analizar estos datos y llegar a las debidas conclusiones.

Como continuación del capítulo correspondiente del tratamiento de datos, se explicará a partir de los datos ordenados cómo obtenerlos correctamente para definitivamente mostrarlos en una tabla.

Para ello, bastaría con introducir los siguientes comandos en el intérprete, en este caso se mostrarán los datos del mes de marzo:

```
In [43]: top_marzo["trust_flow"]
Out[43]:
[(u'torres.es', 52),
 (u'codorniu.com', 50),
 (u'freixenet.es', 50),
 (u'lopezdeheredia.com', 49),
 (u'marquesdemurrieta.com', 48)]

In [57]: top_marzo["citation_flow"]
Out[57]:
[(u'tobelos.com', 64),
 (u'vidyespiga.es', 50),
 (u'habarcelo.es', 44),
 (u'grupomatarromera.com', 43),
 (u'ontanon.es', 42)]

In [58]: top_marzo["backlinks_externos"]
Out[58]:
[(u'tobelos.com', 2221601),
 (u'vinosjuliana.com', 1470450),
 (u'principedeviana.com', 1262348),
 (u'alfredomaestro.com', 255091),
 (u'valdemar.es', 224362)]

In [59]: top_marzo["dominios_referencia"]
Out[59]:
[(u'vinosjuliana.com', 8571),
 (u'tobelos.com', 5595),
 (u'alvareznolting.com', 4195),
 (u'bodegasosca.com', 4122),
 (u'santacatalina.es', 3686)]
```

Seguidamente, se copian estos datos y se pegan en una tabla:

Sitio web	Trust flow	Sitio web	Citation flow
torres.es	52	tobelos.com	64
codorniu.com	50	vidyespiga.es	50
freixenet.es	50	habarcelo.es	44
lopezdeheredia.com	49	grupomatarromera.com	43
marquesdemurrieta.com	48	ontanon.es	42

Sitio web	Backlinks externos	Sitio web	Dominios referencia
tobelos.com	2221601	vinosjuliana.com	8571
vinosjuliana.com	1470450	tobelos.com	5595
principedeviana.com	1262348	alvareznolting.com	4195
alfredomaestro.com	255091	bodegasosca.com	4122
valdemar.es	224362	santacatalina.es	3686

Tabla 1. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de marzo

Para el mes de abril, habría que realizar lo siguiente:

```
from data_process import *
top_abril = top_majestic(data_abril)
```

La tabla correspondiente con los datos del mes de abril, sería la siguiente:

Sitio web	Trust flow	Sitio web	Citation flow
torres.es	51	tobelos.com	58
freixenet.es	50	vidyespiga.es	53
gonzalezbyass.com	49	bodegasosca.com	49
vega-sicilia.com	49	habarcelo.es	45
barbadillo.com	48	garciacarrion.es	45

Sitio web	Backlinks externos	Sitio web	Dominios referencia
tobelos.com	1439720	vinosjuliana.com	7278
vinosjuliana.com	1009148	tobelos.com	4724
principedeviana.com	796702	bodegasosca.com	3707
garciacarrion.es	434392	santacatalina.es	3287
valdemar.es	198874	bodegamontesacro.com	2541

Tabla 2. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de abril

Para el mes de mayo, habría que realizar lo siguiente:

```
from data_process import *
top_mayo = top_majestic(data_mayo)
```

La tabla correspondiente con los datos del mes de mayo, sería la siguiente:

Sitio web	Trust flow	Sitio web	Citation flow
torres.es	52	tobelos.com	53
freixenet.es	50	vicentegandia.es	46
codorniu.com	49	bodegasosca.com	46
lopezdeheredia.com	48	torredelveguer.com	42
gonzalezbyass.com	48	bodegasolicasado.com	41

Sitio web	Backlinks externos	Sitio web	Dominios referencia
tobelos.com	988345	vinosjeromin.com	5813
vinosjeromin.com	788983	tobelos.com	3432
garciacarrion.es	595758	bodegasosca.com	3425
principedeviana.com	507812	santacatalina.es	3000
maspujado.com	309791	torres.es	2317

Tabla 3. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de mayo

Para el mes de junio, habría que realizar lo siguiente:

```
from data_process import *
top_junio = top_majestic(data_junio)
```

La tabla correspondiente con los datos del mes de junio, sería la siguiente:

Sitio web	Trust flow	Sitio web	Citation flow
mendezmoya.com	52	txakolina-k5.com	53
encomiendadecervera.com	50	vianguera.com	51
fincaalbret.com	49	lasorcas.es	45
bodegasven.com	48	bodegasgutierrezdelavega.es	44
reginaviarum.es	48	suriol.com	42

Sitio web	Backlinks externos	Sitio web	Dominios referencia
txakolina-k5.com	770280	vinedosingulares.com	4991
vinedosingulares.com	669707	bodegasgutierrezdelavega.es	3136
gureahaleginak.com	594066	cellerjoansimo.com	2670
quintamazuela.es	434112	txakolina-k5.com	2624
guimaro.es	309779	mendezmoya.com	2296

Tabla 4. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de junio

En última instancia para el mes de julio, habría que realizar lo siguiente:

```
from data_process import *
top_julio = top_majestic(data_julio)
```

La tabla correspondiente con los datos del mes de julio, sería la siguiente:

Sitio web	Trust flow	Sitio web	Citation flow
mendezmoya.com	52	vianguera.com	50
encomiendadecervera.com	50	bodegasgutierrezdelavega.es	48
fincaalbret.com	49	txakolina-k5.com	47
vegatolosa.com	49	lasorcas.es	43
bodegasven.com	49	bodegasjulioresco.com	42

Sitio web	Backlinks externos	Sitio web	Dominios referencia
txakolina-k5.com	592630	vinedosingulares.com	4169
vinedosingulares.com	552260	bodegasgutierrezdelavega.es	2873
gureahaleginak.com	529496	cellerjoansimo.com	2448
guimaro.es	312324	mendezmoya.com	2307
quintamazuela.es	222521	txakolina-k5.com	1974

Tabla 5. Top cinco sitios a partir de Trust Flow, Citation Flow, Backlinks Externos y Dominios de Referencia según majesticSEO del mes de julio

4. Conclusiones

En este proyecto de investigación sobre el análisis de datos webométricos, se han extraído, tratado y visualizado datos del sector del vino español a partir de diversas fuentes para ofrecer un soporte para mejorar el posicionamiento web de las empresas y sus correspondientes marcas de negocio.

A la hora de la extracción de datos, se han tenido los mayores problemas del trabajo ya que tras haberse realizado un algoritmo extractor, éste debido a las fuentes de datos quedaba inutilizable debido a las restricciones que ponían estas fuentes, por lo que se tuvo sopesar diferentes alternativas. Las cuales, aunque cumplían con su labor, a veces fallaban al obtener ciertos datos o directamente no los obtenían y había que lanzar todo de nuevo. Por ello, como se ha visto durante todo el trabajo algunos indicadores métricos sólo estaban disponibles en ciertos meses debidos a estos problemas de extracción. Con todo esto, se ha aprendido que la extracción de datos de los sitios webs sigue siendo un problema reciente y que tiene su complejidad dependiendo sobre todo de las fuentes de datos, ya que son ellas las que ponen los límites.

En cuanto al tratamiento de datos, como se ha explicado en el apartado 2.2.1 hubo un dilema sobre qué lenguaje de programación elegir a la hora de tratar los datos. Los más aventajados eran R y Python ya que al ser de código abierto y tener una comunidad amplia detrás les daba una gran ventaja sobre los demás. Por otro lado, al principio del trabajo se propuso la idea de hacer esta parte en R, debido a que es más complejo que Python y que está hecho específicamente para el análisis de datos. Pero, debido a que el tiempo para realizar esta parte no era suficiente para poder realizar todos los cálculos previstos se optó por Python. Y no sólo por ello, sino también porque este último tiene todas las librerías necesarias para satisfacer cual necesidad a la hora de realizar todo tipo de cálculos, además de que habiendo programado en otros lenguajes de programación su sintaxis y simplicidad ayudó a entenderlo rápidamente y a usarlo con cierta ligereza pasado un tiempo.

Es importante señalar que sin la ayuda de toda la comunidad que hay el aprendizaje no hubiera sido tan rápido. Por tanto, se puede concluir que Python es una buena decisión a la hora del tratamiento de datos y no sólo para ello ya que se desenvuelve igual de bien y algunos casos incluso mejor que cualquier otro lenguaje de programación de alto nivel. Gracias a sus innumerables *frameworks* y sobre todo a la gran cantidad de documentación que existe de todo tipo, incluido ejemplos, para lograr aprender de forma rápida y correcta.



En cuanto a la visualización de datos se refiere, ésta ha sido igual de satisfactoria que el tratamiento de datos debido a que se ha utilizado con el mismo lenguaje de programación, es decir, con Python. En esta parte del trabajo, también se partía completamente de cero, pero debido a la cantidad de ejemplos ^[24] que hay totalmente documentados la curva de aprendizaje ha sido mínima. Por otro lado, también indicar que ha habido casos de gráficas que no se han podido realizar debido a que no se tenían los conocimientos adecuados y que después de ciertas pruebas y buscando posibles soluciones en foros no se han podido representar, sin embargo, se encontró otras alternativas igual de válidas.

En última instancia, a partir de los resultados que se han ido obteniendo a lo largo del trabajo, se sacarán un conjunto de conclusiones que a continuación se exponen con mayor detalle.

4.1 Conclusiones de los resultados

Dada la cantidad de conclusiones que se pueden llegar a alcanzar a partir de los resultados obtenidos y de los criterios aplicados, se expondrán aquellas conclusiones que se han creído más interesantes y relevantes a partir de los resultados recabados en el presente trabajo.

Los primeros resultados obtenidos fueron el tráfico de países que más visitan las bodegas de vino españolas según la fuente de datos similarWeb. Como es lógico, el principal país que realiza estas visitas es España cuyos índices han llegado a ser el doble comparado con el segundo país, el cual es Estados Unidos. Esto puede ser un punto de inflexión para las bodegas, ya que si por parte de este país hay un gran interés (con la visita a los sitios web se demuestra) por los vinos españoles, éstas podrían enfocarse en ese mercado, primero haciendo el sitio web multi-idioma, es decir, que en función de la ubicación del usuario se muestre la web en español o en inglés. Con esta simple adición los usuarios ya se sienten más atraídos y con la sensación de estar bienvenidos.

Si con ello al volver a realizar otro análisis, los indicadores mostrasen que hay un mayor aumento de visitas por parte de este país, se podría barajar la opción de exportar vinos a este país y con ello introducirse en el mercado del país. Continuando con los demás países, se puede apreciar como aparecen dos países de la zona europea como es Reino Unido y Francia, algo que sin duda también se puede tomar en consideración a la hora de ampliar el negocio. Y teniendo además la ventaja de que ambos países están más cerca que Estados Unidos, y sobre todo Francia al ser país vecino.

²⁴ Ejemplos de matplotlib. http://matplotlib.org/api/pyplot_api.html



Por último comentar, como se mantiene de estable México, mientras que Francia y Reino Unido en los últimos meses ya no aparecen, éste permanece de forma estable a lo largo de los meses, algo que sin duda juega a su favor.

Los siguientes resultados de los extraídos son los países que más visitan las bodegas de vinos españolas por parte de la fuente de datos Alexa. Estos datos como se ha podido comprobar en la gráfica del apartado 3.2.2 son poco fiables, sí es verdad que en cabeza de todos está España, como es lógico, pero el resto de países ofrecen unos resultados mínimos que dejan poco que considerar de ellos. Al comparar ambas fuentes de datos, se ha podido observar como similarWeb ha ofrecido datos bastante interesantes con los que se puede trabajar a la hora de mejorar el posicionamiento web de la empresa. Pero con los de Alexa, poco se puede hacer.

A continuación, se tienen los resultados de los cinco sitios que más han hecho referencia a las bodegas de vinos, además de los sitios destinos que visitan los usuarios después de acceder al sitio web, todo esto a partir de la fuente de datos de similarWeb. En cuanto a los sitios que hacen más referencia a las bodegas, esto viene a decir, que en los sitios indicados en la primera gráfica correspondiente al apartado 3.2.3, hay ciertos enlaces que redirigen a distintos sitios de bodegas de vino españolas ya sea por algún convenio que se tenga entre ambos sitios o porque hay diferentes noticias relacionados con esas bodegas, entre otras muchas opciones. Es destacable que tanto en junio como en julio ha habido un corte de cero datos, debido a que los sitios que se mostraban en esos meses ofrecían resultados tan mínimos que no merecían la pena representarlos, es decir, que a partir de esos dos meses hubo un bajón importante de webs que hacían hasta entonces las mayores referencias.

Por su parte, a las de destinos les ocurre lo mismo para el mes de junio y julio. Pero para los tres anteriores meses correspondientes a febrero, marzo y abril, se puede apreciar como el líder es la red social de Facebook. Esto puede ser debido a que los sitios webs de las bodegas tengan una página creada en Facebook para añadir diferentes noticias sobre su marca y productos. Algo que sin duda juega a su favor, pero para los que no es así es una desventaja importante debido a como se ha comentado en la introducción, una empresa sin ningún tipo de cuenta en redes sociales pierde el estar al tanto de lo que las personas opinan sobre la misma. Por su parte, en esta página de destinos también se echa en falta la red social de Twitter, algo igual de imprescindible que Facebook.

Continuando con el tema de las redes sociales, se tienen los resultados en cuanto a las redes sociales más utilizadas según la fuente de datos de similarWeb. Los siguientes datos, no hay que tomárselo al pie de la letra, pero sí tenerlos en cuenta. Ya que, estos datos son obtenidos de aquellas bodegas que tenían disponible este indicador lo cual es el 5-10% de todas las bodegas. La razón de que este porcentaje se tan bajo es debido a que los sitios webs de las bodegas son poco conocidos o generan muy poca actividad entre usuarios, entre otros posibles problemas. Dicho esto, se puede observar como el dominador es sin duda Facebook con un 68% de dominio sobre los demás, algo

que a día de hoy no debería extrañar a nadie ya que es la red social más popular hasta el momento.

En segundo lugar, está Twitter, que como se ha comentado anteriormente debería tener mucho más porcentaje ya que es igual de importante que Facebook dada la gran actividad que genera. Al igual que Twitter, un porcentaje menos está Youtube, la red social de los vídeos por excelencia, que, aunque su porcentaje es igual de bajo que Twitter, no tiene la misma influencia como es el caso de éste.

Siguiendo con los resultados, se consideran los diferentes indicadores que tienen en común Alexa y similarWeb, que son el “bounce rate”, “search visits”, “time on site” y “page views”. Los primeros indicadores se han representado en una misma gráfica debido a que comparten la misma unidad de medida, que es el porcentaje. Estas comparaciones de indicadores entre ambas fuentes de datos tenían como objetivo ver hasta qué punto son fiables y cuál es la diferencia entre ambos. En cuanto al primer campo, el “bounce rate” se puede apreciar como ambos tienen prácticamente el mismo valor lo que sugiere que al tener cada uno un método obtención de datos diferente llegan a la misma conclusión lo que hace de dato fiable y a tenerlo en cuenta en cualquier análisis.

En cuanto al indicador de “search visits” éste indica el porcentaje de veces que la visita al sitio web se ha realizado desde un motor de búsqueda como Google o Bing. Aquí se puede observar que entre ambas fuentes de datos hay una diferencia de un 33% una diferencia bastante elevada, que sugiere que hay discrepancias entre ambos y que no son fiables estos datos. Sin embargo, se puede tener una aproximación media que sirva de orientación a la hora de tener en cuenta este indicador cuando se vaya a mejorar en un sector u otro del posicionamiento web.

Por último, los indicadores de “time on site” y “page views” son aquellos que indican el tiempo que un usuario permanece de media en un sitio web y la cantidad de páginas distintas que accede en ese tiempo, respectivamente. Al observar la gráfica de “time on site” se puede apreciar como hay una diferencia notable, siendo de Alexa 3.3 y de similarWeb 2.2 una diferencia de un minuto entre ambas métricas es algo a destacar y tal como se ha señalado para el indicador de “search visits”, de este campo no es posible fiarse en cierta medida, pero sin embargo, sí tener una aproximación a tener en cuenta. En último término, el campo “page views”, sí que es posible fiarse en mayor medida que el analizado anteriormente ya que los datos mostrados tienen poca diferencia entre ambos lo que llevado a la conclusión final se puede decir que de media un usuario que entra a un sitio de una bodega de vinos españolas suele visitar alrededor de tres páginas del sitio total, lo que puede ayudar a estas a mejorar el contenido de estas páginas para que el usuario visite no sólo una sino más páginas.

En último término, se tienen las tablas con el top de los cinco sitios ordenados de mayor a menor según los campos “trust flow”, “citation flow”, número de “backlinks externos” y número de “dominios referencia” y además se muestran los resultados por cada mes. El objetivo de extraer estos datos, era analizar si el primer sitio de un top

según un campo implicaba ser el primero en otro campo y lo mismo con las consiguientes posiciones.

Como se puede observar en las diferentes tablas a lo largo de los meses, en los tres primeros correspondientes a marzo, abril y mayo el líder tanto en el campo “citatio flow” como en el número de “backlinks externos” es el mismo, en este caso “tobelos.com” lo que a primera vista podría implicar un campo con el otro. Sin embargo, al observar las siguientes posiciones se aprecia como ninguna posición del primer campo tiene que ver con la pertinente en el otro campo, llegando incluso a que los sitios web son distintos coincidiendo sólo en el primero. Este suceso si bien podría ser una casualidad, habría que demostrarlo con más datos comparativos de muchos más sitios.

En lo que a los demás campos se refiere, se puede ver como no tienen nada en común en cuanto a los cinco sitios web que están en el top de cada uno. Si bien podría significar que los campos no son dependientes de otros y uno no implica lo otro, para tener certeza de ello, se tendría que realizar un estudio de análisis de estos dos campos con una cantidad mayor de sitios web y no restringida a un sector como es este caso el vino, sino con datos globales de la fuente de datos.

4.2 Conclusiones futuras

De cara a futuros estudios o investigaciones, se puede decir que este trabajo puede ser un comienzo o una base para estudios más avanzados en la materia, que con las experiencias adquiridas en este trabajo se puede avanzar sin ningún tipo de límite y aprovechar con ello el potencial que pueden tener las herramientas y lenguajes de programación utilizados a lo largo del trabajo.

Tal como se ha comentado anteriormente, la gran limitación han sido las fuente de datos y sus diversas restricciones a la hora de obtener los indicadores, algo que si no hubiera ocurrido se hubiesen obtenido más datos y más fiabilidad en ellos.

Algo que sin duda ha dado buenos frutos ha sido el lenguaje de programación utilizado para el tratamiento y visualización de datos, es decir, Python. El lenguaje en sí permite analizar miles de datos en apenas segundos lo que para el posicionamiento web y el big data en general es una ventaja a tener en cuenta.

Sin embargo, para estudios más avanzados también se tiene muy cuenta el lenguaje R, que como se ha visto mediante gráficas, en la industria del software es muy popular y lo prefieren antes que a Python. Por tanto, este lenguaje hay que tenerlo presente a la hora de analizar grandes cantidades de datos o cuando se vayan a realizar cálculos complejos.

Por otra parte, un mayor número de webs a analizar no sólo de España sino de otros países otorgaría más alicientes para un gran proyecto de investigación de

indicadores métricos. Por su parte, una clasificación detallada de cada una de las bodegas proporcionaría un análisis más detallado y conciso.

En última instancia, y al igual que se ha realizado este estudio para el sector del vino, perfectamente puede ser aplicable a cualquier otro. Este trabajo por consecuencia, aporta su granito de arena al análisis del posicionamiento web y la webometría donde este sector cada día va en aumento con más profesionales y nuevas herramientas que ayudan a las empresas a encontrar su nicho de mercado.

5. Referencias

- McKinney, Wes. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. Editorial O'Reilly Media; Edición: 1 (1 de noviembre de 2012).
- Björneborn, L. and Ingwersen, P. (2004), *Toward a basic framework for webometrics*. *J. Am. Soc. Inf. Sci.*, 55: 1216–1227. doi:10.1002/asi.20077

6. Anexos

6.1 Anexo A

Tabla 6. Sitios web de las bodegas

Sitio web
10dabril.com
10sentits.cat
15albas.es
2859.es.all.biz
2amigos.com
3asesvino.com
40gradosnorte.com
4kilos.com
7magnifics.com
a2vinoycultura.com
aalto.es
abadal.net
abadiadeacon.com
abadiadearibayos.es
abadialaarroyada.es
abadia-retuerta.com
abarando.com
abeica.com
abiotxakolina.com
acontia.es
acoroa.com
actualidad.campante.com
acuerpoderey.com
acusticceller.com
adegaalgueira.com
adegaaraujo.com
adegaavelina.com
adegadapinguela.com
adegadofortes.multiespaciosweb.com
adegadosan.com
adegaideos.com
adega-illo.com

adegamanuelrojo.es
adegameisotero.com
adegamelillas.com
adegamoraima.com
adegasamedo.com
adegasameiras.com
adegasarousa.com
adegasgalegas.es
adegasgranvinum.com
adegaslameiro.com
adegasluz.es
adegasmoure.com
adegasribouzo.com
adegasterrasanta.com
adegasvaldavia.es
adegasvalminor.com
adegavaldes.com
adegavieites.com
adernats.cat
adoblasmartos.com
aecovi-jerez.es
agnesdecervera.com
agricolacastellana.com
agricolacorbera.com
agrodebazan.com
agrodebazansa.es
agrovicolamontserrat.com
agryenca.com
aguardientesdegalicia.es
aguiuncho.com
agustitorellomata.com
ailalawine.com
akiliawines.com
aladro.es
alandeval.com
alavadelacruz.blogspot.com.es
albertoledo.com
albetinoya.cat
albetinoya.com
alcardet.com
alceno.com
alcoholesiglesias.com
alcortavino.com

alcovibodega.com
aldahara.es
aldeanueva.com
aldeavino.com
aldonia.es
alellavinicola.com
alfredomaestro.com
algramar.com
algueira.com
aliaguillerabodega.com
alianzadegarapiteros.es
allozo.com
almazcaramajara.com
alodia.es
alonsocuesta.com
alsinasarda.com
altaalella.cat
altapavina.com
altavins.com
altolandon.com
altorredondo.es
altosdecristimil.com
altosdelcuadrado.com
altosdelmarques.com
altosdelterral.com
altosderioja.com
altosdesanesteban.com
altosdetorona.com
altovela.com
alturia.es
alvareznolting.com
alvarezydiez.com
alvarodomecq.com
alvear.es
alvides.com
alvisa.es
alzania.es
amadisdegaula.com
ametller.com
ampans.cat
ampriuslagar.es
amunategi.eu
anadareal.com

andresmorate.com
annegra.com
antabanderas.com
antiguabodegasanblas4.es
antiguacasadeguarda.net
antiguacasadeguardia.com
antiguausanza.com
antoniomontero.com
apricus.es
arabarte.es
arabate.com
araex.com
aranleon.com
aranleon.es
arboledamediterranean.com
arcadevitis.com
archscellers.com
arcobu.com
arcocuriel.com
ardoa.com
arenaltrading.es
arenaltrading.squarespace.com
aretey.com
argueso.es
aribauroja.com
arinzano.es
arloren.com
armeroiadriver.com
aroawines.com
arpelaza.es
arrayan.es
arrocal.com
artadi.com
artcava.com
artiga-fustel.com
artoje.blogspot.com.es
artuke.com
aruspide.com
arvavitis.com
arzuaganavarro.com
asenjo-manso.com
asfurnias.com
astrales.es

asunciongumiel.com
atroca.eu
attisbyv.com
attisbyv.es
aurumwine.com
avanteselectagrupo.com
avanthiawines.com
avanvinos.com
avbodeguers.com
avelinovegas.com
avenencia.es
avgvstvs.es
avgvstvsforvm.com
avilamajo.com
axpesagardotegia.com
aymnavarro.com
azpilicuenta.com
azulygaranza.com
bach.es
badennumen.es
bagordi.com
bagordi.es
balbas.es
baldisabena.com
barahonda.com
barbadillo.com
barbadillo.net
barbarot.com
barbarot-wines.com
barcolobo.com
bardinet.es
barondalba.com
barondeley.com
barondelsolar.com
baroniadeturis.es
baronia-m.com
barrancooscuro.com
basagoiti.es
basarte.net
batandesalas.com
batlliudesort.cat
bbesalduchvalls.es
beadeprimacia.com

beiraaral.com
beldui.com
belezos.com
bellorivinos.com
belondrade.com
benetakoa.com
benitosantos.com
berarte.es
berberana.com
bernabeleva.com
bernavi.com
beronia.com
beronia.es
berralmiro.com
betolaza.es
bhrubio.com
bikanditxakolina.com
binifadet.com
binigrau.es
biob.es
bioca.es
biurkogorri.com
bjn1963.com
blancher.es
boadochaves.es
bobaldesanjuan.com
bocopa.com
bocoswine.com
bodegaabadiadelaoliva.com
bodegaacevedo.com
bodegaalanis.com
bodegabelloberganzo.com
bodegabenaya.com
bodegabenitorodrigo.es
bodegabernal.com
bodegaberroja.com
bodegabiniagual.com
bodegablecua.com
bodegacamilin.com
bodegacarres.com
bodegacasadepueblas.com
bodegacastillodeeneriz.com
bodegacastillodepeñaranda.com

bodegacastrorei.com
bodegacauzon.blogspot.com.es
bodegacerrosancristobal.es
bodegacezar.com
bodegacheca.com
bodegaclassica.com
bodegacobertizo.com
bodegacomarcal.es
bodegacomarcalguimar.com
bodegacontador.com
bodegaconvento.com
bodegacooperativacigales.com
bodegacortijolafuente.es
bodegacosecheros.com
bodegacrayon.es
bodegacuatrovientos.es
bodegacyan.es
bodegadehesadeloscanonigos.com
bodegadelabad.com
bodegadelaserna.com
bodegadelcamino.com
bodegadeljardin.es
bodegadeparrado.es
bodegadesada.com
bodegadesarria.com
bodegadeseron.es
bodegadiazbayo.com
bodegadiegodelemos.blogspot.com
bodegadiegodelemos.blogspot.com.es
bodegadiezmonuevo.com
bodegadominiodelbendito.com
bodega-edra.com
bodegaeduardogarrido.es
bodegaeduardopenha.es
bodegaehd.com
bodegaelangosto.com
bodegaelarcadenoe.com
bodegaelfabulista.com
bodegaellomo.com
bodegaelmocanero.com
bodegaelvalle.es
bodegaencarnacion.es
bodegaestebanaraujo.com



bodegafeo.es
bodegaflors.com
bodegafuentevictoria.es
bodegagarciadelarosa.es
bodegagarnachaaltoalberche.com
bodegagilarmada.com
bodegagordaliza.es
bodega-gotica.com
bodegahemar.com
bodegahiriart.es
bodegahmesa.com
bodegahorta.cat
bodegahoyosdebandama.com
bodegainiesta.es
bodegainurrieta.com
bodegairi.com
bodegairius.com
bodegajaviersanz.com
bodegajesusromero.com
bodegajf.es
bodegakieninger.com
bodegalacapuchina.es
bodegalaencina.com
bodegalamilagrosa.com
bodegalamioga.com
bodegalaserrana.com
bodegalasgranadas.com
bodegalasolana.es
bodegalasolidad.com
bodegalatercia.com
bodegalesuseres.es
bodegalosmatucos.com
bodegalosolmos.com
bodegalospalomares.com
bodegaluz.com
bodegamanoamano.com
bodegamaranones.com
bodegamarenas.com
bodegamarquesdelpuerto.com
bodegamartinezyebra.es
bodegamasasturias.com
bodegamaslaltet.com
bodegamassana.com

bodegamatarromera.es
bodegamatsu.com
bodegamedinaceli.com
bodegamendiko.com
bodegamipanas.com
bodegamonasterio.com
bodegamonastrell.com
bodegamondelo.com
bodegamontesacro.com
bodegamustiguillo.com
bodegaobalo.com
bodega-ottobestue.com
bodegapagodecubas.com
bodegaparis.com
bodegapirineos.com
bodegaproexa.wordpress.com
bodegapuenteajuda.com
bodegaramanya.com
bodegaramonramos.com
bodegaregantio.es
bodegarento.es
bodegarequiemhispania.com
bodegareveron.com
bodegaribas.com
bodegariberadelarga.blogspot.com.es
bodegaroandi.com
bodegaruizdevinaspre.com
bodegasaavedra.com
bodegasababol.com
bodegasabanico.com
bodegasabaxterra.com
bodegasabinasa.com
bodegasacosta.multiespaciosweb.com
bodegasadria.com
bodegasafersa.com
bodegasagapita.com
bodegasagrovello.com
bodegasaguirre.es
bodegasalabanza.com
bodegasalaude.com
bodegasalbamar.es
bodegasalcubilla.com
bodegasalfonsogarciahernando.es



bodegasaliciarojas.com
bodegasalion.com
bodegasalmijara.com
bodegasalmocaden.com
bodegasalore.com
bodegasaltanza.com
bodegasaltoalmanzora.com
bodegasaltomoncayo.com
bodegasaltona.es
bodegasaltun.com
bodegasalvar.com
bodegasalzagal.es
bodegasamadorgarcia.com
bodegasamaren.com
bodegasamezola.net
bodegasamosurtubia.es
bodegasanacacio.com
bodegasanatorio.com
bodegasanblas.es.tl
bodegasandion.com
bodegasandrade.es
bodegasangabriel.com
bodegasangel.com
bodegasangelortiz.com
bodegasangregorio.com
bodegasanhelo.com
bodegasani.com
bodegasanisidro.es
bodegasanjose.com
bodegasanmames.com
bodegasanmarcos.com
bodegasanmartin.com
bodegasanmiguelsc.es
bodegasanpedroapostol.com
bodegasanprudencio.com
bodegasanprudencio.es
bodegasanrafael.com
bodegasanroman.com
bodegasanroquedelaencina.com
bodegasantacruz.com
bodega-santamaria.com
bodegasantamarta.com
bodegasantiagoruiz.com

bodegasantidoto.com
bodegasantonioalcaraz.com
bodegasantoniobenito.com
bodegasantoniocandela.com
bodegasantonionadal.es
bodegasanzil.es
bodegasaquitania.com
bodegasaraco.com
bodegasaragonesas.com
bodegasarane.com
bodegasarbas.com
bodegasarlanza.com
bodegasarmando.es
bodegasarmentero.com
bodegasarraez.com
bodegasarribesdelduero.com
bodegasartajona.com
bodegasarturo.es
bodegasasensio.com
bodegasaslaxas.com
bodegasaster.com
bodegaseca.es
bodegasaugustabilbilis.com
bodegasayagar.es
bodegasayerra.com
bodegasayles.com
bodegasayuso.es
bodegasazpea.com
bodegasbaigorri.com
bodegasballabriga.com
bodegasbalmoral.com
bodegasbargondia.com
bodegasbaron.es
bodegas-barreda.com
bodegasbarreiroamado.com
bodegasbarrialba.com
bodegasbasilioizquierdo.com
bodegasbegastri.com
bodegasbellod.com
bodegasbenitoblazquez.com
bodegasbentayga.com
bodegasbentomiz.com
bodegasberamendi.com

bodegasbernabenavarro.com
bodegasbernardoalvarez.com
bodegasberrueco.com
bodegasbg.es
bodegasbilbainas.com
bodegasbiosca.es
bodegasblancareyes.com
bodegasbleda.com
bodegasbleda.es
bodegasbohorquez.com
bodegasbordeje.com
bodegasbordoy.es
bodegasborsao.com
bodegasbriego.com
bodegascabestrero.com
bodegascalar.com
bodegascamilocastilla.com
bodegascampillo.com
bodegascampina.net
bodegascamposreales.com
bodegascanalva.com
bodegascandido.com
bodegascandido.net
bodegascano.com
bodegscanopy.com
bodegscanrich.com
bodegascantamora.com
bodegascarballo.es
bodegascardema.com
bodegascarlosplaza.com
bodegascarlossampedro.com
bodegascarreno.com
bodegascarsalo.com
bodegascasadelavina.com
bodegas-casajuan.com
bodegascasajus.com
bodegascasaldearman.com
bodegascasaprimicia.com
bodegascasas.com
bodegascasis.com
bodegascastano.com
bodegascastejon.com
bodegascastelar.com

bodegascastiblanque.com
bodegascastillejo.com
bodegascastillejoderobledo.com
bodegascastresana.com
bodegascastromartin.com
bodegascaudalia.com
bodegascave.com
bodegascelaya.com
bodegascenit.com
bodegascerrolaza.com
bodegascerrosol.com
bodegascesardelrio.com
bodegascesarvelasco.com
bodegaschaves.com
bodegaschesa.com
bodegaschivite.com
bodegaschp.es
bodegaschvinosdecubillas.com
bodegascollado.com
bodegascoloma.com
bodegascomeche.com
bodegasconcejodeollauri.com
bodegascondales.com
bodegas-contreras.com
bodegasconvento.com
bodegasconventodelasclaras.com
bodegascorellanas.com
bodegascorunadelconde.com
bodegas-cotoredondo.com
bodegascovilor.com
bodegascrisve.com
bodegascrisve.es
bodegascruzconde.com
bodegascruzconde.es
bodegascubero.com
bodegascueva.es
bodegascunqueiro.com
bodegascunqueiro.es
bodegasdanielpuras.es
bodegasdavila.com
bodegasdearanda.com
bodegasdelgado.com
bodegasdelmedieval.com

bodegasdelmuni.com
bodegasdelosriosprieto.com
bodegasdelromero.com
bodegasdelrosario.es
bodegasdelsaz.com
bodegasdelsenorio.com
bodegasdelsocorro.com
bodegasderioja.com
bodegasdesantiago.es
bodegasdiaz.com
bodegasdiezllorente.com
bodegasdiosares.com
bodegasdiosbaco.com
bodegasdocampo.com
bodegasdomecodejarauta.com
bodegasdominguez.es
bodegasdonatoda.com
bodegasdondiego.es
bodegasdunviro.com
bodegasecadero.com
bodegaseguia.com
bodegaseguiluz.es
bodegaseidosela.com
bodegasejeanas.com
bodegaseladiosantalla.com
bodegaselcastillo.com
bodegaselcidacos.com
bodegaselgallo.es
bodegaselhoyo.com
bodegaseliasmora.com
bodegaselmonte.com
bodegaselnido.com
bodegaselosegui.com
bodegaselpenitente.es
bodegaselprogreso.com
bodegaselsoto.com
bodegaseltanino.com
bodegasenguera.com
bodegasentremontes.com
bodegasercavio.com
bodegaservilio.com
bodegasescudero.com
bodegasestada.com

bodegasesteban.es
bodegasestraunza.com
bodegaseverinosanz.es
bodegasevine.com
bodegasexcelencia.com
bodegasexeo.com
bodegasfabregas.com
bodegasfarina.com
bodegasfarran.com
bodegasfaustino.com
bodegasfaustinogarcia.es
bodegasfaustinogonzalez.com
bodegasfdzeguiluz.com
bodegasfederico.com
bodegasfelixcallejo.com
bodegasfelixsalas.com
bodegasfelixsanz.es
bodegasfermingilar.com
bodegasfernandez.es
bodegasfernandez.net
bodegasfernandocastro.com
bodegasferris.com
bodegasfillaboa.com
bodegasfincavaqueros.com
bodegasfindesiglo.com
bodegasfondarium.es
bodegasfontana.com
bodegasforcada.com
bodegasfos.com
bodegasfranciscogomez.es
bodegasfrontaura.com
bodegasfrutosvillar.com
bodegasfuenmayor.com
bodegasfuenmayor.es
bodegasfuentevieja.com
bodegasfundadorpedrodomecq.com
bodegasfusion.com
bodegasfuso.com
bodegasgailur.com
bodegasgallardosl.com
bodegasgallegas.com
bodegasgallegozapatero.com
bodegasgancedo.com



bodegasgarcia.com
bodegasgarciaburgos.com
bodegasgarciadeolano.com
bodegasgarciadeverdevique.com
bodegasgarcianoguerol.com
bodegasgarciamirez.com
bodegasgarmendia.com
bodegasgarridomedrano.com
bodegasgerardomendez.com
bodegasgomeznevado.com
bodegasgonzalezlara.com
bodegasgonzalovalverde.es
bodegasgrant.com
bodegasgrm.com
bodegasgualda.com
bodegasgurrutxaga.com
bodegasgutierrezdelavega.es
bodegashabla.com
bodegasharveys.com
bodegashcalvente.com
bodegashcmerino.com
bodegasheredadanson.com
bodegashermanospecina.com
bodegashermosilla.com
bodegashijosdelisardogarcia.multiespaciosweb.com
bodegashispanosuizas.com
bodegashmartin.com
bodegasholgado.es
bodegashuertas.com
bodegasibaiondo.com
bodegasidiaquez.com
bodegasierranorte.com
bodegasietecerros.com
bodegasigarmi.com
bodegasiglesias.com
bodegasillana.com
bodegasimperiales.com
bodegasinarcas.es
bodegasinsulares.es
bodegasiranzo.com
bodegasisidromilagro.com
bodegasisla.com
bodegasitsasmendi.com

bodegasiturbide.com
bodegas-iturria.com
bodegasizquierdo.com
bodegasjalon.com
bodegasjer.es
bodegasjesusdiazehijos.com
bodegasjosecordoba.com
bodegasjsantos.com
bodegasjuangil.com
bodegasjuliocrespo.com
bodegasjuncales.es
bodegaskefren.com
bodegaslaaurora.com
bodegaslabastida.com
bodegaslaciarrera.com
bodegaslacort.com
bodegaslacorte.com
bodegaslacus.com
bodegasladairo.com
bodegaslodefensa.es
bodegasladrero.es
bodegaslaemperatriz.com
bodegaslaespada.com
bodegaslaestacion.es
bodegaslagardezabala.com
bodegaslahorra.es
bodegaslahoz.com
bodegaslalaguna.com
bodegaslalanne.es
bodegaslambuena.com
bodegaslan.com
bodegaslandaluce.es
bodegas-langa.com
bodegaslapurisima.com
bodegaslarchago.com
bodegaslarodetta.com
bodegaslarraz.com
bodegaslasangrederonda.es
bodegaslasierra.es
bodegaslasoledad.com
bodegaslasoterrana.com
bodegaslastirajanas.com
bodegaslatente.com



bodegaslaukote.com
bodegaslauna.com
bodegaslaunion.com
bodegaslaus.com
bodegaslaval.com
bodegaslavalle.com
bodegaslavia.com
bodegaslecanda.com
bodegaslecea.com
bodegasleda.com
bodegasleganza.com
bodegasleopoldo.com
bodegasleza.com
bodegaslicinia.es
bodegaslleiroso.com
bodegasloeda.com
bodegaslolicasado.com
bodegaslopezmorenas.com
bodegaslopezoria.com
bodegaslopezromero.com
bodegasloreto.com
bodegaslosfrailes.com
bodegaslospinos.com
bodegas-lozano.com
bodegasluisangelcasado.com
bodegasluismarin.com
bodegasluisperez.com
bodegaslunares.com
bodegasluque.es
bodegasluva.es
bodegasluzon.com
bodegaslyng.es
bodegasmacaya.com
bodegasmadai.com
bodegasmadronal.com
bodegasmaillo.com
bodegasmalagavirgen.com
bodegasmanuelbastias.com
bodegasmanueldelaosa.com
bodegasmanuelmoneva.com
bodegasmarba.com
bodegasmargon.com
bodegasmarinpalacios.com

bodegasmartinezalesanco.com
bodegasmartinezcorta.com
bodegasmartinezherrero.es
bodegasmartinezpalacios.com
bodegasmartinezsaez.es
bodegasmartinezserantes.com
bodegasmartinon.com
bodegasmasalta.com
bodegasmasquevinos.com
bodegasmasvida.com
bodegasmateos.com
bodegasmauro.com
bodegasmaximoabete.com
bodegasmedina.net
bodegasmedinaytoro.es
bodegasmedranoirazu.com
bodegasmeler.com
bodegasmelwa.com
bodegasmendieta.com
bodegasmendoza.com
bodegasmerayo.com
bodegasmiguelaguado.com
bodegasmiguelgrueso.com
bodegasmitos.com
bodegasmocen.es
bodegasmonfil.es
bodegasmonje.com
bodegasmonroy.es
bodegasmontebaco.com
bodegas-monteblanco.es
bodegasmontecastro.es
bodegasmontsalla.es
bodegasmoral.es
bodegasmoralia.es
bodegasmorate.com
bodegasmoraza.com
bodegasmorosanto.com
bodegasmorte.com
bodegasmoyapalafox.es
bodegasmuga.com
bodegasmunoz.com
bodegasmunozmartin.com
bodegasmuriel.com

bodegasmurilloviteri.com
bodegasmuro.es
bodegasmurua.com
bodegasmuseum.com
bodegasnaia.com
bodegasnairoa.com
bodegasnajerilla.es
bodegasnanclares.es
bodegasnaovictoria.com
bodegasnaranjo.com
bodegasnavajas.com
bodegasnavarioja.com
bodegasnavarro.com
bodegasnavarro.es
bodegasneo.com
bodegasnexus.com
bodegasnilo.com
bodegasnodus.es
bodegasnoroeste.com
bodegasnubori.com
bodegasnuevavalverde.com
bodegasocas.com
bodegasocellumdurii.es
bodegasochagavia.com
bodegasochoa.com
bodegasolabarri.com
bodegasolarra.es
bodegasolartia.com
bodegasolcaviana.com
bodegasole.com
bodegasoledad.com
bodegasolivares.net
bodegasoliveros.com
bodegasolmaza.com
bodegasolvena.com
bodegasondarre.es
bodegasortega.com
bodegasortubia.com
bodegasorusco.com
bodegasosca.com
bodegasotero.es
bodegaspacogarcia.com
bodegaspaezmorilla.com

bodegaspagodealmaraes.com
bodegaspagosdearaiz.com
bodegaspagosdemogar.com
bodegaspalacio.com
bodegas-palmera.com
bodegaspalomarsanchez.com
bodegaspaniza.com
bodegaspaniza.es
bodegasparcent.com
bodegaspascual.com
bodegaspasiego.com
bodegaspastrana.es
bodegaspatrocinio.com
bodegaspavoni.com
bodegaspazodevillarei.com
bodegapedromoreno1940.es
bodegaspegalaz.com
bodegaspeique.com
bodegaspelaez.es
bodegaspenafalcon.com
bodegaspenafiel.com
bodegasperal.es
bodegasperezarquero.com
bodegasperezhidalgo.es
bodegasperica.com
bodegasingon.com
bodegaspintia.com
bodegaspintocaballero.com
bodegaspinuaga.com
bodegaspiquer.com
bodegaspiqueras.es
bodegapisuerga.com
bodegasportia.com
bodegaspozanco.com
bodegaspregalado.acontia.com
bodegasprimitivocollantes.com
bodegasprotos.com
bodegaspuelles.com
bodegaspujanza.com
bodegasquirolga.com
bodegasqumran.es
bodegasramonbilbao.es
bodegasramonsaenz.com

bodegasraposo.es
bodegasrasohuete.com
bodegasraulcalvo.com
bodegas-real.com
bodegasrecuero.com
bodegasregalia.es
bodegasreymar.com
bodegasreymos.com
bodegasrezuelo.com
bodegasribeirao.com
bodegasriberadelduraton.com
bodegasriberadepelazas.com
bodegasriberalta.com
bodegasriberalta.es
bodegasribon.com
bodegasricardobenito.com
bodegasricardopalacios.es
bodegasriojanas.com
bodegasriojasantiago.com
bodegasrobalino.com
bodegasrobeal.es
bodegasrobertoblanco.es
bodegasrobles.com
bodegasroblesgonzalez.com
bodegasrochal.com
bodegasrodero.com
bodegasroman.es
bodegasromero.com
bodegasrosas.com
bodegasrubicon.com
bodegasruconia.com
bodegasruedaperez.es
bodegasruizdevinaspre.com
bodegas-saac.com
bodegassaenzdesantamaria.com
bodegassalado.com
bodegassalzilla.com
bodegassanchezrosado.com
bodegassandionisio.com
bodegassanesteban.com
bodegassangregoriomagno.com
bodegassanpablo.com
bodegassantiagoapostol.com

bodegas-santo-cristo.com
bodegasantodomingo.com
bodegassauci.es
bodegassentencia.com
bodegasserrano.com
bodegasserrano.es
bodegassers.es
bodegassierra.com
bodegassierradeguara.es
bodegassillero.com
bodegassimbolo.com
bodegassimon.com
bodegassolorca.com
bodegastabula.es
bodegastaron.com
bodegastavera.com
bodegastempore.com
bodegasteneguia.com
bodegasterradart.com
bodegasterranatura.com
bodegasterry.com
bodegasthesaurus.com
bodegastiopepe.com
bodegastirajana.com
bodegastobia.com
bodegastoribio.com
bodegastornero.com
bodegastorreduero.com
bodegastorremoron.com
bodegastorresteban.es
bodegas-torrevellisca.com
bodegastososecologica.com
bodegastradicion.com
bodegastradicion.es
bodegastraslascuestas.com
bodegastriton.es
bodegastrobat.com
bodegastrus.com
bodegasurabain.com
bodegasusabiaga.com
bodegasutiernas.com
bodegasvalbusenda.es
bodegasvalcarlos.com



bodegasvaldelana.com
bodegasvaldelomar.es
bodegasvaldesneros.com
bodegasvaldovinos.com
bodegasvalduero.com
bodegasvalentinpascual.com
bodegasvalero.com
bodegasvalhondo.com
bodegasvallelajjar.es
bodegasvalpincia.com
bodegasvaral.com
bodegasvegamar.com
bodegasveganzones.com
bodegasvelascohijos.com
bodegasven.com
bodegasverduguez.com
bodegasveredareal.com
bodegasverum.com
bodegasvetus.com
bodegasvicentesanz.com
bodegasvictorianas.com
bodegasvidal.com
bodegasvidular.es
bodegasvillagran.com
bodegasvilleza.com
bodegasvinamonte.com
bodegasvinasdelcabriel.com
bodegasvinasoro.com
bodegasvinosdeleon.es
bodegasvirgenaguila.com
bodegasvirgendeteben.es
bodegasvirgendelavega.es
bodegasvirgendelorea.com
bodegasvitulia.com
bodegasviyuela.com
bodegasvizar.es
bodegasvolver.com
bodegasvq.es
bodegasxalo.com
bodegaszamoranas.es
bodegas-zarate.com
bodegaszarraguilla.es
bodegaszarzavilla.com

bodegaszurbal.com
bodegatafuriaste.com
bodegatalleri.com
bodegateodororuizmonge.com
bodegatikalo.com
bodega-torresanmillan.com
bodegauniondelvalle.com
bodegavaleroquilez.es
bodegavalleoro.com
bodegavalmadrigal.com
bodegavetas.com
bodegavico.com
bodegavidal.com
bodegavidas.com
bodegavilladorta.com
bodegavinabayona.com
bodegavirgendelasierra.com
bodegavolcan.com
bodegavulcano.es
bodeguerosquintaesencia.com
bodeval.es
bod-fernandez.com
bogarve1915.com
bohigas.es
bonillayaceitunasdemesa.com
bonjorne.es
bosquedemasnos.es
bosquedetoledo.es
bouquetdalella.com
bouzadecarril.com
bouzadorei.com
bras1960.es
brovalero.com
brutnature.com
bsi.es
buezo.com
builgine.com
bujillo.com
bujorn.cat
burdigala.es
burgoviejo.com
butimasana.com
butroi.com

bvaloria.com
ca.cooperativagarriguella.com
caballero.es
cabovasa.com
cabrinana.es
cachazo.com
cadozos.com
calcostas.com
calderico.com
calporrera.com
cambrico.com
caminodelbosque.com
caminodelnortevinos.com
camposdedulcinea.es
camposenanzo.com
camposgoticos.es
campoviejo.es
canalscanals.com
canalscasanovas.com
canalsdomingo.com
canalsimunne.com
canalsnadal.com
canalsnubiola.com
canbonastre.com
canfeixes.com
canjustvicultors.com
canmajoral.com
canonigoareal.com
canrafolsdelscaus.com
cantaburros.es
cantauri.com
canvidalet.com
capafons-osso.com
capdebarbaria.com
capel-vinos.es
capilladelfraile.com
capitavidal.com
carabal.es
carabibas.com
caranordceller.com
carballal.net
carballaldesande.com
carceldecorpa.es

carchelo.com
carewines.com
carlania.com
carlesdelavern.com
carlosserres.com
carmelitano.com
carodorum.com
carrascalejo.com
carrascas.com
carrieldelsvilars.com
carrilcruzado.com
carviresa.com
casaantonete.com
casacastillo.es
casadeguardia.com
casadelaermita.com
casadomorales.es
casagrandelaalmuina.com
casagrandelsiurana.com
casagualda.com
casalarcon.com
casalbor.com
casalbor.es
casalcaeiro.com
casaldearman.net
casaldoscelenis.com
casalfuentes.es
casalnovos.es
casa-lo-alto.es
casalobos.es
casaloboswine.com
casamaguila.com
casamariol.com
casamoreiras.com
casapardet.blogspot.com.es
casapatau.com
casaquemada.es
casardeburbia.com
casarojo.com
casaruralterrasanta.com
casasicilia1707.es
caserioduenas.es
casonamicaela.com

castadiva.es
castaro.com
castellblanch.com
castelldage.com
castelldebiart.es
castelldelreimei.com
castelldelssorells.com
castelldencus.com
castelldor.com
castellmiquel.com
castellroig.com
castellsantantoni.com
castellsimontoliu.com
castelodemedina.com
castelodepedregosa.com
castibell.es
castillodecapmany.es
castillodecuzcurrita.com
castillodeliria.com
castillodemaluenda.com
castillodemendoza.com
castillodemonesma.com
castillodesajazarra.com
castillo-de-sajazarra.com
castillolaserna.com
castopequeno.com
castrobarona.com
castrobrey.com
castromendi.com
castroventosa.com
castuo.es
cat.cooperativagarriguella.com
catera.es
catolicoagricola.com
cavaberdie.com
cavabertha.com
cavabonet.com
cavacaldamia.blogspot.com.es
cavaduran.com
cavagabarro.cat
cavagibert.com
cavagirol.com
cavaguilera.com

cavallograula.com
cavaludens.cat
cavamartinsoler.com
cavamasolive.com
cavamiquelpons.com
cavandreu.com
cavapereabadal.com
cavapererius.com
cava-portell.com
cavareverte.com
cavarocabruna.com
cavas-arestel.com
cavasbolet.com
cavasferret.com
cavashill.com
cavashill.es
cavasmarkel.com
cavasmasachs.com
cavasmirvidal.wordpress.com
cavasrovellats.com
cavavarias.com
cavavarias.es
cava-vivesambros.com
cavesartium.com
cavescolomer.com
cavesfelixmassana.com
caydsa.es
cedoanguera.com
cellerabadia.com
cellerabadia.es
celleraibar.eu
cellerarchepages.com
cellerbalaguercabre.blogspot.com.es
cellerbarbarafors.com
cellerbartolome.com
cellerbatea.com
cellercalpla.com
cellercanpujol.com
cellercansais.com
cellercapcanes.com
cellercastellet.cat
cellercecilio.com
cellercercavins.com

cellerclua.com
cellercredo.cat
cellerdelhospital.cat
cellerdelroure.es
cellerelmoli.com
celler-escodasanahuja.com
cellerespelt.com
cellerespolla.cat
cellerfrisach.com
celler-frisach.com
cellerjoansimo.com
cellerjordidomenech.com
cellerlabollidora.com
cellerlagranada.com
cellerlamuntanya.com
cellerlaurona.com
celler-lo.com
cellermartinfaixo.com
cellermasbaste.com
cellermasbella.com
cellermasroig.com
cellermenescal.com
cellermpages.com
cellerpahi.com
cellerpardas.com
cellerpasanau.com
cellerpascual.com
cellerpinol.com
cellersabate.com
cellersatroca.com
cellersblanch.com
cellerscampdetarragona.blogspot.com.es
cellerscanblau.es
cellerscarol.com
cellerscartoixa.com
cellersdenguilla.com
cellersdescaladei.com
cellersjoandanguera.com
cellersolergibert.com
cellersroset.cat
cellerssantrafel.com
cellerstarrone.com
cellerstriada.com

celler-suigeneris.com
cellersunderground.com
cellersunio.com
cellervell.cat
cellervell.com
cepa21.com
cepado.com
cerrogallina.com
cerrolabarca.com
cervoles.com
cesarprincipe.es
cescavicent.com
cescavicentpriorat.com
chaodocouso.com
cherubino.es
cherubinoblog.com
chesteagraria.com
chinchillawine.com
chivite.com
chivite.es
chozascarrascal.es
ciadevinos.com
cialu.es
ciegodelrey.com
cienypico.com
cigarralsantamaria.com
cillardesilos.es
cincel.net
cincoalmudes.es
cinemawines.es
cinglesblaus.com
ciriacoynenez.com
ciriacoynenezimaginador.com
clos93.com
closabatlet.com
closbarenys.com
closberenguer.com
closcorvi.com
closdagon.com
closdelavall.com
closerasmus.com
closfigueras.com
closgalena.com

closlentiscus.cat
clasmogador.com
clasmontblanc.com
claspns.com
clatdelessoleres.com
cocafito.com
codorniu.com
codorniu.es
colet.cat
coloniasdegaleon.com
comalats.com
comalats.net
comaroma.net
comenge.com
comvins.com
concejobodegas.com
condadodeojancos.com
condedecaralt.com
condedemontornes.es
condesdealbarei.com
consejodelaalta.com
consejodelaalta.es
constantia.es
contrerasruiz.com
conventooreja.net
coop-barbera.com
coopbot.com
coop-cabeza.com
cooperativacastalla.com
cooperativacristodesantaana.es
cooperativadelcamposanisidro.com
cooperativadesotes.com
cooperativaelprogreso.com
cooperativafrontera.com
cooperativagarriguella.com
cooperivalamuella.com
cooperivalaunion.com
cooperativamentrida.es
cooperivarua.com
cooperivasanbernabe.com
cooperivasanisidro.com
cooperivasanroque.com
cooperivasisante.com

cooperativauci.com
coopetrer.es
coopfuendejalon.com
coopgandesa.com
cooplosremedios.com
coopontinyent.com
coopsalelles.com
coopsalomo.com
coopsanisidroentrin.es
coop-santvicent.com
coopulldemolins.com
coopvillena.com
coopvitolocal.com
copaboca.com
copvilar.com
corcovo.com
cornudella.net
cortijodejara.es
cortijosaguilares.com
cosechaextremadura.com
cosecherosdesoutomaior.com
cosmic.cat
costaval.com
costersdelpriorat.com
costersdelros.com
costersdelsiurana.com
costersio.com
cotodegomariz.com
covaserodia.com
covides.com
covijerez.es
covila.es
covilalba.com
covilda.com
covinca.es
covisan.net
covitoro.com
crapulawines.com
craterbodegas.com
cregoemonaguillo.com
criadoresderioja.es
crinroja.es
crismona.com

cruzdealba.es
cruzvalle.com
cuatorrayas.es
cuatorrayas.org
cuatrovides.com
cuestaroa.es
cuevassantoyo.com
cuevassantoyo.vinos.tienda
cumbresdeabona.es
cunadereyes.es
cusoberga.com
cvalvanera.com
cvne.com
cvne.es
cvsolterra.com
dagon.es
dagonbodegas.es
dainsua.es
dalmau-hnos.com
danielbelda.com
dardell.es
davidmoreno.es
dcueva.com
dealberto.com
decantia.com
dehesadelcarrizal.com
dehesadelosllanos.com
dehesadeluna.com
dehesalagranja.com
dehesalosprietos.com
dehozsualdea.com
delgadozuleta.com
demenciawine.com
demuller.es
descalzosviejos.com
descregut.com
destileriasybodegascascallana.es
devinssi.com
dgvicultors.com
dibodegas.es
diez-caballero.es
diezllorente.com
diezmerito.com

diezsiglos.es
dimobe.es
dinastiavivanco.com
dinfante.com
discosta.com
divinaproporcionbodegas.es
domaineslupier.com
dombardo.com
domecqbodegas.com
dominiobasconcillos.com
dominioberzal.com
dominiobuenavista.com
dominiodeatauta.com
dominiodeberzal.com
dominiodecair.com
dominiodelaguila.com
dominiodelavega.com
dominiodelurogallo.com
dominiodenobleza.com
dominiodepunctum.com
dominiodetares.com
dominiodobibi.com
dominiodostares.com
dominiolasierpe.com
dominiomaestrazgo.com
dominiomaestrazgo.es
dominiomarcelino.com
dominiopingus.com
donalbino.com
donbernardino.com
donienegorrondona.com
donjacob.es
donjuandelaguila.es
donolegario.com
doriasbaixas.com
dosterras.com
dsgvineyards.es
duerowine.net
duquedemontebello.com
ebanovinedosybodegas.com
ebesa.org
eccociwine.com
eclemente.es



ecosetrill.com
ecovitres.com
edetaria.com
egomei.es
eguren.com
egurenugarte.com
eguzkialdetxakolina.blogspot.com.es
eladiopineiro.es
elarcadenoesanasensio.es
elbruguervell.com
elcoto.com
elgrifo.com
elgrillo.net
elguiso.es
elmasferrer.com
elnispero.es
eloilorenzo.es
elregajal.es
elternero.com
elvillar.com
elvinoprodigo.com
elviwines.com
elxamfra.com
emendis.es
emilio-lustau.com
emiliomoro.com
emina.es
empordalia.com
enanzo.com
enate.es
encomiadadecervera.com
enkvistwines.com
ennak.com
enologicawamba.es
entrevinosypagos.com
envidia.es
epifaniorivera.com
equiponavazos.com
escocesvolante.es
es-fangar.com
espadaojeda.com
espectaclevins.com
espeltvicultors.com

estanciapietra.com
estebanmartin.com
esteldargent.com
estradaalacio.com
esverger.es
etim.cat
etim.es
etxetxo.com
evohegarnacha.com
exopto.net
explotacionessantiago.com
familiabelasco.com
familiaescudero.com
familiamartinezbujanda.com
fanbar.es
fargas-fargas.com
faustinatorivero.com
fefinanes.com
felixsolis.com
felixsolisavantis.com
fernandezdearcaya.com
fernandodecastilla.com
ferratus.es
ferreicatasus.com
ferrerbobet.com
ferretguasch.com
ficariavins.net
fincaalbret.com
finca-allende.com
fincaanfora.com
fincaardal.com
fincabarronte.com
fincabarronte.weebly.com
fincacanestella.com
fincacollado.com
fincacondal.com
fincaconstancia.es
fincadelarica.com
fincadelosarandinos.com
fincadelregajo.es
fincaelretamar.com
fincalablanca.es
fincalaestacada.com

fincalagramanosa.com
fincaloranque.com
fincalosalijares.com
fincalosaljibes.com
fincamanzanos.com
fincamoncloa.com
fincanueva.com
fincaonegar.com
fincaparera.com
fincarionegro.com
fincasanblas.com
fincasandoval.com
fincateira.com
fincatobella.com
fincavaldeguinea.net
fincavalldosera.es
fincaviladellops.com
fincavinoa.com
florentinomartinez.com
florislegere.com
fonpinet.com
fontalbacapote.es
franchete.com
francoespanolas.com
francoislurton.es
francosanchez.es
fraygerman.com
freixedas.com
freixenet.com
freixenet.es
friadelval.com
frontos.es
f-schatz.com
fuentegalana.com
fuentenarro.com
fuidio.com
furorwines.com
gaineta.com
galandembrilla.com
galanportero.es
galmesiferrer.com
galmesiribot.com
garciacarrion.es

garciadeolano.com
garciarevalo.com
gargalo.es
garva.com
gelama.cat
genesi.cat
geniumceller.com
gilberzal.com
gilfamily.es
giluna.es
girodelgorner.com
giroribot.es
godelia.es
godelleta.com
godeval.com
gomara.com
gomezaguirre.com
gomezcruzado.com
gomezdesegura.com
gomezrial.es
gontes.com
gonzalezbyass.com
gonzalezbyass.es
gonzalezpalacios.com
gordonzello.com
goyavinos.com
gramona.com
granbierzo.com
grandesbodegas.com
grandespagos.com
grandesvinos.com
granfeudo.com
granfucares.com
granrecosind.com
granviu.com
gratavinum.com
gratiaswines.com
graudoria.com
grauvell.cat
gregoriomartinez.com
grupoartevino.com
grupobaco.com
grupobodegasolarra.com

grupoestevez.es
grupofaustino.es
grupogarvey.com
grupoliveda.com
grupomatarromera.com
grupopesquera.com
gruporioja.es
gruporuberte.com
grupoyllera.com
guelbenzu.com
guelbenzu.es
guimaro.es
gureahaleginak.com
gurpegui.com
gurpegui.es
guspi.com
gutierrezcolosia.com
gutierrez-colosia.com
guzmanaldazabal.com
habarcelo.com
habarcelo.es
haciendaalbae.com
haciendadelcarche.com
haciendaespino.com
haciendagrimon.com
haciendamolleda.com
haciendamonasterio.com
haciendas-espana.com
haciendasolano.com
haciendavillarta.com
haciendavillarta.es
herascordon.com
heredadaduna.com
heredadbaroja.com
heredadlinares.com
heredadluzuriaga.com
heredadmartinezcastillo.es
heredad-pallares.com
heredadpanguasodupe.com
heredadugarte.com
heretatdetaverners.com
heretatferrerdelavall.es.tl
heretatnavas.com

herrerobodega.com
hesvera.es
hidalgo.com
hijosderufinoiglesias.com
himafesa.es
hispanobodegas.com
honoriorubio.com
hornillosballesteros.es
hotelmastinell.com
hoyadecadenas.es
hoyadelcastillo.com
huertadealbala.com
ibizkus.com
ignacionmarin.com
ijalba.com
ilurce.com
iniza.net
irache.com
isasi.es
ishern.com
izadi.com
janebaques.cat
janesantacana.com
janeventura.com
jaumandreu.blogspot.com.es
jaumellopartalemany.com
jaumemesquida.com
jeanleon.com
jesusdelnero.es
jgc.es
jimenezlandi.com
jmiqueljane.com
joandelacasa.com
joansarda.com
joaquinrebollo.com
jordandeasso.com
jorgeordonez.es
josecarlosquintasperez.es
josepariente.com
josepforaster.com
jsviticultor.com
jtrias.com
juancarlossancha.com

juanramirezbodegas.com
juliabernet.com
juliansoler.com
juvecamps.com
kairosvino.com
kiriosdeadrada.com
kronosseleccion.com
labodegadealboloduy.com
labodegadelasestrellas.com
labodegadepinoso.com
labodegadeserapio.com
laboella.com
labotera.com
lacana.es
lacasadelasvides.com
lacavaaragonesa.es
lacavadelbierzo.com
lacepaalta.com
lacooperativadealcublas.com
laderasdemontejurra.com
ladrondelunas.com
laestacada.es
lafou.net
lafuina.com
lagarblanco.es
lagardacondesa.com
lagardebesada.com
lagardecervera.com
lagardecosta.com
lagardepintos.com
lagardesabariz.com
lagardofreixo.com
lagardomerens.com
lagarisilla.es
lageria.com
lagitana.es
lagravera.com
lagrima-negra.com
laguita.com
lagunamadrid.com
lahigueramayor.com
laisleta.es
lajaraba.com

lalegua.com
lamejorada.es
lamelonera.com
lapamelita.com
laperladelpriorat.com
laraseleccion.com
lardepaula.com
lardevinas.com
laremediadora.com
lascaraballas.com
lasetera.es
lasmoradasdesanmartin.es
lasorcas.es
latidosdevino.com
latorreagrovinicola.com
latue.com
latue.es
lavicolamestre.com
lavinyeta.es
laxarmada.com
lazarodelcorsobodegas.blogspot.com.es
lazaruswine.com
lazotur.com
legaris.com
legaris.es
lembolic.com
lescousins.es
leyendadelparamo.es
lezaun.com
lezcano-lacalle.com
libaydeleite.com
liberalia.es
linajegarsea.com
llagaresvalveran.com
llagarherminio.com
llanorubio.es
llicorellavins.com
lliria.coop
llopart.com
llopart.es
loculto.com
lodmc.info
loess.es

londono.es
lonecesario.es
lopezcristobal.com
lopezdeheredia.com
lopezpanach.com
lordina.net
lorigancava.com
losadavinosdefinca.com
losalbertos.com
los-barrancos.es
losbermejós.com
losmarcosbodegas.com
losraigones.com
lostrovadores.com
loxarel.com
luberry.com
luculo.es
luguswines.com
luisalegre.com
luiscanas.com
lunabeberide.es
lusco.es
lustau.es
lynus.es
maciabatle.com
madiaz.com
maestrosierra.com
maetierra.com
magalartelezamatxakolina.com
magrinyacalaf.com
mahara.es
maimona.com
maiordemendoza.com
maisulan.com
maiusvicultors.com
malandrinwines.com
maldivinas.es
malondeechaide.com
malondro.es
malumbres.com
mambrilla.com
mamertodelavara.es
mancuso.es

manuelmanzaneque.com
marcelinoserrano.com
marcoabella.com
mardeenvero.es
mardefrades.com
mardefrades.es
mariabargiela.com
mariacasanovas.com
mariaoliverporti.com
mariarigolordi.com
mariscal.es
marquesdealella.com
marquesdearviza.es
marquesdebembibre.com
marquesdecaceres.com
marquesdeirun.es
marquesdelatrio.com
marquesdelciego.com
marquesdemontecierzo.com
marquesdemurrieta.com
marquesdereinosa.com
marquesderiscal.com
marquesdeteran.com
marquesdetomares.com
marquesdeulia.com
marquesdevaldueza.com
marquesdevargas.com
marquesdevelilla.com
marquesdevillalua.com
marquesdevitoria.com
marquesdevizhoja.com
marquesrigol.com
marsilea.es
martinberdugo.com
martincodax.com
martinezbujanda.com
martinezlacuesta.com
martinezlaorden.com
martinezserantes.com
martinsancho.com
martiserda.com
martivilli.com
martue.com



masaveubodegas.com
masbertran.com
masblanchijove.com
mascancolome.com
mascandi.com
mascaro.es
mascodina.com
mascomtal.com
masdecosme.com
masdelabasserola.cat
masdelabundancia.com
masdelacasadora.com
masdelarealdesella.es
masdelesvinyes.com
masdelhereu.com
masdenblei.com
masdencanonge.com
masdengil.com
masderander.com
masdesantiscle.com
masdeuicampos.blogspot.com.es
masdoix.com
masestela.com
maset.com
masetplana.com
masgarciamuret.com
masgarrian.dopriorat.com
masgoma1724.com
masia.i-commweb.net
masiaduch.com
masiaelmas.es
masiaserra.com
masiavila-rasa.com
masigneus.com
masjust.com
maslamola.com
masllunes.es
masmartinet.com
masoller.es
masperinet.com
maspolit.com
maspujado.com
masrodo.com

massananoya.com
massinen.com
mastinell.com
masvicens.com
matacoloma.com
matamangos.com
matarredonda.es
matiasitorres.com
mayogarcia.com
mayol.eu
mayordemigueloa.com
megallum.com
melgarajo.es
melispriorat.com
melquior.com
menade.es
mendezmoya.com
mengoba.com
meoriga.com
merencio.es
merumpriorati.com
mesquidamora.com
mestres.es
microbodegabio.blogspot.com.es
miguelmerino.com
miqueloliver.com
mirosalvat.com
mitarte.com
mmonistrol.com
molidelscapellans.com
monasteriodecorias.es
monjardin.es
monsataperdigon.es
montalvowilmot.com
montaudesadurni.com
monteabellon.com
monteaixa.com
monteaman.com
montebrione.cat
montegaredo.com
montelaguna.es
montelareina.es
montelciego.com

monteodina.com
montepedroso.es
monteplata.es
monterebro.com
monterruiz.com
montevannos.es
montevirgen.com
montferrant.com
mont-marcal.com
mont-reaga.com
montrubi.com
morachacon.com
morgadio.com
morlanda.com
mostosa.com
movialsa.es
munozymazon.es
murviedro.es
nadal.com
navaltallar.com
navarrolopez.com
navarrsotillo.com
naveran.com
navisa.es
nekeas.com
nestareseguzabal.com
nilwines.com
nita.cat
nivarius.com
noguerals.com
nondum.es
nosriqueza.com
numanthia.com
nunci.com
nuviana.com
obac.es
obanca.com
obarreiro.com
obergo.es
ojuelwine.com
olartia.com
olivardots.com
olivareros.com

olivebatllori.es
olivellagalimany.com
olivellaibonet.es
olivera.org
oliverconti.com
olivermoragues.com
olivier-riviere.com
ollerdelmas.com
ondalan.es
ontanon.es
operello.com
oreidecampoverde.es
organicsignaturewines.es
oriolrossell.com
oroaragon.com
orodecastilla.com
ortegaezquerro.com
ortovins.com
orvalaiz.es
osborne.es
osoti.es
ossianvinos.com
ostatu.com
otazu.com
oventosela.com
ovidiogarcia.com
oxinbaltza.com
pablopadin.com
pacolola.com
pagesentrena.com
pagoayles.com
pagobolote.com
pagocalzadilla.com
pagocasadelblanco.com
pagocasagran.com
pagodecarraovejas.com
pagodecirsus.com
pagodeguzque.com
pagodeina.com
pagodelarrainzar.com
pagodelarrea.com
pagodelasencomiendas.es
pagodeloscapellanes.com

pagodelvicario.com
pagodetharsys.com
pagoflorentino.com
pagolosbalancines.com
pagopenafiel.com
pagosdefamilia.es
pagosdegalir.com
pagosdehibera.com
pagosdeleza.com
pagosdelmoncayo.com
pagosdelrey.com
pagosdematanegra.es
pagosdequintana.com
palaciodeazcona.com
palaciodebornos.com
palaciodelavega.com
palaciodelerma.com
palaciodemuruzabal.jimdo.com
palaciodevillachica.com
palaciosduque.com
palaciosvinoteca.com
paramoarroyo.com
paramodecorcos.es
paramodefresnedo.com
paramodeguzman.es
parato.es
pardevalles.es
parentegarcia.com
paresbalta.com
parmipriorat.com
partal-vinos.com
partidacreus.com
parxet.es
pascona.com
pascualberganzo.com
pascuallarrieta.com
pastordiaz.com
paternina.com
payva.es
pazoacapitana.com
pazoasbarreiras.es
pazobaion.com
pazobarrantes.com

pazocasanova.com
pazodastapias.com
pazodegalegos.com
pazodemonterrey.es
pazodesenorans.com
pazodevieite.es
pazodomar.com
pazopondal.com
pazoquinteirodacruz.es
pazosdelrey.com
pedralonga.es
pedroescuderoviticultor.com
pedroromero.es
pelio.es
penadasdonas.com
penadelvalle.es
penalarosa.com
penascal.es
pereguardiola.com
perelada.com
pereseda.com
pereventura.com
perez-baigorri.com
perezbarquero.com
perezcarames.es
perezpascuas.com
pernodricarbodegas.com
pernod-ricard-espana.com
picocadro.com
piemonte.com
pirola.com
pinnafidelis.com
pinnafidelisonline.com
pinord.com
pinord.es
pinyolsabate.com
piodelramo.com
pittacum.com
planasalbareda.com
pomardeburos.es
pombares.com
pontedaboga.es
portaldelmontsant.com



portaldelpriorat.com
pradaatope.es
pradodeolmedo.com
pradorey.com
pradorey.es
preceptwine.com
prediodevascarlon.com
presasocampo.com
prietopicudo.com
primitivoquiles.com
primumbobal.com
principedeviana.com
priordepanton.com
propiedadgrial.net
puentedelea.com
puentederus.com
puigmunts.com
puigpriorat.com
puig-romeu.com
puragarnacha.com
quadernavia.com
qubel.com
quinola.es
quintaavenida.es
quintacouselo.com
quintadapeza.es
quintadelobispo.com
quintadobuble.com
quintamazuela.es
quintamilu.com
quintaquietud.com
quitapenas.es
rafabernabe.com
rafaelcambra.es
rafaelpalacios.com
rafaelreverte.es
raimat.com
raizyparamodeguzman.es
ramirezdelapiscina.com
ramirowinecellar.com
ramonroqueta.com
raventos.com
rayalaehijos.com

realcortijo.com
realdivisa.com
realjunta.com
realrubio.es
recaredo.es
recoveco.es
reginaviarum.es
regoa.es
reinadecastilla.es
reinodelosmallos.es
rejadorada.com
remelluri.com
remelluri.es
remesal.es
remirezdeganuza.com
rendemasdeu.cat
renebarbier.com
resalte.com
restaurantebodegasmonclus.com
rexachbaques.com
ribasdelcua.com
riberadelguadiana.eu
riberadelornia.es
ricardana.com
ricardobenito.com
rimarts.net
rioja1808.com
riojalta.com
riojanadealcanadre.com
riojavega.com
riolanc.com
rivalia.com
rmrioja.com
roda.es
rodriguezdevera.com
rodriguezsanzo.com
rogergoulart.com
roigparals.cat
romagosatorne.com
romaila.es
romale.com
romate.com
romeralvinicola.com



rompesedas.com
ronadelles.com
ronseldosil.com
roquers.com
roquesan.es
rosamariatorres.com
roselliformosa.com
rosellmir.com
rosmas.com
rotllantorra.com
roura.es
rovirabaques.com
rubiejo.com
rudeles.com
ruiztorres.com
sacovaibiza.com
saforana.com
saiÑas.com
salacasanovas.com
salcis.es
salvadorpoveda.com
salvadorpoveda.es
salvueros.com
sampayolo.com
san-alejandro.com
sanantonioabad.es
sanantoniobodegas.com
sancebrin.com
sanchezmuliterno.com
sanchezzahonero.com
sancius.com
sanclodiobodega.com
sanero.com
sangenisivaque.com
sanmartindeabalos.com
sanstrave.com
santacatalina.es
santalba.com
santamariadenogueira.com
santiagoroma.com
santjosepwines.com
santodomingodeguzman.es
santpatrici.com

sanvalero.com
sanvicente.com
saodelcoster.com
satninojesus.com
sauvella.com
sebastiapastor.com
sebiran.es
sedellavinos.com
seguraviudas.es
seldaiz.com
senderoroyal.com
senoriocincovillas.com
senoriodearana.com
senoriodearnedo.com
senoriodeastobiza.com
senoriodecanada.com
senoriodelaestrella.es
senoriodelasvinas.com
senoriodelosarcos.es
senoriodejava.es
senoriodenevada.es
senoriodeurarte.com
senoriodevalazul.es
senoriodevarei.com
senoriodevillafames.com
serradalmos.com
serradecavalls.com
serradellaberia.com
serraladademarina.com
serratdemontsoriu.wordpress.com
serres.net
sialba.es
sierracantabria.com
sietepecados.org
sietesoles.es
silvanogarcia.com
silvanogarcia.es
simodepalau.com
simodepalau.es
simodepalau.net
sinforianobodegas.com
sistemavinari.com
slibrares.com

sobreno.com
solabal.es
soladares.wordpress.com
solaguen.com
solaraventos.com
solardeayala.com
solardesamaniego.com
solardeurbezo.es
solarmsancho.com
solarviejo.com
solerabodegas.com
somanilla.es
sonartigues.com
sonbordils.es
sonprim.com
sonpuig.com
sonramon.com
sonsierra.com
sonvives.com
sotaelsangels.com
spiracp.es
srubios.com
ssobral.net
stratvs.com
suertesdelmarques.com
sumarroca.es
suriol.com
tabaibal.es
tagonius.com
tajinaste.net
talaiberri.com
tallerdebaco.com
tamaral.com
tampesta.com
tanajara.com
tandem.es
taninia.com
tapiasmarinhan.com
tardencuba.es
tavasa.es
tejeruelas.es
telmorodriguez.com
tenoiragayoso.com

teofiloreyes.com
terciadeulea.com
terminodemiranda.es
ternobradordevi.com
terradeasorei.com
terradocastelo.com
terraduro.com
terraesteban.com
terraminei.com
terramoll.es
terranostravinos.com
terraremota.com
terrasdecompostela.com
terrasdelantano.com
terrasdocigarron.com
terrasgauda.com
terror.net
terresdevidalba.com
terroir-al-limit.com
the-haciendas.com
thewinelove.com
tiannanegre.com
tierradefrontos.com
tierrantigua.com
tierrasdeaponte.com
tierrasdemollina.net
tierrayvino.com
tilenus.com
tintoarroyo.com
tintobel.com
tintocarme.es
tintofiguero.com
tintolerma.com
tintoralba.com
tionio.com
tobelos.com
tollodouro.com
tomadadecastro.com
tomascusine.com
tomaspostigo.es
torello.es
toroalbala.com
torrecastillo.com

torredelveguer.com
torredeona.com
torrederos.com
torremaciel.com
torremilanos.com
torreoria.es
torres.es
torresfiloso.org
traslanzas.com
tritium.es
trossosdelpriorat.com
txakoliaizpurua.com
txakoliameztoi.com
txakolibizkaia.com
txakolidealava.com
txakoligorkaizagirre.blogspot.com.es
txakolilarringan.wordpress.com
txakolina-k5.com
txakolirezabal.com
txakoli-uriarte.com
txominetxaniz.com
umesufan3.com
uncastellum.com
urbinavinos.com
urium.es
urkitxakolina.com
uvaguilera.com
uveste.es
valbornedo.com
valdamor.es
valdeaguila.com
valdehermoso.com
valdelares.com
valdelosfrailes.es
valdemar.es
valdemazon.com
valdemonjas.es
valdepablo.com
valderiz.com
valdesil.com
valdesouto.com
valdevid.es
valdevinas.es

valdrinal.com
valdubon.com
valdumia.com
valenciso.com
valldebaldomar.com
valldolina.com
vallebravo.es
vallegarcia.com
vallemayor.com
vallemonzon.com
vallesdebenavente.org
valleyglesias.com
vallformosa.es
vallllach.com
vallobera.com
valonga.com
valquejigoso.com
valsan1831.com
valsangiacomo.es
valsardo.com
valsarte.com
valserrano.com
valsotillo.com
valtravieso.com
vay.es
vdecamarena.com
vegaaixala.com
vegaclara.com
vegadelcastillo.com
vegaderibes.com
vegadeyuco.es
vegadeyuso.com
vegaenix.com
vegaesteban.com
vegalfaro.com
vegamoragona.com
vegasauco.com
vega-sicilia.com
vegatolosa.com
vegaval.com
vegazar.com
veigaserantes.com
vendrellrived.com

ventadaubert.ch
ventadelpuerto.com
ventalavega.com
ventomareiro.com
venuslauniversal.com
veradeestenas.es
vermutyzaguirre.com
veronicasalgado.es
vertijana.es
viñadoresartesanos.com
viaarxentea.com
vialcon.net
vianguera.com
viaromana.es
vicentegandia.es
vidyespiga.es
viejocoso.com
vihucas.com
vilacorona.cat
vilarnau.es
vilarvin.com
vilerma.com
villachica.es
villacreces.com
villadelcorpus.es
villarrica.es
villavid.com
villeza.com
vinaaliaga.com
vinaalmirante.com
vinaane.com
vinaberneda.com
vinablanca.com
vinacampanero.com
vinacerron.com
vinaelcardon.com
vinaelena.com
vinaemiliano.com
vinaemureri.com
vinafarnadas.es
vinalaguardia.es
vinamagana.com
vina-mayor.es

vinamein.com
vinamezquita.blogspot.com.es
vinanora.com
vinaobdulia.blogspot.com
vinaoliva.com
vinaredo.com
vinaruda.com
vinasalceda.com
vinasastre.com
vinasdelportillo.com
vinasdelvero.es
vinasobreira.com
vinateriayanez.com
vinatigo.com
vinatorreblanca.com
vinavilano.com
vinazagros.com
vinedosdealfaro.com
vinedosdelsoto.com
vinedosruizjimenez.es
vinedossingulares.com
vinessens.com
vinexsa.com
vineyard.es
vinicachivana.com
vinicoladearganda.com
vinicoladecastilla.com
vinicoladelcondado.com
vinicoladelnordest.es
vinicoladelpriorat.com
vinicolaguadiana.com
vinicolareal.com
vinicolarequenense.com
vinicolatomelloso.com
vinicolavillarrobledo.com
vinigalicia.es
vinispana.com
vinoarar.com
vinobelardi.com
vinobispo.com
vinobronce.com
vinocontiempo.com
vinolea.es

vinoleapurisima.com
vinolibranza.com
vinopuchero.blogspot.com.es
vinoribeiro.com
vinos.anecoop.com
vinosalborea.com
vinosalex.com
vinosaltoturia.com
vinosartajona.com
vinosbarco.com
vinosbernabeudeyeste.es
vinosbureo.com
vinoscarbballal.com
vinosconrad.com
vinosculebron.com
vinosdealguenya.es
vinosdealiste.com
vinosdeanora.es
vinosdearganza.com
vinosdechiclana.com
vinosdehoz.com
vinosdelbierzo.com
vinosdeltoboso.com
vinosdemadrid.es
vinosderauda.com
vinosdivertidos.es
vinosdoribeiro.es
vinoselcarrascal.com
vinosferrer.com
vinosfiel.es
vinosflorentinoperez.es
vinosfragata.es
vinosga.com
vinosguerra.com
vinoshabla.com
vinosiberian.com
vinosjeromin.com
vinosjuliana.com
vinosladama.com
vinoslaveguilla.com
vinoslazorra.es
vinos-menchero.com
vinosmontillamoriles.com

vinosnavarra.com
vinosomoza.com
vinospalarea.com
vinospasiego.com
vinospatio.blogspot.com.es
vinosperdiguer.com
vinospinol.com
vinosrenalte.com
vinos-roa-rauda.com
vinosrubenramos.com
vinosrubicon.com
vinosruchel.com
vinossanz.com
vinostomillar.es
vinosvaltuille.com
vinosveleta.com
vinosvetusta.com
vinosvillar.com
vinosybodegas.com
vinosylicorespicos.es
vinoteatinos.com
vinotierraranda.com
vinotoro.com
vinscancoleto.com
vinscollideroses.com
vinsdeforesta.cat
vinsdelcomtat.com
vinsdelsud.wix.com
vinsdeltros.com
vinsdepedra.es
vinsdetaller.com
vinselep.com
vinsgrau.com
vinsguillamet.com
vinsjaumedepuntiro.com
vinsjosepvicens.com
vinslaconreria.com
vinsmiquelgelabert.net
vinsnadal.com
vinspadro.com
vinssingulars.com
vinstonigelabert.com
vintae.com

vinyaescude.com
vynaivo.wordpress.com
vynalhereu.com
vinyanatura.com
vinyasacudia.com
vinyataujana.es
vinyesdelalba.com
vinyesdelsaspres.cat
vinyesdengabriel.com
vinyesdomenech.com
vinyesmortitx.com
viñaherminia.es
virgendelaasuncion.com
virgendelasnieves.com
virxendosremedios.es
viticultoresdebarros.com
vitisnatura.com
vitisterrarum.es
vitivinos.com
vivancoculturadevino.es
vivanza.es
viverdecelleristes.concadebarbera.cat
vivirelvino.com
vizcarra.es
vocarraje.es
vsantamarina.com
web.bodegasmunana.com
williams-humbert.com
winepalace.es
www.jilaba.es
xamos.es
xdt.es
xenysel.com
ximenezspinola.com
xlsebio.es
xn--bodegacastillodepearanda-ilc.com
xn--bodegasyviedosluismariscal-qrc.es
xn--marquesdecasapardias-l7b.com
yuntero.com
zacariasbivian.com
zanata.net
zarate.es
zearra.com

zifar.com
zuazogaston.com
zuger.com