



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

CAMPUS D'ALCOI

Análisis del Impacto socioeconómico del Covid-19 mediante técnicas estadísticas

MEMORIA PRESENTADA POR:

Daniel Barandiarán Torres

GRADO DE ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

Convocatoria de defensa: Septiembre 2020

Resumen:

El actual virus propagado a nivel mundial, Covid-19, tiene efectos negativos que afectan a toda la sociedad por la interconexión entre países, y la influencia de los Estados en la población que sujetan. Al ser un tema de actualidad, y que se encuentra en curso, es interesante conocer qué consecuencias graves está acarreado.

El presente proyecto tiene por objetivo realizar un análisis estadístico para averiguar qué factores se ven afectados por el virus, así como analizar el impacto que ocasiona a nivel macroeconómico y social. Se pretende dar conclusiones claras y fiables sobre la evolución del virus y su influencia.

Para llevar a cabo el estudio se aplicarán técnicas estadísticas diversas. Se buscarán bases de datos e información referente para contextualizar y poder entender dichos datos. A partir de distintas bases de datos y fuentes seleccionadas se realizará el análisis de los datos.

Palabras clave: Covid-19; base de datos; técnicas estadísticas; análisis estadístico; Árboles de Decisión CART; programa estadístico R.

Summary:

The current globally spread virus, Covid-19, has negative effects that affect the whole society because of the interconnection between countries, and the influence of States on the population they support. As it is a current issue, and one that is underway, it is interesting to know what serious consequences it is bringing about.

This project aims to carry out a statistical analysis to find out which factors are affected by the virus, as well as to analyse the impact it causes at a macroeconomic and social level. The aim is to give clear and reliable conclusions on the evolution of the virus and its influence.

Various statistical techniques will be applied to carry out the study. Databases and reference information will be sought in order to contextualize and understand these data. The analysis of the data will be carried out from different databases and selected sources.

Keywords: Covid-19; database; statistical techniques; statistical analysis; CART Decision Trees; statistical programme R.

Índice de Contenidos

1.	Introducción y objetivos.....	7
1.1.	Motivación	8
1.2.	Hipótesis.....	10
1.3.	Objetivos	11
1.4.	Estructura del proyecto.....	12
2.	Metodología y Técnicas.....	13
2.1.	Antecedentes	14
2.2.	Metodología	18
2.3.	Técnicas	20
2.3.1.	Técnicas aplicadas con “R”	20
2.3.2.	Herramienta estadística CART con “R”	21
3.	Cálculos y Resultados	22
3.1.	Descripción de variables.....	23
3.2.	Análisis univariante	26
3.3.	Análisis bivariante	36
3.4.	Análisis multivariante.....	51
3.4.1.	Modelos de la variable dependiente: número de fallecidos por Covid-19.....	52
3.4.2.	Modelos de la variable dependiente: número de contagiados por Covid-19.....	59
3.4.3.	Modelos según nivel económico.....	65
4.	Conclusiones y futuras líneas de investigación	72
4.1.	Conclusiones.....	73
4.2.	Futuras líneas de investigación	76
5.	Bibliografía	77
6.	Anexo.....	82
	Base de Datos.....	83
	Gráficos no incluidos	87

Índice de Ilustraciones

Ilustración 1. Gráfica comparativa de virus, parásitos y bacterias. Fuente: informationisbeautiful.net.....	14
Ilustración 2. Impacto de la gripe en España. Fuente: vacunasaep.org.....	15
Ilustración 3. Gráfico de sectores de la variable CONFI.....	27
Ilustración 4. Gráfico de barras de la variable JABON.....	28
Ilustración 5. Gráfico de caja y bigotes de la variable CAMAS.....	29
Ilustración 6. Gráfico de caja y bigotes de la variable MORTA.....	30
Ilustración 7 e Ilustración 8. Gráficos de caja y bigotes de las variables MED y GASTO.....	30
Ilustración 9 e Ilustración 10. Gráficos de caja y bigotes de las variables URB y DENSI.....	31
Ilustración 11 e Ilustración 12. Gráficos de caja y bigotes de las variables AGLO y DESNUT.....	32
Ilustración 13. Histograma de la variable MIG.....	32
Ilustración 14. Histograma de la variable DIAS.....	33
Ilustración 15. Histograma de la variable PIB.....	34
Ilustración 16. Histograma de la variable PARO.....	34
Ilustración 17 e Ilustración 18. Gráficos de caja y bigotes de las variables FALLE y INFECT.....	35
Ilustración 19. Gráfico XY de las variables CAMAS e INFECT.....	38
Ilustración 20. Gráfico XY de las variables MIG e INFECT.....	39
Ilustración 21. Gráfico XY de las variables GASTO y FALLE.....	40
Ilustración 22 e Ilustración 23. Gráficos XY de la variable DESNUT respecto a las variables PIB y PARO.....	40
Ilustración 24 e Ilustración 25. Gráficos de caja y bigotes de las variables FALLE e INFECT en grupos de la variable JABON.....	41
Ilustración 26 e Ilustración 27. Gráficos de caja y bigotes de las variables FALLE e INFECT en grupos de la variable CONFI.....	42
Ilustración 28 e Ilustración 29. Gráficos de caja y bigotes de las variables PARO e PIB en grupos de la variable CONFI.....	43
Ilustración 30. Árbol de decisión del modelo 1 de la variable dependiente FALLE.....	52
Ilustración 31. Árbol de decisión del modelo 2 de la variable dependiente FALLE.....	54
Ilustración 32. Árbol de decisión del modelo 3 de la variable dependiente FALLE.....	56
Ilustración 33. Árbol de decisión del modelo 1 de la variable dependiente INFECT.....	60
Ilustración 34. Árbol de decisión del modelo 2 de la variable dependiente INFECT.....	61
Ilustración 35. Árbol de decisión del modelo 3 de la variable dependiente INFECT.....	63
Ilustración 36. Árbol de decisión del modelo de la variable dependiente INFECT en los países con ingresos altos.....	65
Ilustración 37. Árbol de decisión del modelo de la variable dependiente FALLE en los países con ingresos altos.....	66
Ilustración 38. Árbol de decisión del modelo de la variable dependiente INFECT en los países con ingresos medios.....	67
Ilustración 39. Árbol de decisión del modelo de la variable dependiente FALLE en los países con ingresos medios.....	68
Ilustración 40. Árbol de decisión del modelo de la variable dependiente INFECT en los países con ingresos bajos.....	69
Ilustración 41. Árbol de decisión del modelo de la variable dependiente FALLE en los países con ingresos bajos.....	70

Índice de Tablas

Tabla 1. Comparativa virus y bacterias pandémicas.	16
Tabla 2. Resumen estadístico de las variables.	26
Tabla 3. Variable medida de confinamiento.	27
Tabla 4. Variable JABON (% de población con instalaciones básicas de agua e higiene).	28
Tabla 5. Matriz de correlaciones de variables cuantitativas.	37
Tabla 6. Países según nivel económico.	43
Tabla 7. Matriz de correlaciones de variables cuantitativas de los países con ingresos bajos...	45
Tabla 8. Matriz de correlaciones de variables cuantitativas de los países con ingresos medios.	47
Tabla 9. Matriz de correlaciones de variables cuantitativas de los países con ingresos altos....	49

1. Introducción y objetivos

Esta primera sección se divide en cuatro apartados: motivación, donde se explica las motivaciones personales, así como los conocimientos aplicados y una breve explicación del virus; la hipótesis, donde se enuncia lo que se pretende conseguir con el proyecto; los objetivos, donde se enuncia el objetivo general y específicos para llevar a cabo el proyecto; y la estructura, donde se explica brevemente la estructura posterior del proyecto.

1.1.Motivación

Este proyecto ha sido elaborado gracias a conocimientos adquiridos durante la carrera. Las asignaturas clave para el desarrollo del mismo han sido: Métodos Estadísticos en Economía para la interpretación de las medidas estadísticas básicas, Econometría para la interpretación de correlaciones, y elaboración de la bases de datos y de modelos estadísticos; y por último, Métodos Cuantitativos para la Ayuda a la Toma de Decisiones para la interpretación de árboles de decisión. Además, se han aplicado conocimientos externos a las asignaturas de la Universidad en la elaboración y comprensión de Árboles de Decisión CART y para utilizar el programa R.

Debido a que la estadística es una de las ramas que más me han interesado y despertado curiosidad durante la carrera, el proyecto está enfocado en el análisis estadístico de un tema de actualidad. Además, la obtención de datos, resultados y el análisis de los mismos, al tratarse de un estudio reciente puede ser de ayuda y motivación para otros proyectos, que siempre es de agradecer. Por lo tanto, decidí enfocarlo en un tema que actualmente está afectando a todas las personas y a países, el Covid-19.

Según la OMS, “Los coronavirus son una extensa familia de virus que pueden causar enfermedades tanto en humanos como animales. En los humanos, se sabe que varios coronavirus causan infecciones respiratorias y pulmonares, que pueden ir desde el resfriado común hasta enfermedades más graves como el síndrome respiratorio de Oriente Medio (MERS) y el síndrome respiratorio agudo severo (SARS)”.

El coronavirus actual, o comúnmente conocido como Covid-19, según la OMS, es la enfermedad infecto-contagiosa causada por el coronavirus que se ha descubierto más recientemente.

Los síntomas más habituales de la COVID-19, según la OMS, “son la fiebre, la tos seca y el cansancio. Otros síntomas menos frecuentes que afectan a algunos pacientes son los dolores y molestias, la congestión nasal, el dolor de cabeza, la conjuntivitis, el dolor de garganta, la diarrea, la pérdida del gusto o el olfato y las erupciones cutáneas o cambios de color en los dedos de las manos o los pies. Estos síntomas suelen ser leves y comienzan gradualmente. Algunas de las personas infectadas solo presentan síntomas levísimos” [1].

También comenta la OMS que alrededor del 80% de las personas se recuperan de la enfermedad sin necesidad de tratamiento hospitalario, y que alrededor del 20% de las personas que contraen la COVID-19 acaba presentando síntomas graves y experimenta dificultades respiratorias. Las personas mayores y las que padecen afecciones médicas previas son los más vulnerables, sin embargo, cualquier persona puede contraer la COVID-19 y caer gravemente enferma.

El Covid-19 comenzó en Wuhan, China, en diciembre de 2019, de ahí su nombre. Algunos países ya advertían de su peligro para la sociedad, pero la vida siguió su curso como si el virus estuviese aislado en China. Se extendió levemente por algunas zonas de Asia, y posteriormente por Europa, que a mediados de febrero provocó un caos y una crisis económica, social y sanitaria por encima del resto de países, siendo los más afectados entonces Italia y España. En abril, finalmente se extiende con ferocidad por algunos países de América del norte y sur, principalmente Estados Unidos y Brasil.



Ha sido evidente para todos que se ha tratado de una pandemia a nivel global que ha afectado a todos los países, ¿pero se han preguntado por qué unos países se han visto afectados más que otros, aun estando en condiciones similares?

1.2. Hipòtesis

En este trabajo se trata de analizar los factores que pueden influir significativamente en la propagación e impacto del Covid-19 en la sociedad, o mejor dicho, distintas sociedades.

A partir del análisis de las variables se podrá averiguar los países con más riesgo y porque son los más afectados. Así como las variables que no influyen en el impacto del virus en los países.

A partir del análisis conjunto de las variables se pretende obtener un modelo que explique el impacto social del actual coronavirus y que variables son las más influyentes para el número de fallecimientos y contagiados por el Covid-19.

1.3. Objetivos

En cuanto a los objetivos, el objetivo general del proyecto es determinar qué variables, de las escogidas, son influyentes sobre el nivel de fallecidos e infectados por Covid-19, y por consiguiente, capaces de formar un modelo explicativo.

Los objetivos para llevar a cabo el proyecto son:

- Examinar el virus del Covid-19.
- Comparar con otros virus de la historia.
- Considerar variables a introducir en la base de datos.
- Recopilar los datos de las variables.
- Contrastar datos entre fuentes.
- Discriminar variables con pocos datos, o mucha relación con otras.
- Introducir datos en la base de datos.
- Organizar los datos de las variables con los países coincidentes.
- Definir las variables establecidas.
- Analizar las variables de forma individual (univariante).
- Analizar las variables por parejas de dos con las variables principales (bivariante).
- Analizar las variables de forma múltiple (múltivariante).
- Definir modelo explicativo a partir de árboles de decisión.
- Desarrollar conclusiones en base al análisis de los datos.
- Determinar la influencia de las variables aportadas en el impacto del virus.

1.4. Estructura del proyecto

El proyecto se estructura de la siguiente manera:

Metodología y técnicas: empieza con una introducción sobre los antecedentes, realizándose un análisis de distintos virus para poder tener una idea general de estos. En cuanto a la metodología y técnicas, en ellas se explican los procesos que se han llevado a cabo para el proyecto así como las herramientas y técnicas estadísticas aplicadas.

Cálculos y resultados: aquí se hallan varios apartados. Se comienza con una descripción de todas las variables utilizadas en el proyecto para comprender adecuadamente la interpretación de los resultados. Seguidamente se analizan las variables individualmente para encontrar matices y puntos anómalos. Después de analizarse una a una, se procede por analizarlas por parejas, con el objetivo de encontrar variables que se relacionen con las variables de interés y poder interpretar. Por último, se analizan todas las variables conjuntamente con el fin de obtener un modelo que explique qué variables son las que más influyen en el impacto del virus.

Conclusiones y futuras líneas de investigación: en estos apartados se interpretan los resultados obtenidos más relevantes. Además de dar pie a la realización de otros proyectos similares a quienes les sirva de motivación o utilidad.

2. Metodología y Técnicas

En esta sección se explica la metodología aplicada para la investigación. En primer lugar, se exponen los antecedentes, donde se realiza una investigación documental analizando los distintos tipos de virus para entender la magnitud del virus en cuestión. Seguidamente, en el apartado de metodología, se presenta el procedimiento seguido en el proyecto. Y por último, se exponen en el apartado de técnicas las herramientas utilizadas con el programa R.

2.1. Antecedentes

Durante el transcurso de la historia han existido numerosos virus y enfermedades que han penetrado en la vida humana perjudicando el normal funcionamiento de la sociedad. Se debe tener en cuenta que dos de los factores que determinan el grado de agresividad de los virus en cuanto al perjuicio de las personas son la mortalidad y la contagiosidad. En el siguiente gráfico se puede observar muchos de los virus según estas dos variables, en el eje Y nos encontramos el porcentaje de mortalidad, mientras que en el eje X nos encontramos con el factor R, es decir el grado de contagio que puede llegar a tener el virus, si el R es igual a 3 quiere decir que un infectado llega a contagiar a 3 individuos más, produciéndose una progresión geométrica a raíz de los nuevos infectados. Los datos aportados posteriormente están actualizados a la fecha de marzo de 2020, mientras que el gráfico mostrado en la Ilustración 1 es de 2016 (gráfico actualizado en link).

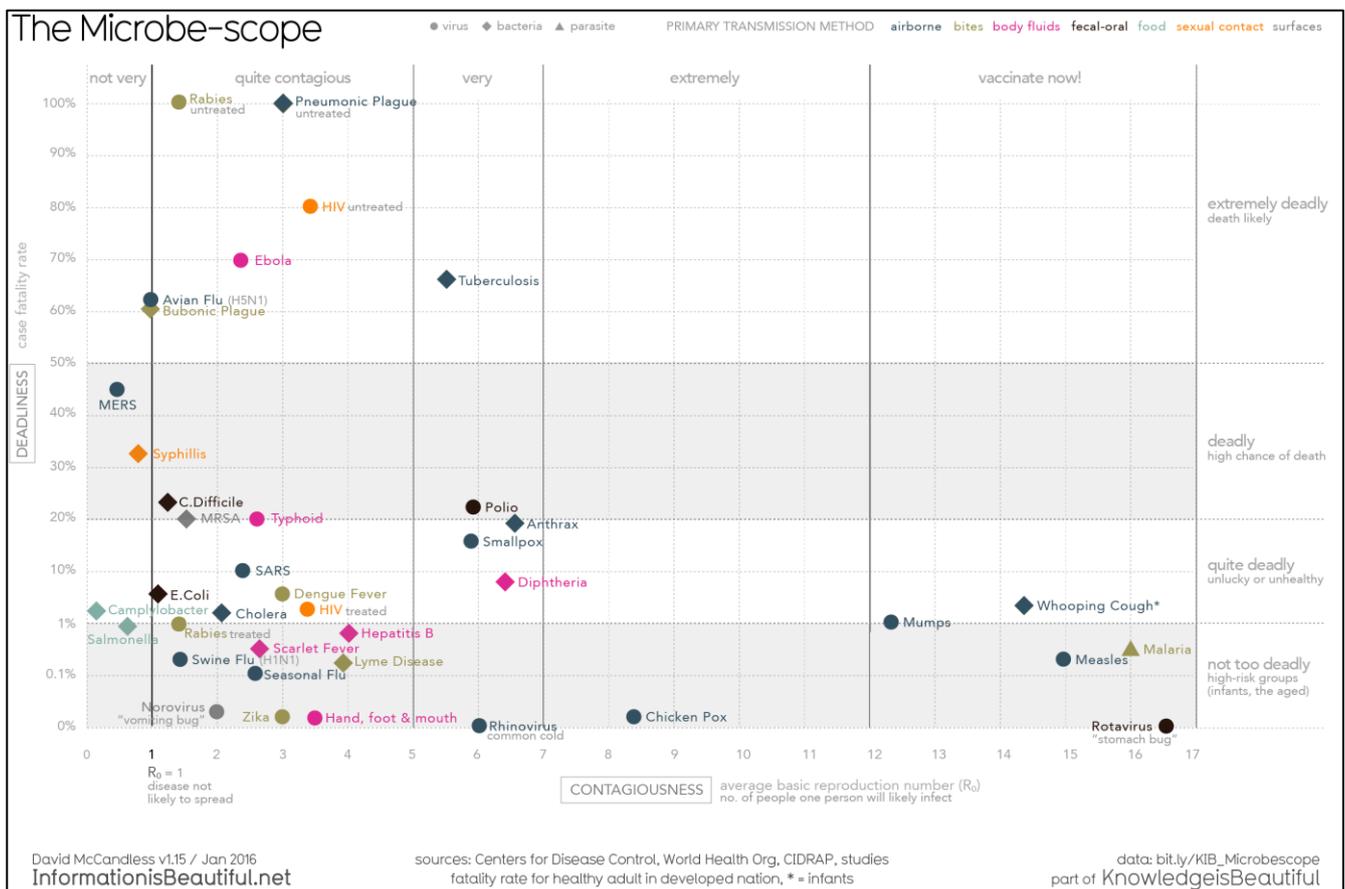


Ilustración 1. Gráfica comparativa de virus, parásitos y bacterias. Fuente: informationisbeautiful.net.

El actual Covid-19 se sitúa actualmente con un R₀ de 2 a 3 y con una tasa de mortalidad alrededor del 4%. Esto le sitúa en una posición relativamente baja en comparación con otros virus. Pero que sea más bajo no quiere decir que perjudique menos a la sociedad, de hecho, tener esta proporción equilibrada entre mortalidad y factor reproductivo (R₀) suficiente para extenderse como una pandemia es lo que ha originado tanto descontrol sobre el virus. Si observamos la ilustración 1 y lo comparamos con un virus cercano y similar como el MERS,

este último tiene una mortalidad del 45% y un R_0 menor que 1, lo que significa que los infectados del MERS apenas lo contagiaron, y además, estos infectados presentaron síntomas fuertes y más rápidamente lo que facilitó la identificación de estos y su puesta en cuarentena. Al igual pasa con la rabia, un virus que generalmente se contagia a través de perros y murciélagos, que posee una mortalidad del 100% unido a un R_0 igual a 10, y que al año se encuentran unos 13.000 casos nuevos. La rabia es un virus que una vez causa síntomas es difícil la supervivencia, cabe pensar que una enfermedad tan mortal y tan contagiosa como esta es raro que no se haya propagado a nivel mundial tan rápidamente como el Covid-19 pero esto es gracias a que se presentan síntomas de forma clara y se puede evitar su propagación por medio de la cuarentena del infectado.

También se sabe que la gripe común ha llegado a ser un virus del que preocuparse. A nivel mundial se contagian al año 4 millones de personas y que causa, de media, 400 mil muertes. En el caso de España, la temporada de 2017/2018 fue más grave que la del periodo posterior que se encontraban los resultados mostrados en la ilustración 2:



Ilustración 2. Impacto de la gripe en España. Fuente: vacunasaep.org.

En el caso de Estados Unidos las estimaciones de la actual temporada de gripe son de alrededor de 46 millones de enfermos (desde muy leves a muy graves), 500.000 hospitalizaciones y 40.000 muertes, con una tasa de mortalidad alrededor de 0,05%.

La principal diferencia entre estos dos virus es que el coronavirus no tiene vacuna mientras que la gripe sí.

Para hacer una comparativa de los principales virus que han llegado a suponer una pandemia se encuentran: la fiebre hemorrágica de Marburgo; virus de Ébola, virus de Hanta; gripe aviar; fiebre de Lassa; fiebre del Dengue; Peste Bubónica; Gripe española; la rabia; VIH; y el actual covid-19.

Tabla 1. Comparativa virus y bacterias pandémicas.

Nombre del virus	Fecha de mayor impacto	Mortalidad	Contagio (R ₀)	Nuevos casos por año	Muertes por año	Observaciones
Marburgo*	Años 60	52%	1,3	22	19	Clínicamente similar al ébola.
Ébola*	1997-2006; 2014-2016	50%	2,2	41	4	El mayor brote se produjo en 2016.
Hanta*	Desde 1959	35%	2,1	175.000	n/A	Tratamiento de apoyo.
Gripe Aviar*	1997-2005	58	1,2	2.000	18	La prevalencia ha aumentado con el aumento de población mundial de aves de corral en los 90's.
Tuberculosis	Desde 1882	13% (58% sin tratamiento)	2,3	9.000.000	1.200.000	Alrededor de 53 millones de vidas salvadas entre 2000 y 2016 por diagnóstico y vacunación exitosos.
Dengue	Desde 1780	0,04% (22% sin tratamiento)	11	101,1 mill.	37.800	La incidencia global ha aumentado gravemente en las últimas décadas. Se estima que alrededor de la mitad de la población global está en riesgo.
Peste bubónica	1347-1353	15% (60% sin tratamiento)	3,5	650	117	En la Edad Media mató a alrededor de la mitad de la población mundial (50 mill.).
Gripe española	1918-1920	5,5%	2,2	0	0	Mató de 50 a 100 mill. de personas (del 3% al 5% de la población mundial).
La rabia	Se remonta incluso antes de la antigua Grecia (550 A.C)	100%	10	13.300	13.300	A pesar de la vacunación generalizada de perros y personas todavía provoca decenas de miles de muertes al año, principalmente en Asia y África.
VIH	Desde 1981	2,1% (80% sin tratamiento)	0 (6,6 sin tratamiento)	1.900.000	1.000.000	Si no se trata causa SIDA. Actualmente no hay cura pero los tratamientos pueden suprimir el virus.
Covid-19	Desde 2019	De 0,7 a 3,4%	De 2 al 3	Presente (6mill. Aprox.)	Presente (330.000 Aprox.)	Actualmente su tasa de mortalidad estimada es del 4,5% [2].

*Sin tratamiento específico.

En la tabla 1 aparecen por fila los virus más conocidos en las últimas décadas. Por columna encontramos: la fecha en la que apareció el virus o bacteria y el periodo de mayor incidencia, incluso si hubo una réplica en otro periodo; el grado de mortalidad en porcentaje, con y sin tratamiento; el número de casos nuevos que se encuentran a día de hoy; y el número de

muerres al año por dicha enfermedad; y una pequeña observación de cada una para completar información.

Por lo que se puede observar, el Covid-19 se encuentra con valores que son comparables a otras grandes enfermedades provocadas por otros virus. Se debe tener en cuenta que, con el actual virus, se contabiliza en la mayoría de países por datos confirmados. Esto provoca que en el caso de los infectados sea difícil estimar la cantidad de población que padeció la enfermedad sin estar confirmado, y es probable, que sea superior proporcionalmente a la cantidad de muertes no confirmadas.

Actualmente, se realizan estudios de la genética del coronavirus puesto que existen distintas cepas según las mutaciones que va experimentando el virus durante la incubación en otro nuevo huésped, ya que existen distintos casos de mortalidad y síntomas en distintos países e incluso en el mismo. Se está estudiando si puede ser causa de la distinta genética del virus, o cepas, o de la distinta genética de las personas.

Y además, con suerte, en unos meses se podrán desarrollar vacunas y paliativos para aminorar el impacto de una posible réplica. Podemos decir que somos afortunados de vivir en una sociedad avanzada y podemos actuar con más rapidez ante problemas nuevos y desafiantes en casi todos los ámbitos científicos.

Según Matemáticas y sus fronteras, marzo 2020: “Aunque el SARS-nCoV-2 (inicial y comúnmente conocido como Covid-19) es un nuevo virus y nos pone ante un nuevo desafío, no resulta totalmente desconocido por pertenecer a la familia de los coronavirus. No obstante, la pandemia que ha generado no sólo tiene que ver con los aspectos biológicos del virus sino, quizá con mayor importancia, con aspectos sociales y demográficos” [3].

2.2. Metodología

En primer lugar se debe averiguar qué factores pueden ser significativos a la hora de una propagación de un virus. A priori, no sabemos qué factores pueden ser significativos hasta que no lo analicemos por lo que se escogerán varios factores de distinta naturaleza para averiguar si influyen: principalmente en número de muertes y número de infectados; y secundariamente, en la variación de paro y variación de PIB. Estas últimas, son las variables principales que medirán el impacto del Covid-19.

Una vez escogidas las variables, se procede a la obtención de los datos de los distintos países. Alrededor de 140 países se analizarán sobre estas variables, de modo que sea lo más representativo posible. Exceptuando las variables de variación de paro y PIB que se analizarán con unos 50 países.

Se trata de llevar a cabo un análisis de las variables de forma individual y de forma colectiva para averiguar si existe relación o si son influyentes sobre las principales variables. A través del análisis se formalizará un modelo a través de los árboles de decisión CART.

A modo de esquema, se muestra la consecución de tareas para llevar a cabo la construcción de un modelo y análisis de las variables:

- Recopilación de información relacionada con el coronavirus actual.
- Planteamiento y selección de variables.
- Construcción de la base de datos (sucia).
- Depuración de la base de datos (limpia).
- Análisis univariante.
- Análisis bivariante.
- Análisis multivariante.
- Validación de modelos.
- Comparación de modelos.
- Selección de modelos.

A continuación, se desglosa los apartados anteriores del análisis, mostrándose los análisis que se han realizado en concreto:

Análisis univariante:

- Análisis descriptivo.
- Tabla de frecuencia.
- Gráfico de sectores y gráfico de barras.
- Diagrama Cajas y Bigotes.
- Histograma.

Análisis bivalente:

- Matriz de correlaciones.
- Gráficos XY.
- Diagrama Cajas y Bigotes.

Análisis multivariante:

- Árboles de Decisión (CART).

Validación de modelos:

- Validación con distintos datos.

Comparación de modelos:

- Variables que intervienen.
- Orden importancia variables.
- Signo de parámetros.

2.3. Técnicas

2.3.1. Técnicas aplicadas con “R”

Para el análisis univariante se utilizarán las siguientes técnicas estadísticas:

- Medidas estadísticas básicas: mínimo; 1er cuartil; mediana; media; 3er cuartil; y máximo.
- Tabla de frecuencias: indica la repetición de los valores de una variable.
- Gráfico de sectores: representación gráfica de los valores de la variable cualitativa por sectores.
- Gráfico de barras: representación gráfica de los valores de la variable cualitativa en barras.
- Gráfico de caja y bigotes: son una presentación visual que describe varias características importantes, al mismo tiempo, tales como la dispersión y simetría. Consiste en una caja rectangular, donde los lados más largos muestran el recorrido intercuartílico. Este rectángulo está dividido por un segmento vertical que indica donde se posiciona la mediana y por lo tanto su relación con los cuartiles primero y tercero (el segundo cuartil coincide con la mediana). Esta caja se ubica a escala sobre un segmento que tiene como extremos los valores mínimo y máximo de la variable. Las líneas que sobresalen de la caja se llaman bigotes, el resto de datos o casos que no se encuentre dentro de este rango es marcado e identificado individualmente (puntos anómalos) [5].
- Histograma: representación gráfica de los valores en intervalos de la variable cuantitativa en barras.

En el análisis bivalente se utilizan las matrices de correlación para determinar la influencia de unas variables respecto a otras. Además, las variables de mayor correlación se analizarán más específicamente a través de Gráficos XY, que muestra en una gráfica los valores en base a dos variables, eje X y eje Y. Y por último, también se utiliza el Diagrama Caja y Bigotes clasificando la variable según los valores de otra variable cualitativa.

Finalmente, para el análisis multivariante se utilizan árboles de decisión, CART (árboles de clasificación y regresión, Breiman et al. 1984) es un método que construye árboles binarios con los casos de la muestra, en los que cada nodo se divide en dos ramas o subconjuntos de la muestra. Los nodos terminales, las hojas del árbol, permiten predecir una clase de pertenencia o un número real, por lo que se puede utilizar como método de clasificación o de regresión. En el caso de la regresión, se utiliza para variables cuantitativas, las divisiones buscan minimizar el cuadrado del residuo o error de predicción (criterio de mínimos cuadrados), y la predicción se basa en la media ponderada para cada nodo del árbol. Se trata de dividir en cada paso los datos en dos partes homogéneas, de modo que los casos dentro de cada parte sean similares entre sí, y distintos de los casos de la otra parte [6].

2.3.2. Herramienta estadística CART con “R”

Para ejecutar las herramientas y realizar los cálculos para los análisis se usa en este trabajo el programa “R”, que es un software estadístico que permite manejar y analizar datos.

R es un ambiente de programación formado por un conjunto de herramientas muy flexibles que pueden ampliarse fácilmente mediante paquetes, librerías o definiendo nuestras propias funciones. Además es de código abierto [4].

Esto quiere decir que las operaciones se realizan a través de instrucciones en un terminal, por lo que no es un programa gráfico. Al ser un lenguaje de programación, tenemos la posibilidad de ampliar su funcionalidad en la medida de nuestros conocimientos.

R-Commander es una Interfaz Gráfica para usuario que permite acceder a muchas capacidades del entorno estadístico R sin que el usuario tenga que conocer el lenguaje de comandos propio de este entorno.

La herramienta CART (árboles de clasificación y regresión, Breiman et al. 1984) es un método que construye árboles binarios con los casos de la muestra, en los que cada nodo se divide en dos ramas o subconjuntos de la muestra. Los nodos terminales, las hojas del árbol, permiten predecir una clase de pertenencia o un número real, por lo que se puede utilizar como método de clasificación o de regresión. En el caso de la regresión, se utiliza para variables cuantitativas, las divisiones buscan minimizar el cuadrado del residuo o error de predicción (criterio de mínimos cuadrados), y la predicción está basada en la media ponderada para cada nodo del árbol. Se trata de dividir en cada paso los datos en dos partes homogéneas, de modo que los casos dentro de cada parte sean similares entre sí, y distintos de los casos de la otra parte [6].

Para la construcción de árboles de clasificación la variable respuesta debe ser categórica, mientras que para construir árboles de regresión la variable debe ser numérica, ordenándose los valores para encontrar el punto de corte óptimo que divide los datos en dos grupos, menor y mayor o igual.

El inconveniente de los árboles de clasificación es que tienen un problema de sobreajuste, por lo que a la hora de introducir nuevos datos, o predecir nuevos casos, la precisión puede variar drásticamente. Por lo general, un árbol pequeño es más fácil de interpretar y tiene menos problema de sobreajuste, pero cuanto más grande mayor es también su precisión y pureza conseguida.

En la construcción del árbol es necesario una medida de impureza dentro de un nodo como criterio de división, y luego, un criterio de costo-complejidad para podarlo, obteniendo así un modelo preciso en el que si una rama no aporta una certeza superior a la complejidad sobre el modelo se elimina dicha rama.

Para ajustar los árboles de regresión se utiliza la librería *rpart* de R. El paquete *rpart* permite construir modelos de clasificación o de regresión con estructura general. Tiene dos opciones de medida de impureza para el particionamiento recursivo: el índice de Gini para árboles de clasificación y la entropía para árboles de regresión. Para visualizar los árboles de forma gráfica se utiliza la librería *rpart.plot* de R.

3. Cálculos y Resultados

En la siguiente sección se describen una serie de variables objeto de estudio para el análisis posterior. Dichos factores son de carácter social, sanitario, económico y político.

Seguidamente se realiza un análisis univariante, bivariante y multivariante de dichas variables: con el primer análisis se pretende conocer características de las variables que puedan ofrecer información para el estudio; con el segundo se pretende conocer la relación existente entre dos variables, así como el comportamiento conjunto, y con el tercer análisis se pretende construir un modelo que explique qué variables son determinantes en la magnitud de fallecidos e infectados, a través de segmentación de variables con árboles de decisión CART.

3.1.Descripción de variables

Las variables escogidas para los análisis y planteamiento del modelo, así como su abreviatura y definición, son las siguientes:

- Países: conjunto de países de todos los continentes.
- Número de camas de hospital por cada 1.000 habitantes (CAMAS): las camas de hospital incluyen camas para pacientes hospitalizados disponibles en hospitales públicos, privados, generales y especializados y centros de rehabilitación. En la mayoría de los casos, se incluyen camas para cuidados agudos y crónicos. Datos actualizados al último dato disponible en el periodo 2010-2020 (contrastado con datos banco mundial). **Debido a que el nivel de servicios para pacientes hospitalizados requerido para cada país depende de varios factores, como los problemas demográficos y la carga de morbilidad, no existe un objetivo global para el número de camas de hospital por país. Por lo tanto, aunque 2 camas por cada 1.000 en un país pueden ser suficientes, 2 camas por cada 1.000 en otro pueden ser lamentablemente inadecuadas debido a la cantidad de personas hospitalizadas por la enfermedad.*
- Tasa de mortalidad cruda por cada 1000 habitantes (MORTA): la tasa bruta de mortalidad indica el número de muertes que ocurren durante el año, por cada 1000 habitantes estimados a mediados del año 2017. **(La tasa bruta de mortalidad es un buen indicador del estado general de salud de un área geográfica o población. La tasa bruta de mortalidad no es apropiada para la comparación de diferentes poblaciones o áreas con grandes diferencias en la distribución por edades)*
- Migración neta (MIG): la migración neta es el total neto de migrantes durante el período, en este caso 2017, es decir, el número total de inmigrantes menos el número anual de emigrantes, incluidos ciudadanos y no ciudadanos. Los datos son estimaciones de cinco años. *La División de Población de las Naciones Unidas proporciona datos sobre la migración neta y el stock de migrantes. La migración internacional es el componente del cambio de población más difícil de medir y estimar de manera confiable. Por lo tanto, la calidad y cantidad de los datos utilizados en la estimación y proyección de la migración neta varía considerablemente según el país.*
- Población urbana (URB): porcentaje de la población en el periodo 2018. La población urbana se refiere a las personas que viven en áreas urbanas según lo definido por las oficinas nacionales de estadística. Los datos son recopilados y suavizados por la División de Población de las Naciones Unidas. *Se debe tener especial cuidado al interpretar las cifras del porcentaje urbano de diferentes países. Los países difieren en la forma en que clasifican a la población como "urbana" o "rural". La población de una ciudad o área metropolitana depende de los límites elegidos.*
- Médicos por cada 1.000 habitantes (MED): entre los médicos se incluyen médicos generalistas y especialistas. Datos actualizados al último dato disponible (contrastado con datos banco mundial). ** La OMS estima que se necesitan al menos 2.5 personal médico (médicos, enfermeras y parteras) por cada 1,000 personas para proporcionar*

una cobertura adecuada con las intervenciones de atención primaria (OMS, World Health Report 2006).

- Días transcurridos desde la aparición del virus hasta el primer día de confinamiento (DIAS): se refiere a los días sucedidos desde la fecha en la que se confirma el primer infectado en el país hasta la aplicación de alguna medida de confinamientos sea leve o estricta. Nulo en los casos de no confinamiento, o no aparición de virus en el país. Algunas fechas son aproximadas debido a la falta de información de algunos países.
- Medida confinamiento obligatorio (CONFI): se divide en cuatro grupos según la cobertura u obligatoriedad de las medidas impuestas por los gobiernos en relación al confinamiento de los ciudadanos. Estos cuatro grupos son:

Total: se aplica confinamiento obligatorio a todo el país.

Parcial: se aplica confinamiento obligatorio a una parte del país o a un grupo de personas por factores de riesgo.

Mínima: se aplican restricciones mínimas en el país (como cierre de algunos comercios, limitación de reuniones o confinamiento no obligatorio).

Nulo: no se aplica ningún tipo de restricción hasta la fecha.

Parte de estos datos se han buscado de forma independiente en artículos basados en otros países.

- Gasto en salud, pública + privada (GASTO): nivel de inversiones de capital en salud expresado como porcentaje del PIB. Las inversiones de capital en salud incluyen infraestructura de salud (edificios, maquinaria, TI) y existencias de vacunas para emergencias o brotes. (Datos actualizados al último dato, contrastado con datos banco mundial).
- Población en aglomerados urbanos > 1 millón (% de la población total en 2018) (AGLO): la población en aglomerados urbanos de más de un millón es el porcentaje de la población de un país que vive en áreas metropolitanas cuya población en el año 2000 superaba el millón de personas.
- Densidad de población (personas por kilómetro) (DENSI): la densidad de población se define como la población a mitad de año 2018 dividida por la superficie territorial en kilómetros cuadrados. La población incluye a todos los residentes independientemente de su estado legal o de ciudadanía, con excepción de los refugiados no asentados permanentemente en el país de asilo, que suelen considerarse parte de la población del país de origen. El área de tierra es la superficie total de un país, sin incluir la superficie cubierta por masas de agua interiores (ríos y lagos), los derechos del país sobre la plataforma continental ni las zonas económicas exclusivas.
- Prevalencia de desnutrición (% población en 2017) (DESNUT): la población por debajo del nivel mínimo de consumo de energía en la dieta (también conocida como prevalencia de desnutrición) muestra el porcentaje de la población cuya ingesta de alimentos es insuficiente para satisfacer las necesidades de energía de la dieta de forma continua. Los datos con valor igual a 5% son países con una prevalencia de desnutrición entre el 0% y el 5%.

- Población con instalaciones para lavarse las manos, como agua y jabón (2010-2017) (JABON): el porcentaje de personas que viven en hogares que tienen instalaciones de lavado de manos con agua y jabón disponibles en las instalaciones. Las instalaciones de lavado de manos pueden ser fijas o móviles e incluyen un fregadero con agua del grifo, cubetas con grifos, grifos y jarras o cuencos designados para el lavado de manos. El jabón incluye jabón en barra, jabón líquido, detergente en polvo y agua jabonosa, pero no incluye cenizas, tierra, arena u otros agentes de lavado de manos. (Europa y países de ingresos alto se encuentra al 100% o casi).
- Variación relativa trimestral del PIB (PIB): diferencia porcentual del PIB generado en el país en el primer trimestre del 2020 respecto al trimestre anterior.
- Fallecimientos por el Covid-19 (FALLE): número de muertes en el país por cada 100.000 habitantes a causa directa por el Covid-19 a fecha 18/06/2020.
- Infectados por el Covid-19 (INFECT): número de infectados en el país por cada 100.000 habitantes confirmados de Covid-19 a fecha 18/06/2020.
- Variación absoluta trimestral del desempleo (PARO): diferencia absoluta del nivel de desempleo de la población del país en el primer trimestre del 2020 respecto al trimestre anterior.

*(Recomendaciones y consejos de interpretación aportados por el banco mundial).

3.2. Análisis univariante

A continuación, se va a proceder a calcular las medias, y demás indicadores objeto de análisis de las variables que se plasmarán en la tabla 2. Podemos observar los cálculos en la siguiente tabla:

Tabla 2. Resumen estadístico de las variables.

Variable	Mínimo	1er cuartil	Mediana	Media	3er cuartil	Máximo
CAMAS	0,1	1,3	2,5	3,14	4,33	13,8
MORTALIDAD	1,17	5,82	7,26	7,72	9,5	15,5
MIGRACION	-3.266.243	-77.093	-4.999	27.798	77.998	4.774.029
P. URBANA	13,03	47,55	65,90	62,92	80,69	100
MEDICOS	0,0008	0,53	1,85	2	3,1	8,2
DIAS	-6	9	18	22,25	29,5	82
GASTO SALUD	0,9	5,08	6,5	6,8	8,43	16,8
AGLOMERACION	4,03	14,48	21,61	25,45	32	100
DENSIDAD	3,25	38,225	82,91	346,73	171	19.196
DESNUTRICION	2,5	2,5	4,85	9,6	10,78	51,3
PIB	-7	-3,1	-1,8	-2,01	-0,5	3
FALLECIMIENTOS	0,0	0,395	1,46	7,631	5,28	124,33
INFECTADOS	0,0	18,38	55,24	190,83	222,97	2990,06
PARO	-1,3	0,0	0,3	0,29	0,6	1,4

En las siguientes tablas se muestran dos variables que son cualitativas, por lo que en este caso haremos un cálculo sobre los datos que se repiten de cada variable, mostrando el tamaño en porcentaje sobre el total, para saber que condición se repite más en los países.

Tabla 3. Variable medida de confinamiento.

Medida Confinamiento	Número de países que lo aplican	% sobre el total
Ninguna	14	9,46 %
Mínima	16	10,81 %
Parcial	46	31,08 %
Total	72	48,65 %
Total de países	148	100 %

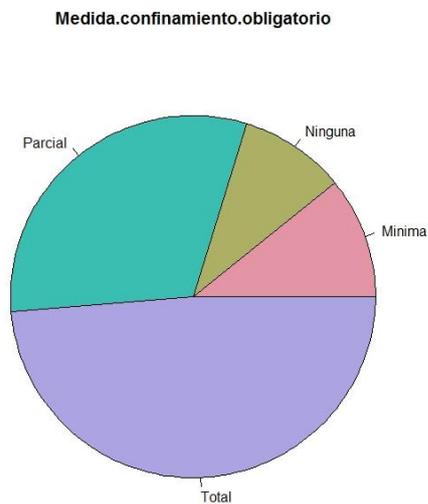


Ilustración 3. Gráfico de sectores de la variable CONFI.

En la ilustración 3, se puede ver por el gráfico de sectores que casi el 50% de los países tomados como muestra en la variable han aplicado medidas de confinamiento obligatorio total sobre la población, el 30% han tomado medidas parciales de confinamiento, mientras que el 10% de los países han adoptado medidas mínimas y el otro 10% no han aplicado ninguna.

Tabla 4. Variable JABON (% de población con instalaciones básicas de agua e higiene).

Variable: JABON (%)	Número de países	% sobre el total
0 a 10	9	7,38 %
10 a 20	4	3,28 %
20 a 30	5	4,1 %
30 a 40	3	2,46 %
40 a 50	6	4,92 %
50 a 60	4	3,28 %
60 a 70	3	2,46 %
70 a 80	6	4,92 %
80 a 90	13	10,65 %
90 a 100	69	56,55 %
Total de países	122	100 %

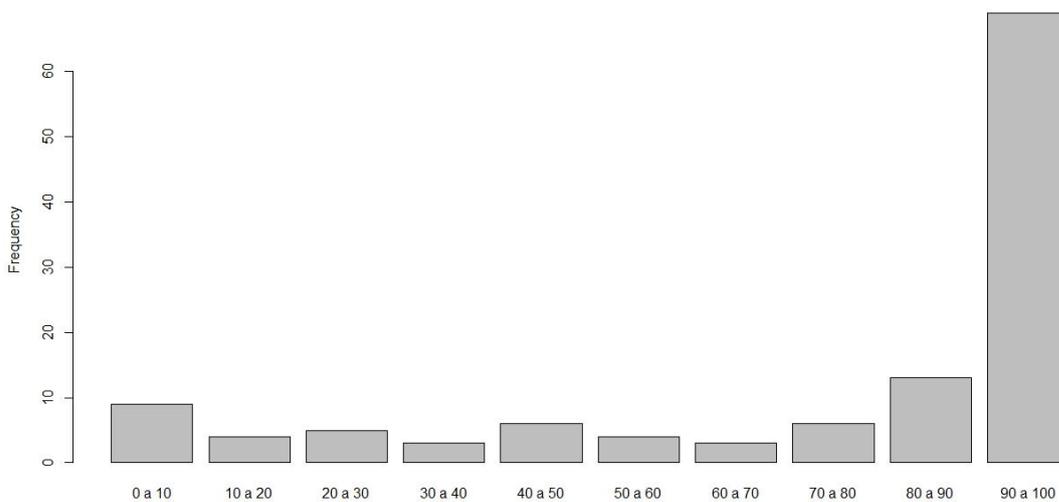


Ilustración 4. Gráfico de barras de la variable JABON.

En esta variable se han dividido por grupos de porcentaje de población del país con instalaciones básicas de agua e higiene. A través de la ilustración 4, se ve de forma directa la gran diferencia entre un grupo y el resto, siendo relativamente positivos los resultados. En la tabla 4, vemos que el grupo de "90 a 100", supone el 56% del total de países con un 90% a 100% de la población con acceso a instalaciones básicas de agua e higiene, mientras que el resto de grupos representan individualmente entre el 3% y 10% del total de países. La única parte negativa de todo esto es que el tercer grupo con mayor tamaño es el "0 a 10", el cuál no debería en ningún caso representar casi el 8% de los países.

En la variable “número de camas de hospital”, existe una diferencia de 13,7 entre el valor mínimo y el máximo, esta diferencia puede estar relacionada con la brecha económica entre países o la inversión en sanidad. También se puede ver que el 50% de los países tienen 2,5 camas o menos y que el 75% tiene 4,33 camas o menos, siendo la media de los países de 3,14. Por lo que podemos decir que algunos pocos países tienen un número de camas más elevado que el resto.

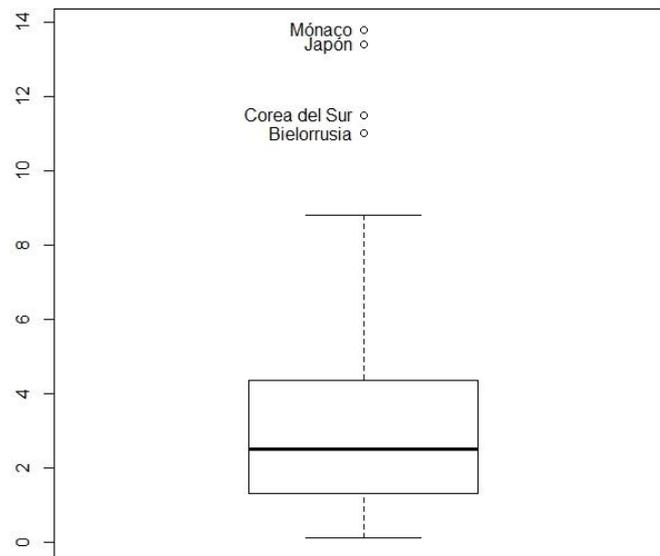


Ilustración 5. Gráfico de caja y bigotes de la variable CAMAS.

En la ilustración 5 vemos que la mayoría de los países se agrupan por debajo de 5 camas por cada 1000 habitantes, sin embargo existen varios países con un número elevado de camas con más diferencia que el resto. Vemos que Mónaco, Japón, Corea del Sur y Bielorrusia poseen un número de camas muy por encima del resto, son los puntos anómalos de la variable CAMAS.

La variable “mortalidad del país” se comporta de una forma más normal, aparentemente con una distribución más regular.

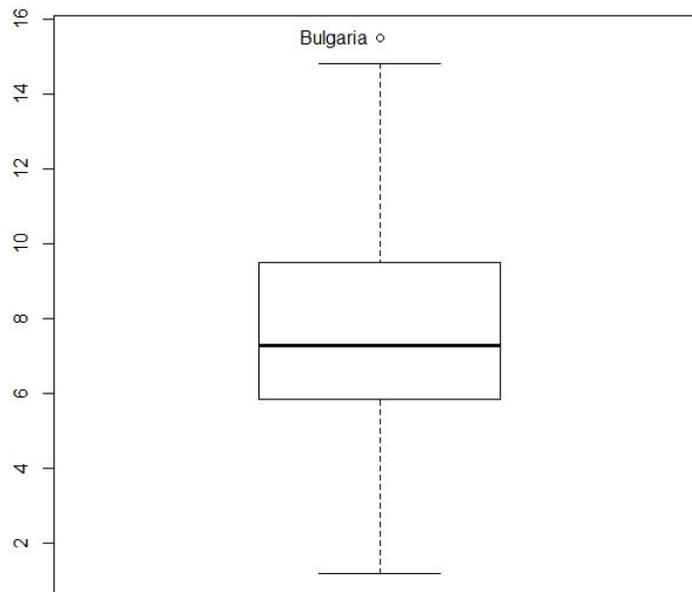


Ilustración 6. Gráfico de caja y bigotes de la variable MORTA.

Precisamente, en el gráfico de la ilustración 6, se aprecia como esta variable tiene una forma regular con la misma distancia de los bigotes respecto a la caja lo que nos indica que se distribuye normalmente. El único punto anómalo es Bulgaria pero sin mucha diferencia respecto al anterior. De todas formas existe notoria diferencia entre el valor mínimo y el máximo, siendo interesante comprobar si influye en los países esta variable.

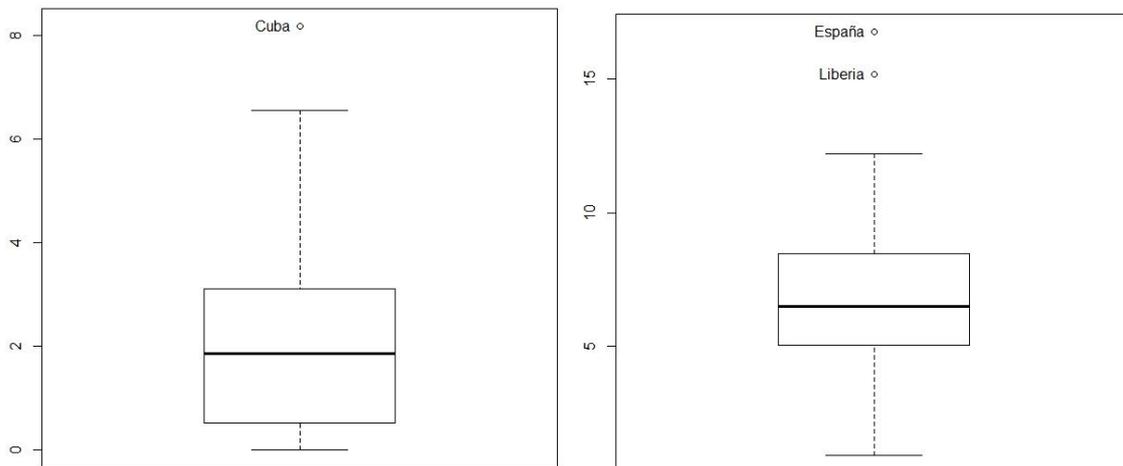


Ilustración 7 e Ilustración 8. Gráficos de caja y bigotes de las variables MED y GASTO.

Como se puede observar en la ilustración 7, la variable “médicos” tiene un solo punto anómalo que es Cuba, con 8,2 médicos por cada 1000 habitantes, el resto se sitúa prácticamente por debajo de 6, con una media total de 2 médicos y con el 50% de países con 2 o menos médicos por cada 1000 habitantes, y el 75% con 3 o menos. Habrá que probar si en este caso son tan

importantes los médicos, o si por tener más médicos no implica un menor impacto en la sanidad. Así como el número de camas.

En la variable “gasto en salud” si se puede apreciar una forma más normal del gráfico de caja y bigotes. Si nos fijamos en la ilustración 8, en este caso, la mayoría de países se encuentran entre el 5% y el 8 % de gasto en salud sobre el PIB. Pero encontramos un país de los más afectados por el Covid-19 y el cual es un punto anómalo de nuestra variable, España se sitúa en el punto máximo, con un valor de 16, duplicando a la mayoría de países, junto con Liberia.



Ilustración 9 e Ilustración 10. Gráficos de caja y bigotes de las variables URB y DENSI.

En la ilustración 9, vemos que distribución de los datos es muy amplia, en donde hay países que se acercan al 100% de población urbana, como países que se acercan al 0%. Por ello no existen puntos anómalos en esta variable. Sin embargo vemos en la ilustración 10, en la variable “densidad de población”, que la situación es contraria donde prácticamente todo el conjunto de datos para esa variable se concentra en un pequeño tramo, de 3,75 a 171 personas por kilómetro cuadrado, se encuentran el 75% de los países, y donde los países, Mónaco, Singapur, Maldivas Malta, Bangladés, Líbano, Barbados, San Marino, Países Bajos y Ruanda son puntos anómalos que se alejan muy por encima del resto, siendo el máximo Mónaco con 19.196 personas por kilómetro cuadrado.

En cuanto a la variable “aglomeración urbana”, encontramos que el 75% de los países tienen el 32% de la población, o menos, en aglomeraciones urbanas. Como se puede observar en la ilustración 11, el 25% restante se distribuyen entre 32% y 55% aproximadamente. Existiendo países con valor muy superior como Emiratos Árabes, Australia, Japón, Kuwait y Singapur, casualmente todos países con ingresos altos, y este último con el 100% de la población viviendo en aglomeraciones urbanas.

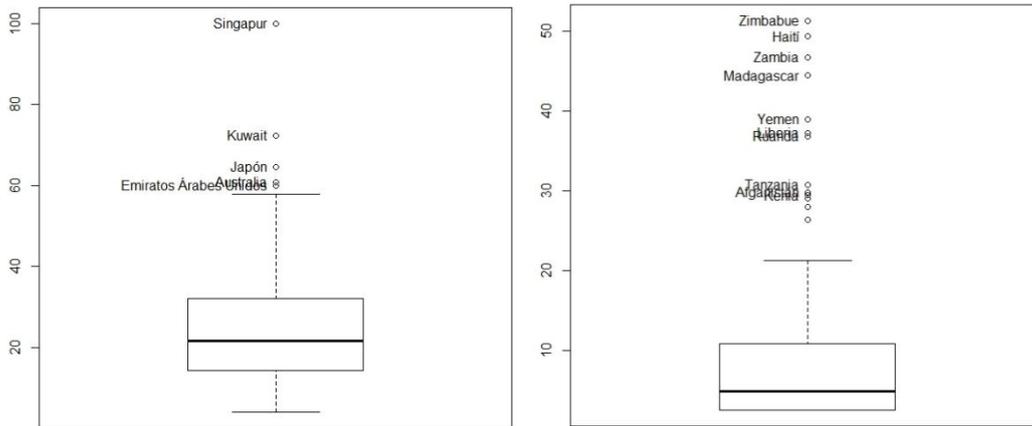


Ilustración 11 e Ilustración 12. Gráficos de caja y bigotes de las variables AGLO y DESNUT.

En el lado contrario, podemos ver en la ilustración 12 que todos los países considerados puntos anómalos en la variable “porcentaje de población con desnutrición” son países con ingresos bajos, por orden descendente: Zimbabwe, Haití, Zambia, Madagascar, Yemen, Liberia, Ruanda, Tanzania, Afganistán y Kenia. Estos se encuentran entre el 25% y el 50% de la población con desnutrición. Por otro lado, el 75% de los países tienen en su población el 10% o menos con desnutrición.

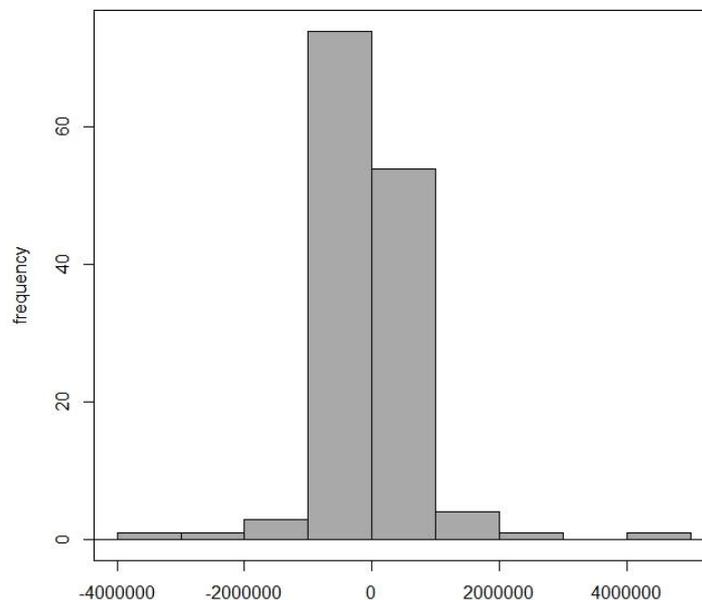


Ilustración 13. Histograma de la variable MIG.

La variable “migración neta” se concentra en los valores cercanos a 0, ya que la migración neta es la diferencia de los que entran menos los que se van. En la ilustración 13, los extremos son países donde, o la población interior del país está emigrando y no quieren entrar del extranjero, o la población del interior permanece en su país y existe mucha inmigración al país.

El 50% de los países en encuentran entre -77.000 y 78.000 de migración neta en su país. Por debajo de 0, encontramos puntos anómalos en orden ascendente empezando por Venezuela, con -3.266.243, seguido por India, Bangladés, China, Pakistán, Zimbabue, Indonesia, Sri Lanka, Rumanía y Filipinas. Y por encima de 0, encontramos puntos anómalos como EE.UU, en cabeza con 4.774.029, seguido por orden descendente por: Alemania, Turquía, Reino Unido, Canadá, Colombia, Rusia, Australia, Italia y Sudáfrica.

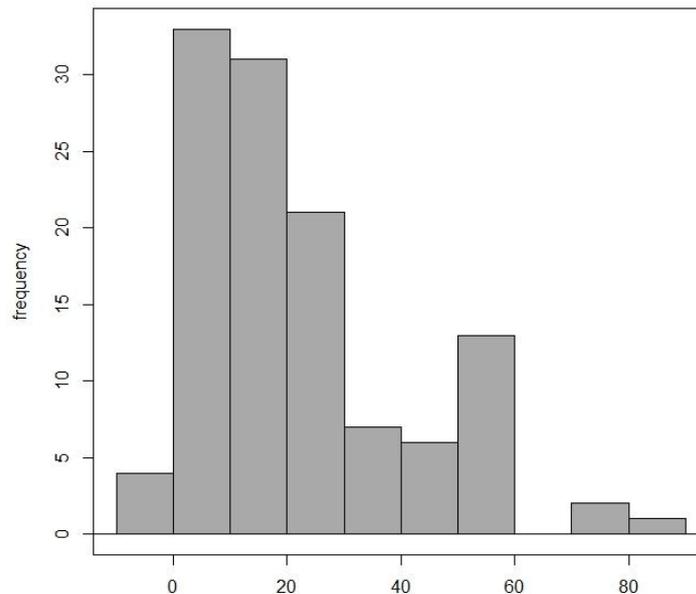


Ilustración 14. Histograma de la variable DIAS.

En cuanto a la variable días transcurridos desde la aparición hasta la aplicación de medidas de confinamiento ya sean laxas o estrictas, se puede observar en la ilustración 14 que gran parte de los países se concentran en los primeros 30 días, habiendo países que han aplicado medidas antes de su aparición, como El Salvador que lo hizo 6 días antes, y países muy por encima del resto como Japón, que no impuso medidas como tal de confinamiento pero si hubo cierres de algunos comercios, o Singapur y Tailandia que también son países que han tardado más en responder de lo habitual. Otros países directamente no aparecen porque nunca han impuesto medidas de confinamiento aunque haya aparecido el virus.

Como se puede observar en la ilustración 15, la mayoría de países han tenido una variación negativa del PIB del primer trimestre de 2020, concretamente el 50% de los países se encuentran entre -3,1 y -0,5 por ciento de decrecimiento del PIB respecto al trimestre anterior.

Islandia es un punto anómalo, y es el país con mayor decrecimiento con un valor de -7% del PIB en el primer trimestre. En el otro lado encontramos unos pocos países que sí que han obtenido una variación positiva del PIB.

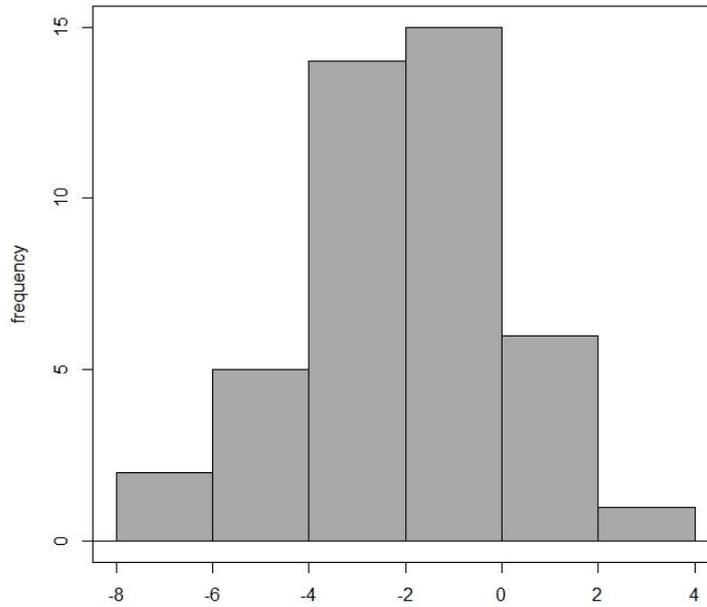


Ilustración 15. Histograma de la variable PIB.

En la ilustración 16, se concentran la mayoría de países en un aumento del paro entre 0 y 0,5, situándose el 50% de los países entre 0,0 y 0,6. No hay prácticamente ningún punto anómalo, ya que se distribuyen de forma regular, y aunque cabe pensar lo contrario, no ha habido ningún país con un aumento mayor del 2% de paro. Aparentemente Grecia es el país con menor variación del paro en el primer trimestre de 2020, de hecho es negativo por lo que son buenas noticias para ellos y pocos países que están en la misma situación.

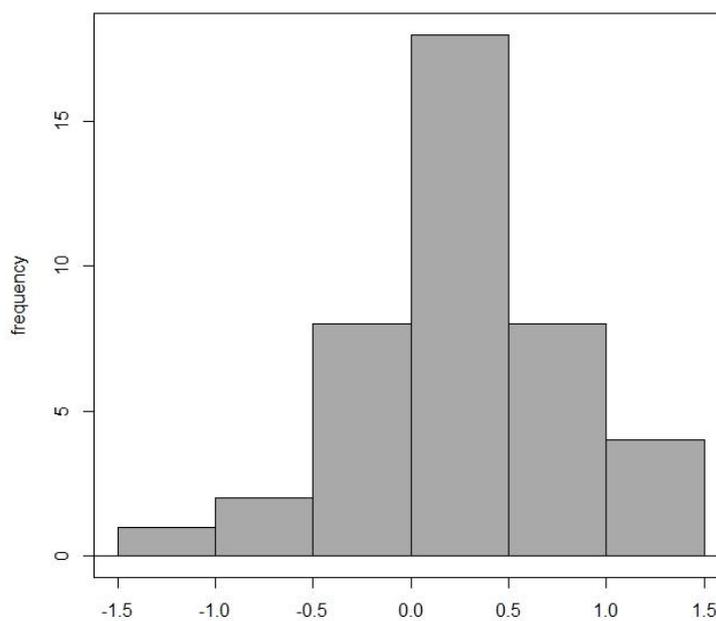


Ilustración 16. Histograma de la variable PARO.

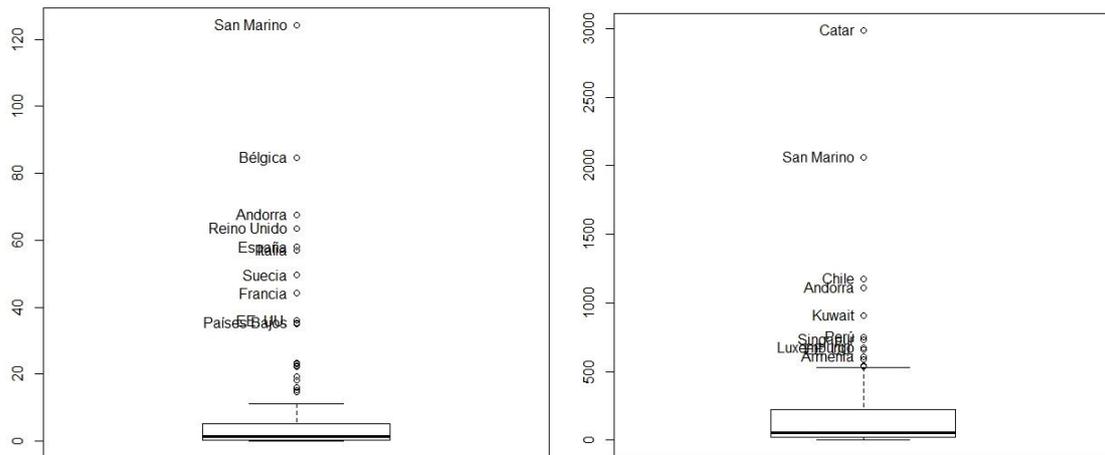


Ilustración 17 e Ilustración 18. Gráficos de caja y bigotes de las variables FALLE y INFECT.

La ilustración 17 nos muestra el número de muertes por cada 100.000 habitantes en cada país. Como ya se iba sabiendo, los países de Europa más afectados por el Covid-19 son puntos anómalos de la variable: San Marino con el mayor valor (124), Bélgica, Andorra, Reino Unido, España, Italia, Suecia, Francia, Países bajos, y fuera de Europa, EEUU. También es de destacar que la mayoría de los países se encuentran con valores bajos comparados con los puntos anómalos. El 50% de los países se encuentran entre 0,39 y 5,3 personas muertas por cada 100.000 habitantes a causa del Covid-19.

En la Ilustración 18 se muestran los números de casos confirmados por cada 100.000 habitantes por Covid-19 en cada país. Al igual que con la variable MUERTES, los países se concentran en valores bajos comparados con los puntos anómalos de la variable. El 50% de los países se encuentran entre 18,38 y 223 personas contagiadas por cada 100.000 habitantes. Catar es el país con mayor número de infectados por cada 100.000 habitantes con 10 veces más de contagiados que el resto (2990), seguido de San Marino, Chile, Andorra, Kuwait, Perú, Singapur, Luxemburgo, EE.UU y Armenia.

Estas diferencias entre las variables a explicar de muertos y contagiados ayudarán a sacar distintas conclusiones sobre por qué en algunos países existen más muertos que otros e igual con los infectados ya que puede ser que varíe la mortalidad según distintos países.

3.3.Análisis bivariante

En el análisis bivariante se analiza la relación que tienen las variables entre sí. La matriz de la Tabla 5 es una matriz de correlación que nos muestra de forma directa la relación de una variable respecto a otra. Con valores entre -1 y 1, indica que cuanto más se acerque a 1 mayor será la relación y si se aproxima a 0 la relación será nula. Los valores negativos indican que cuando aumenta una variable la otra disminuye.

La tabla de correlaciones es la siguiente:

Matriz de correlaciones

Tabla 5. Matriz de correlaciones de variables cuantitativas.

	AGLO	CAMAS	DENSI	DESNUT	DIAS	GASTO	INFECT	MED	MIG	MORTA	FALLE	URB	PIB	PARO
AGLO	1.0000	-0.109	-0.0071	0.0032	0.5293	-0.062	0.0707	-0.3347	0.202	-0.4690	-0.1101	0.265	0.1161	-0.0652
CAMAS	-0.1091	1.000	0.3164	-0.0658	0.2633	0.147	-0.4236	0.1514	-0.069	0.5871	-0.2244	0.022	0.0107	0.0465
DENSI	-0.0071	0.316	1.0000	-0.0461	0.1569	0.202	-0.0150	0.0424	-0.075	0.0053	0.3234	0.426	-0.2259	-0.2077
DESNUT	0.0032	-0.066	-0.0461	1.0000	-0.0498	-0.318	-0.2241	-0.317	-0.409	-0.1193	-0.2002	-0.235	-0.3931	0.3983
DIAS	0.5293	0.263	0.1569	-0.0498	1.0000	0.210	0.0023	-0.1936	0.420	-0.1221	0.2172	0.251	-0.1476	0.1530
GASTO	-0.0623	0.147	0.2015	-0.3183	0.2104	1.000	0.1695	0.5228	-0.036	0.0561	0.3893	0.485	-0.3158	0.1040
INFECT	0.0707	-0.424	-0.0150	-0.2241	0.0023	0.169	1.0000	-0.2878	0.381	-0.3425	0.5668	0.035	0.1833	0.0205
MED	-0.3347	0.151	0.0424	-0.3174	-0.1936	0.523	-0.2878	1.0000	-0.022	0.3175	0.0064	0.103	-0.3089	-0.1855
MIG	0.2017	-0.069	-0.075	-0.4090	0.4205	-0.036	0.3813	-0.0225	1.000	-0.0634	0.2272	-0.208	0.1894	-0.0904
MORTA	-0.4690	0.587	0.0053	-0.1193	-0.1221	0.056	-0.3425	0.3175	-0.063	1.0000	-0.0723	-0.222	-0.0124	0.0457
FALLE	-0.1101	-0.224	0.3234	-0.2002	0.2172	0.389	0.5668	0.0064	0.227	-0.0723	1.0000	0.171	-0.3571	-0.1050
URB	0.2645	0.022	0.4256	-0.2345	0.2505	0.485	0.0354	0.1032	-0.208	-0.2216	0.1710	1.000	0.0433	-0.1541
PIB	0.1161	0.011	-0.2259	-0.3931	-0.1476	-0.316	0.1833	-0.3089	0.189	-0.0124	-0.3571	0.043	1.0000	-0.0058
PARO	-0.0652	0.046	-0.2077	0.3983	0.1530	0.104	0.0205	-0.1855	-0.090	0.0457	-0.1050	-0.154	-0.0058	1.0000

Como podemos ver en la matriz de correlaciones de la tabla 5, existe muy poca correlación entre las variables variación del PIB, variación del paro, muertes e infectados por Covid-19, y el resto. Ninguna de ellas se acerca tan siquiera a 0,6. Esto significa que cuando aumenta o disminuye una variable, no afecta significativamente al cambio de otras variables. Por ejemplo, entre las variables muertes por Covid-19 y número de camas por cada 1000 personas, hay una correlación negativa igual a -0,22, por lo que cuando existe un mayor número de camas, en teoría hay un menor número de muertes, pero esta cifra no es significativa. Al igual que en la variable número de días desde que se descubre el primer caso en el país hasta que se aplica confinamiento.

También se ve que cuando hay una mayor migración neta (MIG) existe aparentemente un mayor número de infectados, aunque esta correlación (0,38) es mayor que la anterior, tampoco es significativa.

A continuación, se procederá a extraer las gráficas XY para las variables con una correlación en torno a 0,4 o mayor (señaladas en amarillo en la matriz de correlación). Esto permite una mejor visualización de la correlación entre variables.

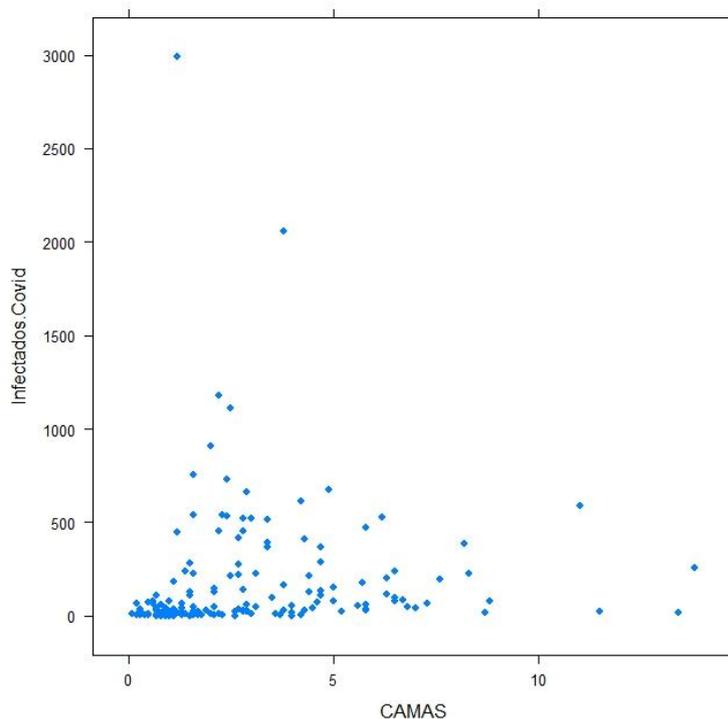


Ilustración 19. Gráfico XY de las variables CAMAS e INFECT.

Como se puede observar en la Ilustración 19, existe muy poca correlación entre variables. Se percibe un ligero aumento en la variable infectados por Covid-19 a medida que disminuye la variable número de camas, con un ligero repunte entre 2 y 3 camas por cada 1000 habitantes.

Por sacar algún sentido a esto, si a medida que hay menos camas por habitante en el país, hay más infectados, puede deberse a la población que no está hospitalizada en camas hospitalarias provocando mayor riesgo de contagio al resto de la población.

En cambio, se puede apreciar en la Ilustración 20 como existe una distribución distinta, concentrándose los puntos alrededor de 0 en ambas variables. Sin embargo, existen mayor número de infectados a medida que aumenta la variable migración neta, por lo que tiene sentido que si entra más población al país de la que sale se provocará mayor riesgo de contagio.

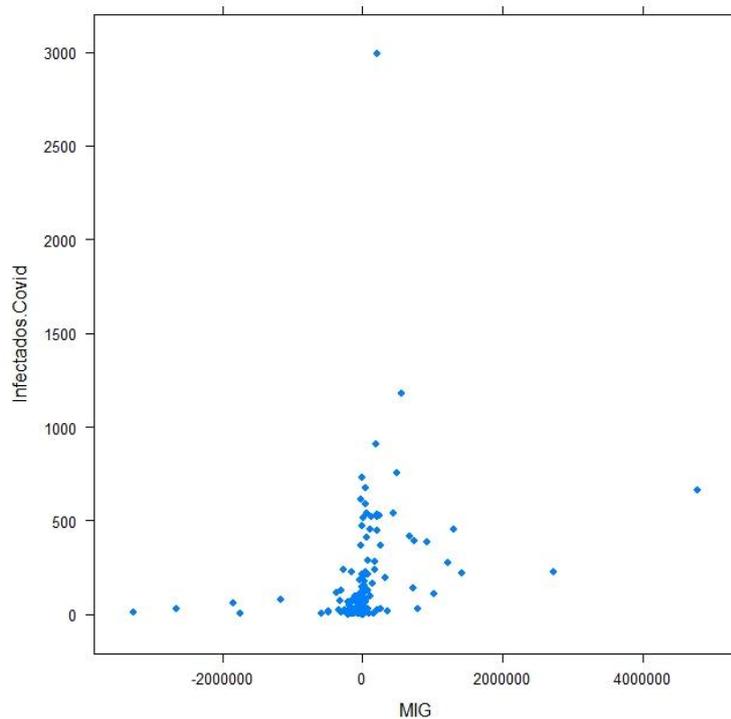


Ilustración 20. Gráfico XY de las variables MIG e INFECT.

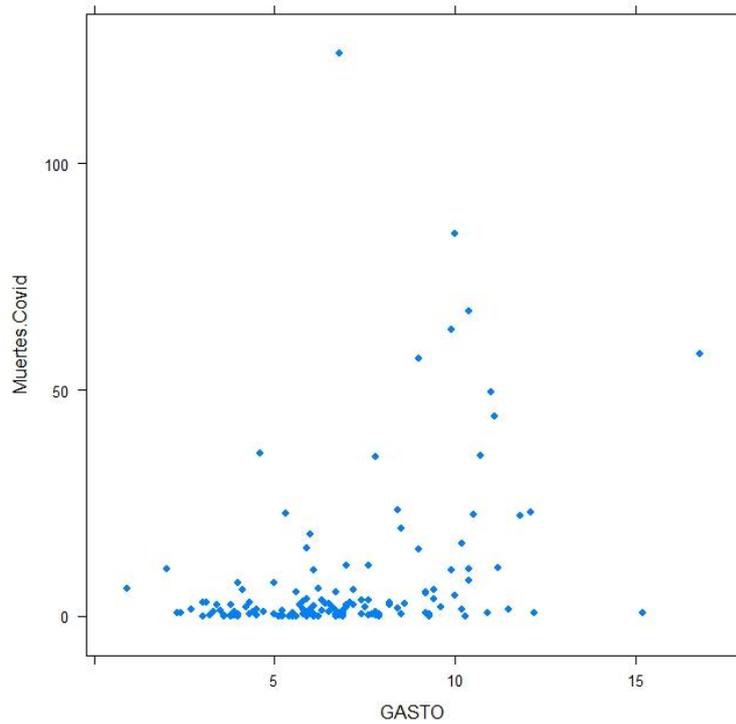


Ilustración 21. Gráfico XY de las variables GASTO y FALLE..

En este último caso se puede observar como la relación en sí no tiene sentido, ya que se aprecia una ligera relación positiva entre variables. Pero no tiene sentido pensar que a mayor gasto en salud por parte del país vaya a provocar mayor número de muertes por Covid-19, en todo caso se pensaría al contrario. Esto es debido a otros factores que influyen con más gravedad sobre la cantidad de fallecimientos por el Covid-19. Según comentarios de algunos especialistas en medicina, se puede deber a que este mayor gasto en salud se traduzca en una mayor capacidad para diagnosticar pacientes, y a su vez, mayor capacidad para contabilizar fallecimientos e infectados.

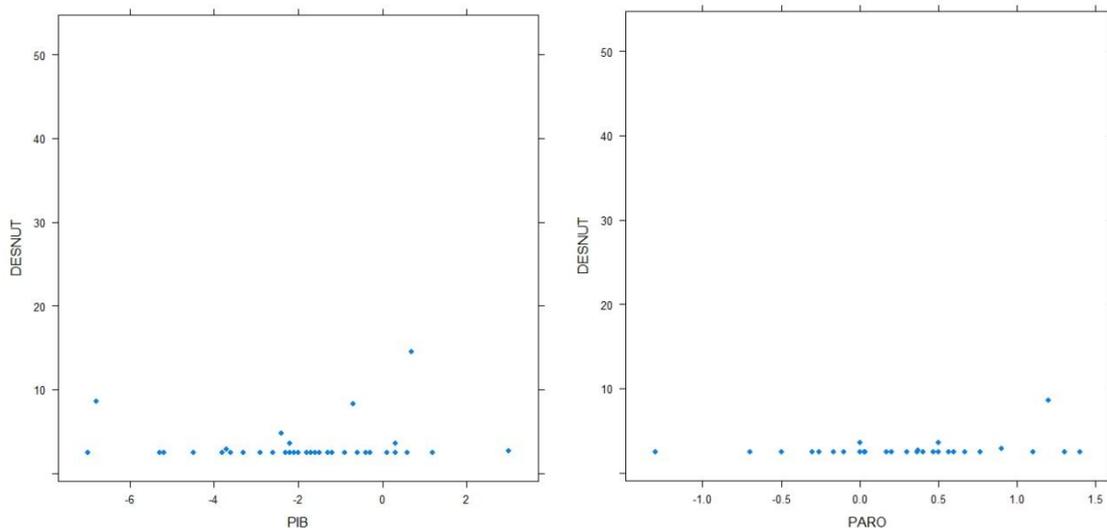


Ilustración 22 e Ilustración 23. Gráficos XY de la variable DESNUT respecto a las variables PIB y PARO.

En cuanto a las variables, variación trimestral del PIB y del paro, la mayor correlación que tienen es con la variable % de población con problemas de desnutrición. Al ser variables económicas, cabe pensar la poca relación con la salud y la nutrición. En ambos casos las gráficas XY carecen de forma, semejantes a una línea recta con pequeñas desviaciones. Sin embargo en el caso del PIB, la correlación es negativa, apenas apreciándose en la Ilustración 22, por lo que en teoría, a mayor desnutrición en el país la variación del PIB es menor.

Comparando ambas gráficas de las Ilustraciones 22 y 23, apenas existe diferencia visual, en cambio la variable variación del paro posee la misma correlación que el PIB respecto al % de desnutrición en el país pero en positivo. Lo que quiere decir que a medida que hay mayor desnutrición en el país la variación del paro es mayor.

Por lo que en este caso podemos afirmar que la desnutrición afecta negativamente a la variación del PIB y del paro de un país, aunque no significativamente. Hay que tener en cuenta la menor cantidad de datos de la muestra para estas dos últimas variables, ya que no son tan representativas como otras variables en las que la muestra es mucho mayor.

Para analizar la relación existente entre las variables cualitativas junto a las variables a explicar se procede de una forma diferente. Se utilizarán las primeras para clasificar por grupos y así comprobar la influencia/impacto en los distintos grupos.

A continuación se procederá al análisis por medio de diagrama de caja y bigotes de las variables a explicar, dentro de los grupos de cada variable cualitativa.

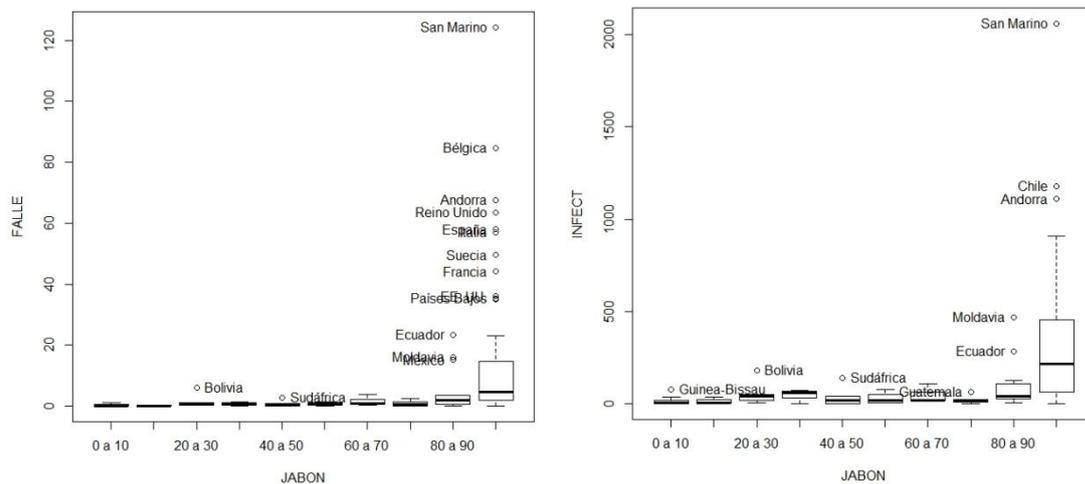


Ilustración 24 e Ilustración 25. Gráficos de caja y bigotes de las variables FALLE e INFECT en grupos de la variable JABON.

A partir de las Ilustraciones 24 y 25 podemos obtener información relevante sobre el nivel de incidencia según el grupo de porcentaje de población del país con instalaciones básicas de agua e higiene. A simple vista se puede observar como los países más afectados tanto en

fallecimientos como en infectados por el Covid-19 por cada 100.000 habitantes son los países con mayores recursos higiénicos. Los fallecimientos en los países se sitúan al mínimo de la gráfica hasta a partir del 80% de población con instalaciones básicas de higiene, e igual ocurre con los infectados. Los puntos anómalos que se encuentran en las variables principales, de forma individual, son prácticamente los mismos que se encuentran ahora en el último tramo de las gráficas (del 90 al 100%). Puntos anómalos de fallecimientos en orden ascendente de higiene: de 20 a 30% Bolivia; de 40 a 50% Sudáfrica; de 80 a 90% Ecuador, Moldavia y México; y por último, y más afectados, San Marino, Bélgica, Andorra, Reino Unido, España, Italia, Suecia, Francia, EE.UU y Países Bajos. Los puntos anómalos en los infectados aparecen más en cada grupo: de 0 a 10% se encuentra Guinea-Bissau; de 20 a 30% vuelve a aparecer Bolivia; de 40 a 50% Sudáfrica; de 70 a 80% Guatemala; de 80 a 90% de nuevo Moldavia y Ecuador; y finalmente, San Marino, Chile y Andorra.

Por tanto, podemos afirmar que la Higiene como tal no afecta al nivel de muertes e infectados por Covid-19, sino que puede estar influenciado por otros motivos asociados al nivel económico del país.

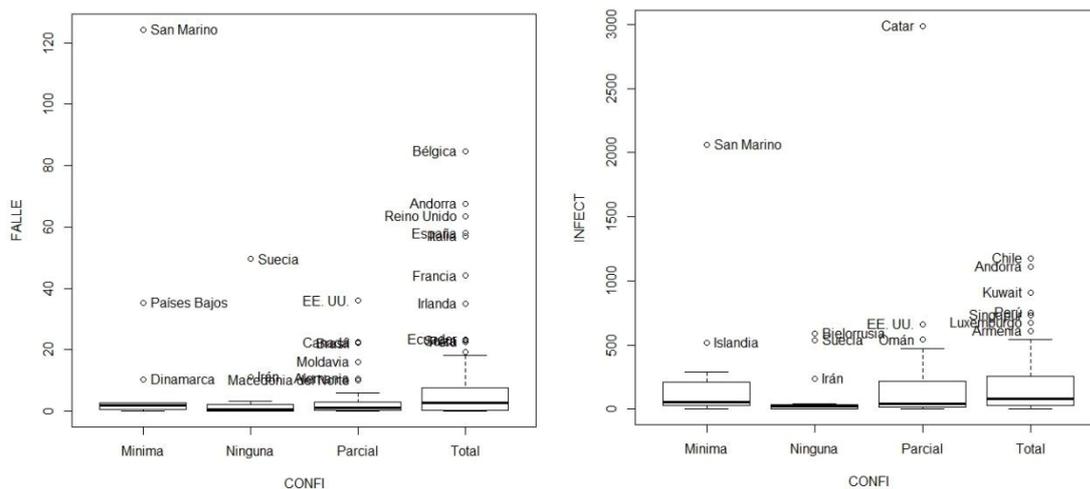


Ilustración 26 e Ilustración 27. Gráficos de caja y bigotes de las variables FALLE e INFECT en grupos de la variable CONFI.

En las Ilustraciones 26 y 27 se puede observar la relación entre las muertes e infectados de un país y la medida de confinamiento aplicada en dicho país. En cuanto a las muertes según la medida de confinamiento, se observa como los países que han aplicado medidas de confinamiento totales en el país son los más afectados respecto al resto, habiendo países con mayor impacto como San Marino el cuál aplicó medidas de confinamiento mínimas.

En el caso de los infectados es diferente, el cual es más regular entre las medidas de confinamiento, Siendo Catar el más afectado aplicando medidas de confinamiento parciales.

Según los resultados, podemos decir que al aplicar medidas de confinamiento restrictivas no repercute en un menor impacto en cuanto a muerte e infectados, ya que los países que no han aplicado medidas de confinamiento no han sido los más afectados en proporción a otros

países que si han aplicado medidas. Es posible que funcione como atenuante ante una progresión continua de propagación del virus pero no para evitar un daño ya causado y en proceso emergente.

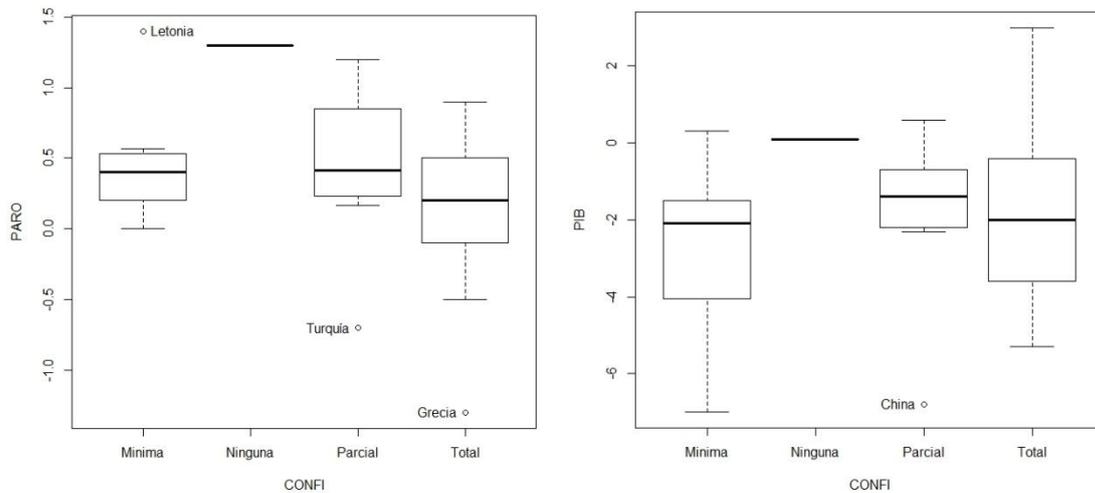


Ilustración 28 e Ilustración 29. Gráficos de caja y bigotes de las variables PARO e PIB en grupos de la variable CONFI.

Al comparar las variables de paro y PIB con las medidas de confinamiento se observan diferencias entre sí. En la Ilustración 28 se compara cuanto a variado el paro según la medida de confinamiento aplicada, en la que el grupo más afectado ha sido los países que han aplicado confinamiento parcial, seguido de los que han aplicado un confinamiento mínimo y los que han aplicado confinamiento total.

En cuanto a la variación del PIB, se puede observar en la Ilustración 29 como los países más perjudicados son el grupo que ha aplicado confinamiento mínimo, seguido de los que han aplicado confinamiento total, y parcial. Donde cabe pensar lo contrario, los países que han aplicado confinamiento mínimo, debería tener una variación positiva mayor del PIB, y una variación menor del paro.

Debido que existe una amplia diferencia entre los países más desarrollados y los menos desarrollados, es conveniente separar por grupos los países según nivel de ingresos y así poder analizar si son los mismos factores los que influyen en los países o si existen correlaciones mayores entre factores. Por ello, se dividirán los países en tres grupos: ingresos altos, medios y bajos, expuestos en la Tabla 6:

Tabla 6. Países según nivel económico.

Países con ingresos altos	Países con ingresos medios	Países con ingresos bajos
Suiza	Costa Rica	Sri Lanka
Noruega	Rusia	El Salvador
Luxemburgo	Argentina	Argelia
Islandia	Malasia	Filipinas
EE. UU.	China	Cabo Verde
Catar	Maldivas	Suazilandia

Dinamarca	Turquía	Bolivia
Irlanda	México	Ucrania
Singapur	Bulgaria	Túnez
Suecia	Brasil	Marruecos
Australia	Montenegro	Papua Nueva Guinea
Países Bajos	Kazajistán	Egipto
Austria	Rep. Dominicana	Honduras
Finlandia	Botsuana	Ghana
Alemania	Libia	India
Bélgica	Líbano	Nigeria
Canadá	Cuba	Bangladés
Emiratos Árabes Unidos	Tailandia	Nicaragua
Israel	Gabón	Uzbekistán
Nueva Zelanda	Serbia	Kenia
Francia	Perú	Pakistán
Reino Unido	Colombia	Camerún
Japón	Guinea Ecuatorial	Zambia
Italia	Bosnia y Herzegovina	Zimbabue
Kuwait	Ecuador	Benín
Corea del Sur	Sudáfrica	Kirguistán
Brunéi	Macedonia del Norte	Nepal
Bahamas	Fiyi	Tanzania
España	Irak	Guinea
Malta	Surinam	Yemen
Eslovenia	Paraguay	Mali
Estonia	Irán	Etiopía
Portugal	Jamaica	Guinea-Bissau
Arabia Saudí	Albania	Ruanda
Rep. Checa	Guyana	Burkina Faso
Grecia	Georgia	Haití
Lituania	Armenia	Gambia
Letonia	Guatemala	Sudán
Barbados	Azerbaiyán	Liberia
Trinidad y Tobago	Belice	Afganistán
Antigua y Barbuda	Jordania	Madagascar
Uruguay	Tonga	Malawi
Hungría	Samoa	Burundi
Omán	Indonesia	
Polonia		
Chile		
Panamá		
Croacia		
Venezuela		
Rumanía		

En el grupo de países con ingresos altos se encuentran los países con más de 12.000 de INB per cápita (50 países). En el de ingresos medios, se agrupan los países entre 4.000 y 12.000 de INB per cápita (44 países). Y en el grupo de países con ingresos bajos se encuentran los países con menos de 4.000 de INB per cápita (43 países).

MATRIZ DE CORRELACIONES DE LOS PAÍSES CON INGRESOS BAJOS

Tabla 7. Matriz de correlaciones de variables cuantitativas de los países con ingresos bajos.

	AGLO	CAMAS	DENSI	DESNUT	DIAS	FALLE	GASTO	INFECT	MED	MIG	MORTA	URB
AGLO	1.000	-0.167	-0.099	-0.055	-0.282	0.5511	-0.137	0.518	-0.060	0.011	0.1585	0.49
CAMAS	-0.167	1.000	-0.174	-0.271	-0.190	0.1064	0.038	0.083	0.726	0.203	0.5042	0.33
DENSI	-0.099	-0.174	1.000	-0.022	0.091	-0.1393	-0.329	0.027	-0.079	-0.640	-0.1014	-0.25
DESNUT	-0.055	-0.271	-0.022	1.000	-0.269	-0.2841	0.372	-0.283	-0.595	0.025	-0.0115	-0.52
DIAS	-0.282	-0.190	0.091	-0.269	1.000	-0.1474	-0.199	-0.195	-0.036	-0.322	-0.0563	-0.41
FALLE	0.551	0.106	-0.139	-0.284	-0.147	1.0000	-0.270	0.926	0.382	0.041	0.0095	0.48
GASTO	-0.137	0.038	-0.329	0.372	-0.199	-0.2704	1.000	-0.360	-0.294	0.306	-0.0256	-0.16
INFECT	0.518	0.083	0.027	-0.283	-0.195	0.9260	-0.360	1.000	0.346	-0.038	0.0523	0.37
MED	-0.060	0.726	-0.079	-0.595	-0.036	0.3824	-0.294	0.346	1.000	0.059	0.1995	0.54
MIG	0.011	0.203	-0.640	0.025	-0.322	0.0411	0.306	-0.038	0.059	1.000	0.0471	0.26
MORTA	0.159	0.504	-0.101	-0.011	-0.056	0.0095	-0.026	0.052	0.200	0.047	1.0000	0.18
URB	0.485	0.332	-0.255	-0.522	-0.410	0.4783	-0.159	0.374	0.539	0.259	0.1754	1.00

Como se puede comprobar en la Tabla 7, las correlaciones han mejorado ligeramente respecto a la inicial, eliminándose por sí solo el sesgo en cuanto al nivel económico reflejado en las variables de número de médicos y gasto en salud (como se podrá observar en el análisis multivariante).

Además, las mayores correlaciones entre variables de las pertenecientes a las variables de fallecimientos e infectados coinciden entre sí. La variable, porcentaje de población en aglomerados urbanos, es la que más significatividad tiene entre los países más pobres con un valor en torno al 0,5 de correlación con las variables de interés. Mientras que la siguiente variable con mayor correlación es el porcentaje de población urbana con un valor aproximado a 0,48 con fallecidos y 0,37 con infectados.

En el análisis multivariante estas variables entran dentro del modelo como las más explicativas, por lo que se puede confirmar que dichas variables influyen de forma significativa en el nivel de infectados y contagiados por Covid-19 en los países. Hay que tener en cuenta que los países de ingresos bajos son los menos afectados por Covid-19.

MATRIZ DE CORRELACIONES DE LOS PAÍSES CON INGRESOS MEDIOS

Tabla 8. Matriz de correlaciones de variables cuantitativas de los países con ingresos medios.

	AGLO	CAMAS	DENSI	DESNUT	DIAS	FALLE	GASTO	INFECT	MED	MIG	MORTA	PARO	PIB	URB
AGLO	1.0000	-0.995	0.0023	-0.261	0.6316	0.7599	-0.703	0.67	-0.747	0.283	-0.902	-0.585	-0.014	0.359
CAMAS	-0.9949	1.000	-0.0492	0.171	-0.6640	-0.7471	0.706	-0.59	0.772	-0.185	0.918	0.501	0.110	-0.293
DENSI	0.0023	-0.049	1.0000	0.758	0.7767	-0.6092	-0.656	-0.25	-0.666	-0.420	-0.433	0.408	-0.702	-0.900
DESNUT	-0.2614	0.171	0.7581	1.000	0.4177	-0.5297	-0.148	-0.78	-0.295	-0.908	-0.089	0.881	-0.954	-0.924
DIAS	0.6316	-0.664	0.7767	0.418	1.0000	0.0043	-0.954	0.23	-0.987	-0.140	-0.904	-0.059	-0.547	-0.469
FALLE	0.7599	-0.747	-0.6092	-0.530	0.0043	1.0000	-0.073	0.51	-0.157	0.281	-0.422	-0.539	0.261	0.768
GASTO	-0.7035	0.706	-0.6562	-0.148	-0.9543	-0.0731	1.000	-0.50	0.967	-0.162	0.918	0.337	0.272	0.269
INFECT	0.6661	-0.588	-0.2518	-0.785	0.2306	0.5053	-0.497	1.00	-0.345	0.900	-0.494	-0.985	0.668	0.639
MED	-0.7470	0.772	-0.6662	-0.295	-0.9871	-0.1565	0.967	-0.35	1.000	0.049	0.960	0.181	0.460	0.325
MIG	0.2826	-0.185	-0.4204	-0.908	-0.1402	0.2807	-0.162	0.90	0.049	1.000	-0.076	-0.941	0.905	0.692
MORTA	-0.9023	0.918	-0.4331	-0.089	-0.9041	-0.4217	0.918	-0.49	0.960	-0.076	1.000	0.353	0.313	0.064
URB	0.3589	-0.293	-0.8997	-0.924	-0.4692	0.7684	0.269	0.64	0.325	0.692	0.064	-0.750	0.802	1.000

En la Tabla 8 se puede observar una mejora muy elevada de las correlaciones. En este intervalo de países se pueden encontrar más diferencias entre ellas, por ello, los resultados obtenidos son más significativos que en los países con muchas semejanzas.

La mayor correlación obtenida es la migración neta, de nuevo coincidente en los modelos del análisis multivariante como se verá posteriormente, con un valor de 0,9, lo que lo convierte en una variable que está totalmente relacionada con el nivel de infectados de los países. Esto de momento solo ocurre en los países con ingresos medios, ya que en los países de ingresos bajos puede deberse a que no existe una migración neta elevada y por ello influyen más otros factores.

En cuanto a las demás variables con mayor correlación respecto a la variable número de infectados, encontramos la misma situación que en los países de ingresos bajos, además de la variable prevalencia de desnutrición. En esta última existe correlación alta y negativa, por lo tanto habría que comprobar a que se debe este resultado ya que no cabe pensar que a mayor prevalencia de desnutrición menor número de infectados. En cambio las siguientes variables con mayor correlación vuelven a ser porcentaje de población en aglomerados y porcentaje de población urbana, en este caso con un valor superior, de 0,6 en ambas variables.

En la columna de correlaciones de fallecimientos por Covid-19 ocurre igual. Se han obtenido mejores resultados y las dos variables con mayor correlación han vuelto a ser las variables porcentaje de población en aglomerados y porcentaje de población urbana, con un valor de 0,76 para ambas variables. La única diferencia es que la variable número de camas tiene una correlación negativa y alta, lo cual tiene un efecto positivo, porque a mayor número de camas menor número de fallecidos, por lo que una insuficiencia de camas para cuidados necesarios puede derivar en mayor número de muertes en países donde si pueden contabilizar adecuadamente.

Por lo tanto, se puede confirmar que las variables, señaladas en amarillo en la matriz de la Tabla 8, son influyentes y tienen relación con las variables explicadas en los países de ingresos bajos y medios.

MATRIZ DE CORRELACIONES DE LOS PAÍSES CON INGRESOS ALTOS

Tabla 9. Matriz de correlaciones de variables cuantitativas de los países con ingresos altos.

	AGLO	CAMAS	DENSI	DESNUT	DIAS	FALLE	GASTO	INFECT	MED	MIG	MORTA	PARO	PIB	URB
AGLO	1.0000	-0.038	0.0038	0.0991	0.550	-0.115	0.0017	0.077	-0.302	0.2220	-0.453	-0.0066	0.1443	0.271
CAMAS	-0.0380	1.000	0.3173	-0.2245	0.311	-0.269	0.0347	-0.495	0.018	-0.0984	0.544	-0.0270	-0.0142	0.026
DENSI	0.0038	0.317	1.0000	-0.1964	0.116	0.306	0.1690	-0.058	0.011	-0.1070	0.020	-0.2614	-0.2313	0.451
DESNUT	0.0991	-0.225	-0.1964	1.0000	-0.120	-0.027	-0.0690	0.715	-0.530	0.0019	-0.314	0.0957	0.5143	0.168
DIAS	0.5499	0.311	0.1160	-0.1203	1.000	0.183	0.2061	-0.061	-0.221	0.4552	-0.021	0.2118	-0.1081	0.274
FALLE	-0.1150	-0.269	0.3056	-0.0271	0.183	1.000	0.3297	0.535	-0.094	0.1854	-0.101	-0.0756	-0.4611	0.134
GASTO	0.0017	0.035	0.1690	-0.0690	0.206	0.330	1.0000	0.070	0.383	-0.1636	-0.174	0.1314	-0.4930	0.503
INFECT	0.0770	-0.495	-0.0583	0.7154	-0.061	0.535	0.0699	1.000	-0.445	0.3248	-0.449	0.1246	0.1622	-0.016
MED	-0.3020	0.018	0.0115	-0.5296	-0.221	-0.094	0.3833	-0.445	1.000	-0.1389	0.109	-0.2635	-0.5456	0.042
MIG	0.2220	-0.098	-0.1070	0.0019	0.455	0.185	-0.1636	0.325	-0.139	1.0000	-0.100	0.1437	0.0026	-0.341
MORTA	-0.4525	0.544	0.0199	-0.3141	-0.021	-0.101	-0.1743	-0.449	0.109	-0.0999	1.000	-0.0907	-0.1764	-0.323
PARO	-0.0066	-0.027	-0.2614	0.0957	0.212	-0.076	0.1314	0.125	-0.263	0.1437	-0.091	1.0000	0.2864	-0.069
PIB	0.1443	-0.014	-0.2313	0.5143	-0.108	-0.461	-0.4930	0.162	-0.546	0.0026	-0.176	0.2864	1.0000	-0.070
URB	0.2708	0.026	0.4513	0.1679	0.274	0.134	0.5029	-0.016	0.042	-0.3408	-0.323	-0.0686	-0.0699	1.000

Finalmente en la Tabla 9 los resultados de los países ricos han cambiado totalmente respecto a los anteriores.

En la columna de la variable de infectados: la variable prevalencia de desnutrición se contradice con el resultado de los países con ingresos medios, por lo que no se puede concluir nada respecto a dichos resultados; en cuanto a las variables número de médicos y camas existe una correlación negativa en torno a 0,47, por lo que tener mayor número de camas y médicos conlleva un menor número de contagiados. Si nos fijamos en los fallecidos también ocurre lo mismo con estas dos últimas variables, y en los países con ingresos medios también se cumple, entonces sí que tiene un efecto positivo en cuanto al nivel de fallecidos e infectados en países con ingresos altos y medios.

Por otro lado, en la columna de la variable de fallecidos: la variable con mayor correlación es respecto a la variable variación de PIB, con un valor de -0,46, lo que significa que en los países más afectados (con ingresos altos) existe un perjuicio en el PIB aparentemente generado por los fallecidos en el país o por sus medidas aplicadas; por otro lado existen dos variables con poca significatividad, que son, densidad de población y gasto en salud, ambas positivas y con un valor de 0,3, pudiendo deberse al mayor riesgo poblacional y a la capacidad de confirmar casos, respectivamente.

Hay que tener en cuenta que los países con altos ingresos poseen altos niveles y muy similares de población urbana, migración neta y población en aglomerados, por ello no ha existido una relación significativa respecto a estas variables. Y al contrario, los países con ingresos bajos y medios tienen niveles similares de número de médicos.

3.4. Análisis multivariante

Para analizar el conjunto de variables como un modelo de regresión de variables independientes con una dependiente se utilizará el método estadístico de árboles de decisión CART (en inglés, Classification And Regression Trees). Como el nombre indica, CART es una técnica en la que se pueden obtener árboles de clasificación y de regresión, usándose de clasificación cuando nuestra variable objetivo es discreta, y de regresión cuando es continua.

La finalidad será averiguar que variables son las más representativas para la fiabilidad del modelo y dividir los datos de esas variables por grupos a modo de filtro. El árbol construido por el modelo será más completo cuantas más ramas tenga sin perder pureza y por lo tanto mayor segregado y explicado será la variable.

Las variables dependientes son el número de fallecimientos e infectados confirmados por cada 100.000 habitantes por Covid-19. El restante son las variables independientes para el estudio del modelo.

Las variables variación trimestral de paro y PIB, se utilizarán para complementar las conclusiones que se obtengan del análisis previo con todos los países, ya que en estas variables solo hay datos de los países con altos ingresos.

3.4.1. Modelos de la variable dependiente: número de fallecidos por Covid-19

A continuación, la primera variable dependiente para el análisis del modelo será el número de fallecidos. Se utilizará toda la muestra para el modelo. La regla obtenida es la siguiente:

- 1) root 147 42954.620 7.630884
- 2) MED < 2.58685 97 2218.371 2.461340 *
- 3) MED >= 2.58685 50 33115.050 17.659800
- 6) MIG < 161318 36 15952.560 9.808889
- 12) DENSI < 234.4628 29 1385.009 5.487931 *
- 13) DENSI >= 234.4628 7 11782.960 27.710000 *
- 7) MIG >= 161318 14 9237.753 37.847860 *

La primera línea es por donde el árbol comienza, en la raíz (root), y las siguientes son las ramas. Lo que significa cada rama es: Nodo) criterio de división junto al punto de corte, el siguiente número es el número de observaciones en cada rama antes de la división (n), el siguiente es el error o desviación (RRS), y por último la media de la predicción final para cada rama (yval). En los asteriscos es donde acaba la rama del árbol, o también conocido como hoja, aunque sean más bien parecidos a una raíz. El árbol se extiende comenzando desde arriba hacia abajo, y van surgiendo ramas, o grupos, según el punto de corte óptimo para la variable que mejor explica el modelo. A través del diagrama del árbol se podrá interpretar mejor:

¿Qué afecta al número de muertes por Covid-19?

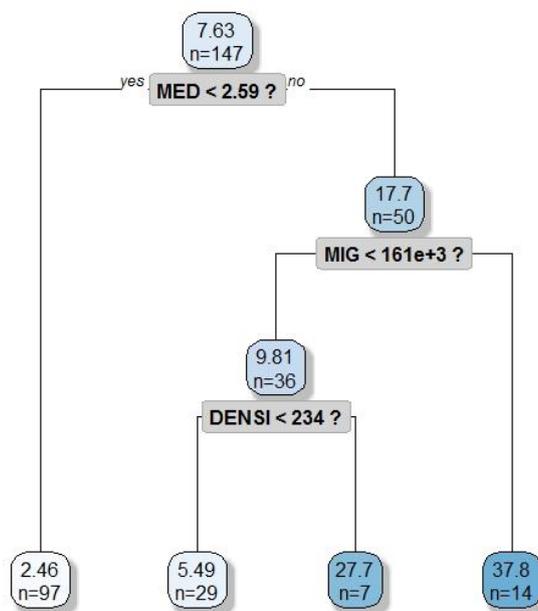


Ilustración 30. Árbol de decisión del modelo 1 de la variable dependiente FALLE.

Aquí se ve algo curioso en el árbol de la Ilustración 30, la primera variable que mejor explica el número de muertes es el número de médicos, pero a lo contrario que cabe pensar, a mayor número de médicos es mayor el número de muertes, explicándose a través de 97 países que con menos de 2,59 médicos por cada 1.000 habitantes los países obtienen de media 2,46 muertes por cada 100.000 habitantes. Esto quiere decir que el que haya mayor atención sanitaria no influye en las muertes, sino que está más motivado por factores económicos ya que si nos fijamos en los 50 países restantes, obtienen que con más de 2,59 médicos, el número de muertes es de 17,7 de media.

Además, como se puede observar en la base de datos, son estos últimos la mayoría países del norte, con altos ingresos económicos, siendo precisamente EE.UU. el punto de corte con 2,59 médicos. Ello explica que el tener más médicos no influye positivamente en el impacto de las muertes.

Algo más normal ocurre en el siguiente nodo, los 50 países con más de 2,59 médicos se dividen también por la migración neta de los países, pero aquí si tiene más sentido que al haber mayor migración neta existe un mayor número de muertes. Con más de 161.000 de migración neta se encuentran los 14 países más afectados de media, con 37,8 muertes por cada 100.000. Estos países son: EE.UU, Canadá, Reino Unido, Francia, Bélgica, Australia, Rusia, España, Italia, Alemania, Suiza, Austria y Suecia.

En el otro lado, los 36 países restantes con menos de 161.000 de migración neta se vuelven a subdividir según la densidad de población, con el punto de corte de 234 personas por kilómetro cuadrado. Encontrándose por debajo de esta cifra 29 países con 5,49 muertes, y por encima, 7 países con 27,7 muertes de media por cada 100.000 habitantes respectivamente.

Por lo tanto, en este árbol se han obtenido 3 nodos de las variables que mejor explican el número de muertes y las condiciones de esas variables que determinan los puntos de corte, a fin de llegar a las 4 terminaciones finales u hojas.

A continuación, se comprobará con la agrupación de solamente 45 países, de los que se tienen los datos de todas las variables para que tengan mayor peso sobre el resto, las variables PIB y paro, debido a que faltan datos de los países en estas variables. Estos países son:

- | | | | |
|-----------|-----------------|-----------------|--------------|
| – Grecia | – Portugal | – Nueva Zelanda | – Bulgaria |
| – Turquía | – Países Bajos | – Dinamarca | – Luxemburgo |
| – Italia | – Rep. Checa | – EE. UU. | – Islandia |
| – Francia | – Austria | – Alemania | – España |
| – Israel | – Polonia | – Chile | – Eslovenia |
| – Croacia | – Corea del Sur | – Eslovaquia | – Canadá |
| – Bélgica | – Japón | – Hungría | – Suiza |
| – Letonia | – Finlandia | – Indonesia | – Rumanía |

- Colombia
- Brasil
- Noruega
- México
- China
- Suecia
- Reino Unido
- Irlanda
- India
- Estonia
- Malta
- Australia
- Lituania

Además estos 45 países superan los puntos de corte de las dos principales variables anteriormente comentadas, número de médicos (con 3,17 de media) y migración neta (con 258.000 de media). La regla obtenida que se obtiene es:

- 1) root 45 18323.8900 15.589110
- 2) MIG < 77999 24 579.8137 4.762917 *
- 3) MIG >= 77999 21 11716.3100 27.961900
- 6) AGLO >= 31.32434 8 1181.2830 13.853750 *
- 7) AGLO < 31.32434 13 7962.8220 36.643850 *

En estos resultados ha desaparecido la variable médicos y densidad de población, en cambio se ha mantenido la variable migración neta y ha aparecido otra nueva variable que es población en aglomerados urbanos de más de 1 millón de personas, similar a densidad de población.

El punto de corte para la variable migración neta ha disminuido respecto al anterior modelo y la otra se encuentra en 31% de la población en aglomerados urbanos.

¿Qué afecta al número de muertes por Covid-19?

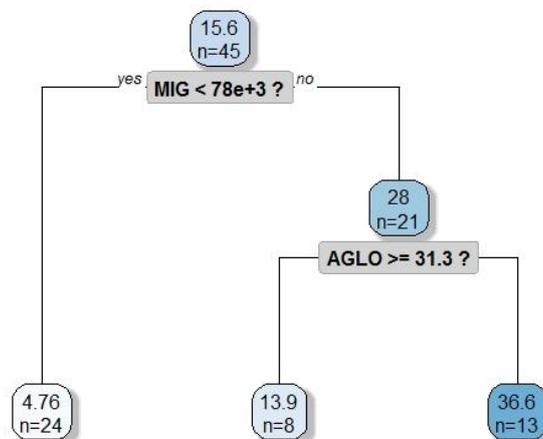


Ilustración 31. Árbol de decisión del modelo 2 de la variable dependiente FALLE.

En la Ilustración 31, El árbol comienza a dividirse en 24 países que han tenido menos de 78.000 de migración neta que han sufrido de media 4,76 muertos por cada 100.000 habitantes, mientras que los 21 países que han tenido más se subdividen por porcentaje de población en aglomerados. Los 8 países que tienen menos de 31,3% de población en aglomerados han obtenido, de media, 13,9 muertes por cada 100.000 habitantes, mientras que los que han tenido más han llegado a 36,6 muertes de media.

Prácticamente todos los países seleccionados para este modelo son países con ingresos altos, y además son muchos de estos los más afectados a nivel mundial. Se puede comprobar en la base de datos como la variación del PIB de todos estos países para los que tenemos datos han tenido una variación trimestral negativa, de media -2%, por lo que efectivamente ha tenido un impacto económico en la mayoría de países. El país más perjudicado de este grupo ha sido china con una variación del PIB de -6% pero casualmente también es el país con menos fallecidos por cada 100.000 habitantes.

En el caso del paro, también ha tenido un impacto social en la mayoría de países. De los 41 países que tenemos datos, el paro ha aumentado de media 0,3%, pero curiosamente de estos países solo 8 han tenido una variación negativa del paro, y por lo tanto no han resultado perjudicados en ese sentido, de hecho, Italia, Francia y Bélgica, que han reducido su paro han sido de los más perjudicados en este grupo.

Ya que eliminando a los países con menos ingresos la variable médicos ha desaparecido, podemos afirmar que esta variable está íntimamente relacionado con esto, y por ello, eliminaremos del primer modelo la variable médicos y gasto en salud de las variables para comprobar si existen mayores diferencias en el modelo. La regla obtenida es la siguiente:

- 1) *root 147 42954.6200 7.630884*
- 2) *MIG< 79999 110 5661.0120 2.957727*
- 4) *URB< 80.129 91 917.4015 2.026923 **
- 5) *URB>=80.129 19 4287.1560 7.415789 **
- 3) *MIG>=79999 37 27749.6400 21.524050*
- 6) *DENSI< 238.4821 30 9374.3260 15.820330*
- 12) *MORTA< 7.8 18 2086.0020 9.810000 *↓*
- 13) *MORTA>=7.8 12 5662.7390 24.835830 **
- 7) *DENSI>=238.4821 7 13216.6000 45.968570 **

Estos resultados parecen más completos debido a que hay más ramas, y por tanto, más variables en el modelo que lo explican.

La variable migración sigue siendo la primer nodo que divide las ramas, pero esta vez el punto de corte es prácticamente el mismo que en el anterior modelo. El punto de corte de la variable población urbana se encuentra en 80%. La variable densidad de población ha aparecido de nuevo con un valor de corte muy cercano al del primer modelo. Y por último, otra variable

nueva es tasa de mortalidad cruda por cada 1.000 habitantes, que se encuentra en 7,8 de punto de corte. El árbol es el siguiente:

¿Qué afecta al número de muertes por Covid-19?

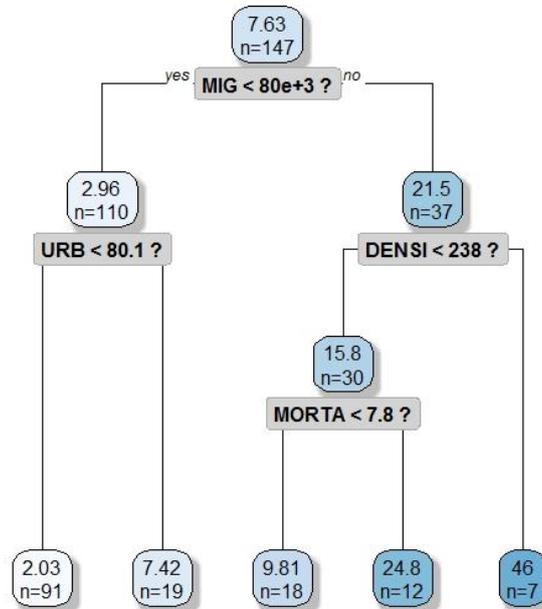


Ilustración 32. Árbol de decisión del modelo 3 de la variable dependiente FALLE.

En la Ilustración 32, el árbol ya tiene 5 hojas o nodos finales, proporcionando más detalles al árbol.

De los 147 países, 110 están por debajo del punto de corte de la variable migración neta (80.000) de los cuales se dividen 91 países con menos del 80% de población urbana y el 19 restante con más del 80%, que sufren de media 2,03 muertes y 7,42 muertes por cada 100.000 habitantes, respectivamente. Estos 19 países fueron:

- | | | | |
|-------------------|-----------------|----------------|------------|
| – México | – Nueva Zelanda | – Jordania | – Islandia |
| – Rep. Dominicana | – Dinamarca | – Luxemburgo | – Malta |
| – El Salvador | – Venezuela | – Países Bajos | – Uruguay |
| – Bahamas | – Líbano | – Argentina | – Singapur |
| – Finlandia | – Gabón | – Israel | |

Por el otro lado, de los 37 países con más de 80.000 de migración neta, estos se subdividen según densidad de población, los cuales por encima de 238 personas por kilómetro cuadrado se encuentran 7 países y son los más afectados con 46 muertes de media (Catar, Reino Unido,

Japón, Bélgica, Países Bajos, San Marino y Mónaco están en este grupo). Aquí se comprueba que tanto el movimiento de personas como la concentración de las mismas es un factor determinante.

Por debajo de 238 personas por kilómetro se encuentran los 30 países restantes que se vuelven a subdividir por tasa de mortalidad, al parecer, está relacionado la tasa de mortalidad de un país con las muertes por Covid-19. El punto de corte se encuentra en 7,8% de tasa de mortalidad, situándose 18 países por debajo y 12 por encima, con una media de muertes de 9,8 y 24,8 respectivamente. Teniendo en cuenta que la mayor tasa de mortalidad es de 15%, se podría decir que el punto de corte se encuentra casi en la mitad. Si nos fijamos en la base de datos, hay 83 países por debajo de la tasa de corte y si sumamos todas las muertes de estos países encontramos una cifra de 302 muertes por 100.000 habitantes (una media de 3,6 muertes por país), pero si sumamos todas las muertes de los 64 países que están por encima de la tasa encontramos una cifra de 819,23 (una media de 12,8 muertes por país) muertes por cada 100.000, casi 3 veces por encima. Aquí se muestran los 18 países que tienen menos de ese 7,8% de mortalidad:

- | | | | |
|--------------------------|------------|-------------|-------------|
| – Emiratos Árabes Unidos | – Ecuador | – Irlanda | – Canadá |
| – Omán | – Turquía | – Brasil | – Tailandia |
| – Kuwait | – Perú | – Nepal | – Noruega |
| – Arabia Saudí | – Colombia | – Australia | |
| – Malasia | – Chile | – Etiopía | |

Estos son países de América y algunos de Asia, pero poco se encuentra de Europa. Sin embargo existe una diferencia notable respecto a los 12 países que sí que tienen más porcentaje de mortalidad, y que la mayoría son europeos:

- | | | | |
|-----------|-------------|--------------|---------------------|
| – Suiza | – Francia | – Austria | – Alemania |
| – EE. UU. | – Suecia | – Rep. Checa | – Rusia |
| – España | – Sudáfrica | – Italia | – Guinea Ecuatorial |

No se sabe de forma fehaciente el porqué de esto, pero puede estar relacionada la tasa de mortalidad con otros factores que también expliquen el porqué. En definitiva, Se ha llegado a encontrar 4 variables que mejor explican por qué ha habido más número de muertes en unos países que en otros y la relación entre ellas.

Respecto a la variable migración neta, se ha observado de forma actual (en este mes de julio) el repunte de contagio y brotes en España a causa de la llegada de temporeros y turistas. Sin ir más lejos, una noticia reciente (el viernes 17 de julio) reveló que se produjo un brote en un



asentamiento ilegal de un campamento donde residían 500 temporeros. 23 personas dieron positivo en los análisis PCR de las 408 realizadas en ese momento [7].

3.4.2. Modelos de la variable dependiente: número de contagiados por Covid-19

A continuación, la variable dependiente objeto de estudio para el modelo será la variable número de contagiados por Covid-19 por cada 100.000 personas. Al igual que el modelo anterior, se utilizarán todas las variables para que el programa vaya aceptando y rechazando las que vea convenientes. En teoría, se debería obtener un modelo similar a los anteriores, ya que el número de muertes y número de infectados son variables similares y relacionadas entre sí. La regla obtenida es la siguiente:

- 1) *root* 147 18900490.00 190.82870
- 2) *URB*< 94.973 140 6348754.00 146.82690
- 4) *MIG*< 41409.5 97 1061739.00 66.40536
- 8) *MED*< 2.81915 77 203175.00 40.34364 *
- 9) *MED*>=2.81915 20 604911.70 166.74300
- 18) *GASTO*< 8.5 12 33487.27 80.33667 *
- 19) *GASTO*>=8.5 8 347442.80 296.35250 *
- 5) *MIG*>=41409.5 43 3244451.00 328.24280
- 10) *AGLO*< 24.37221 20 1349462.00 254.08400
- 20) *MED*< 3.2841 7 42239.36 54.94429 *
- 21) *MED*>=3.2841 13 880151.30 361.31310 *
- 11) *AGLO*>=24.37221 23 1689355.00 392.72870 *
- 3) *URB*>=94.973 7 6859416.00 1070.86600 *

Como se puede observar, han aparecido muchas más variables que en los modelos anteriores, apareciendo de nuevo las variables médicos y gasto en salud, las cuales siguen mostrando un doble rasero, debido a que cuando aumentan estas variables la variable dependiente es mayor por lo que debe estar relacionado de nuevo con el nivel económico. El árbol es el siguiente:

¿Qué afecta al número de contagiados por Covid-19?

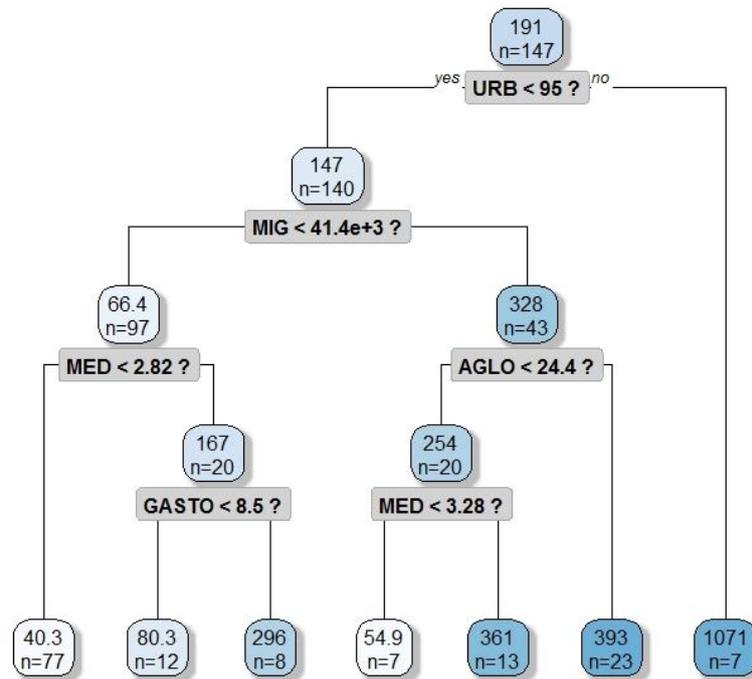


Ilustración 33. Árbol de decisión del modelo 1 de la variable dependiente INFECT.

En la Ilustración 33, se observa como los países con más del 95% de la población urbana son los más afectados. Estos 7 países tienen de media 1.071 contagiados por cada 100.000 personas, siendo dos de ellos los más perjudicados de todos los países, San Marino y Catar. Pero observando los datos de menor a mayor vemos un claro aumento de infectados a medida que aumenta la cantidad de población urbana de los países. Por el otro lado, los 140 países restantes con menos de 95% de población urbana se dividen por migración neta, al igual que en los modelos de fallecimientos, pero esta vez el punto de corte es mucho menor.

Los 97 países que se encuentran por debajo de 41.000 de migración neta se subdividen de nuevo por número de médicos y, posteriormente, en porcentaje de gasto en salud. Los 77 países que se encuentran por debajo de 2,82 médicos por cada 1.000 habitantes acaban teniendo de media 40,3 contagiados, mientras que los que tienen más de 2,82 médicos y además tienen menos del 8,5% del PIB de gasto en salud tienen de media 80,3 contagiados, y los que están por encima de ese 8,5% aumenta la cifra hasta 296 contagiados.

Sin embargo, los 43 países que se encuentran por encima de 41.000 de migración neta se subdividen en porcentaje de población en aglomerados urbanos y posteriormente en número de médicos. Los 23 países que se encuentran por encima de 24,4% de población en aglomerados se han visto afectados con 393 personas contagiadas de media por cada 100.000 habitantes. Mientras que de los 20 países con menos del 24,4% de población en aglomerados, 7 países tienen menos de 3,28 médicos por cada 1.000 habitantes y 54,9 contagiados, y los 13 restantes tienen más de 3,28 médicos y en cambio, tienen 361 contagiados de media.

Por lo tanto podemos afirmar que las variables número de médicos y gasto en salud vuelven a ser contradictorios con los resultados, y que están relacionados con otros factores.

Para eliminar este sesgo se procederá de la misma manera que en los modelos de fallecimientos, primero se planteará un modelo con todas las variables, eliminando únicamente gran parte de los países y dejando los más avanzados (los mismos 45 países) que son los que nos proporcionan las variables de PIB y paro. Y posteriormente se plantea otro con todos los países eliminado únicamente las variables médicos y gasto en salud.

El modelo anterior, eliminando los países menos relevantes, sigue la siguiente regla:

- 1) *root* 45 2556532.0 237.8398
- 2) *MIG* < 39351.5 19 296635.9 103.3105 *
- 3) *MIG* >= 39351.5 26 1664747.0 336.1496
- 6) *PARO* < 0.3 15 371669.9 245.7520 *
- 7) *PARO* >= 0.3 11 1003352.0 459.4191 *

Como cabía esperar, el modelo obtenido es sumamente simple respecto al anterior, pero efectivamente ha conseguido eliminar las variables que, como tal, no tenían sentido en el modelo. La variable migración neta vuelve a tener superioridad, apareciendo en todos los modelos propuestos, con un punto de corte similar al anterior. La variable paro aparece como una variable relevante para el modelo, con un punto de corte de 0,3 de variación trimestral del primer trimestre del 2020 respecto al anterior. En la Ilustración 36 se muestra el árbol obtenido de este modelo:

¿Qué afecta al número de contagiados por Covid-19?

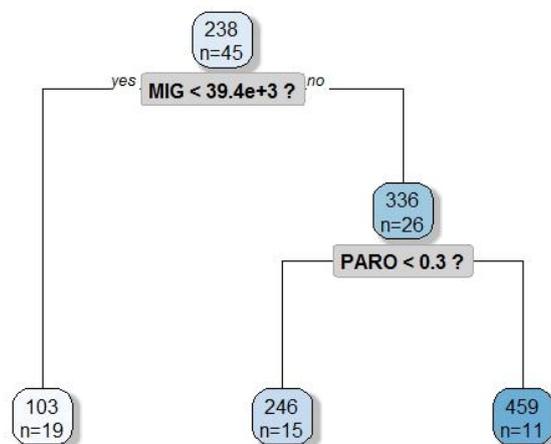


Ilustración 34. Árbol de decisión del modelo 2 de la variable dependiente INFECT.

En este árbol de 45 países de muestra existen dos nodos. El primero separa la muestra dejando 19 países que tienen una migración neta inferior a 39.400 con 103 infectados de media por cada 100.000 habitantes, mientras que los 26 restantes se subdividen según el nivel de variación del paro. Los 15 países con menos de 0,3% de variación de paro han tenido 246 infectados de media, mientras que los países que se encuentran por encima de 0,3% han tenido 459 infectados de media.

Esto no quiere decir que la variación del paro influya en los infectados, sino al contrario, los países que han tenido 246 infectados de media han sufrido a su vez una variación del paro de 0,3% o menos, llegando a cifras negativas, mientras que los que han sufrido una cifra mucho superior de infectados de media, también han sufrido una variación del paro de 0,3 o más.

Veamos entonces por último si, ampliando la muestra y eliminando las variables, número de médicos y gasto en salud, el modelo obtenido es más completo y detallado. La regla que nos muestra para el siguiente modelo es:

- 1) root 147 18900490.0 190.82870
- 2) URB< 94.973 140 6348754.0 146.82690
- 4) MIG< 41409.5 96 1059901.0 66.84969
- 8) AGLO< 29.64051 83 366716.6 46.92554 *
- 9) AGLO>=29.64051 13 449871.0 194.05770 *
- 5) MIG>=41409.5 44 3335060.0 321.32250
- 10) AGLO< 24.37221 22 859127.2 230.88640 *
- 11) AGLO>=24.37221 22 2116070.0 411.75860
- 22) CAMAS>=2.6 12 530987.6 292.43250 *
- 23) CAMAS< 2.6 10 1209180.0 554.95000 *
- 3) URB>=94.973 7 6859416.0 1070.86600 *

Efectivamente se puede ver en los resultados que han aparecido 5 variables explicativas para el modelo, de las cuales una es repetida para distintas ramas. Además, ha aparecido una nueva variable que no había aparecido en los anteriores modelos, el número de camas por cada 1.000 habitantes. Respecto a los puntos de corte, son muy similares al primer modelo de contagiados, con el punto de corte de número de camas en 2,6 camas por cada 1.000 habitantes. El árbol representado es el siguiente:

¿Qué afecta al número de contagiados por Covid-19?

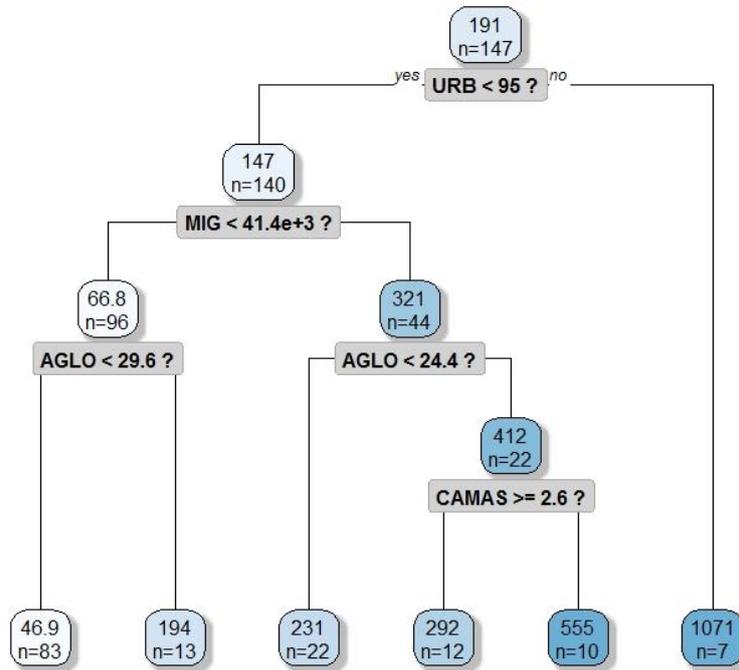


Ilustración 35. Árbol de decisión del modelo 3 de la variable dependiente INFECT.

Como se puede observar en la Ilustración 35, la primera rama del árbol que indica los países con más del 95% de población urbana es la misma que en el primer modelo, siendo esta variable la que mejor explica el número de contagiados y la que más perjudica a los países cuando tienen altos porcentajes de población urbana. Sin ir más lejos, En España los focos de infectados más importantes se produjeron en Madrid y Barcelona. Estos 7 países con más del 95% de población urbana son:

- Uruguay
- San Marino
- Bélgica
- Catar
- Singapur
- Kuwait
- Mónaco

No es el caso de Uruguay, pero los demás países tienen una tasa muy elevada de infectados, de media, contando a los 7, tienen 1.071 infectados por cada 100.000.

Los 140 países con menos del 95% de población urbana se dividen según la migración neta pero aquí las ramas se subdividen por la misma variable. Los 96 países que tienen menos de 41.400 de migración neta se dividen en porcentaje de población en aglomerados, de los cuales 83 países se sitúan por debajo de 29,6% de población en aglomerados y 13 por encima, con 46,9 y 194 contagiados respectivamente.

Los 44 países con más de 41.400 de migración neta se vuelven a dividir en porcentaje de población en aglomerados. Los 22 países que se encuentran por debajo de 24,4% de población en aglomerados tienen 231 contagiados de media. Estos países son:

- Etiopía
- Nepal
- Alemania
- Ucrania
- Rep. Checa
- Países Bajos
- Suecia
- Suiza
- Italia
- Noruega
- Tailandia
- Jordania
- Austria
- Francia
- Dinamarca
- Finlandia
- Rusia
- Irlanda
- Irak
- España
- Reino Unido
- Malasia

Mientras que los 22 restantes que superan los 24,4% se vuelven a desagregar según el número de camas. Los 12 países que tienen 2,6 camas, o más, por cada 1.000 habitantes tienen 292 contagiados. Estos 12 países son:

- Turquía
- Canadá
- Arabia Saudí
- Nueva Zelanda
- Sudáfrica
- EE. UU.
- Israel
- Australia
- Maldivas
- Luxemburgo
- Bielorrusia
- Japón

Sin embargo, hay una diferencia notable cuando es superior a 2,6 camas, puesto que los 10 países restantes con menos de 2,6 camas pasan a tener 555 contagiados por cada 100.000 habitantes:

- Emiratos Árabes Unidos
- Ecuador
- Colombia
- Omán
- Perú
- Guinea Ecuatorial
- Chile
- Brasil
- Panamá
- Andorra

Esto último se puede deber al menor control sanitario que poseen los pacientes cuando no están en camas hospitalarias. O que el número de camas refleja la capacidad sanitaria para hacer frente a la contención del virus para evitar su propagación.

3.4.3. Modelos según nivel económico

Se ha comprobado que el impacto del Covid-19 se refleja más en las variables con trasfondo económico, y que al eliminar dichas variables, los resultados obtenidos han sido más coherentes. Debido a ello, se procederá a dividir los países en tres grupos según su nivel económico calculado para el IDH: Países de ingresos altos, medios y bajos.

La lista de países según la clasificación anteriormente indicada se encuentra en la tabla 6, en el apartado de análisis bivariante.

El grupo de países con ingresos altos se compone de 50 países. Los árboles obtenidos tanto de contagiados, como de fallecidos son los siguientes:

¿Qué afecta al número de contagiados por Covid-19?

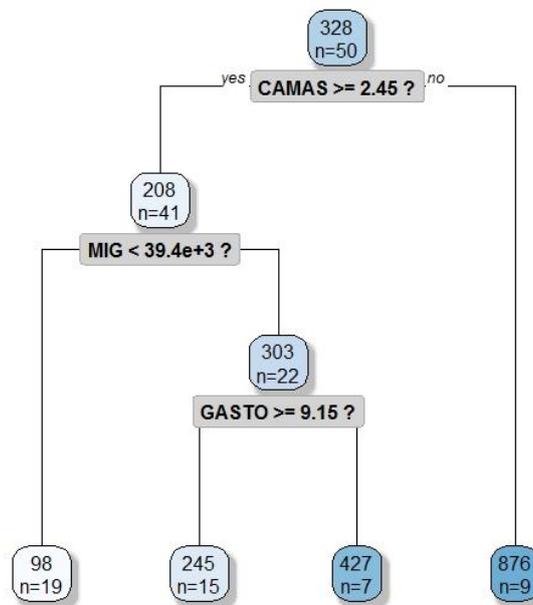


Ilustración 36. Árbol de decisión del modelo de la variable dependiente INFECT en los países con ingresos altos.

¿Qué afecta al número de muertes por Covid-19?

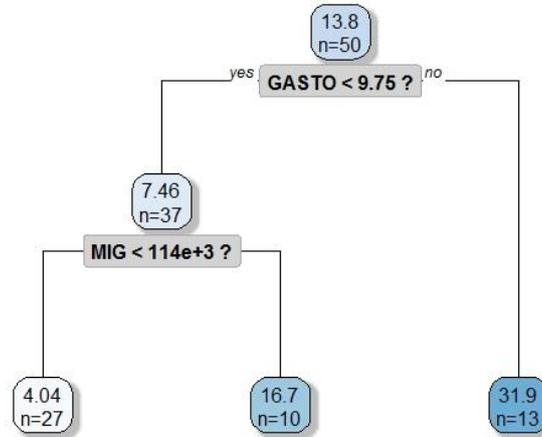


Ilustración 37. Árbol de decisión del modelo de la variable dependiente FALLE en los países con ingresos altos.

Como se puede observar en la Ilustración 36, el número de camas es el factor que mejor explica el modelo de contagiados, siendo 9 países los más afectados con 976 contagios por cada 100.000, de media, al tener un número de camas menor a 2,45 por cada 1.000 habitantes. Los 41 restantes se subdividen según migración neta: por un lado, los 19 países menos afectados de este grupo son los que tienen menos de 39.000 de migración neta, concordando con resultados anteriores, y coincidiendo con el árbol de la Ilustración 39 del modelo de fallecidos. Por último, los 22 países restantes con mayor de 39.000 de migración neta se subdividen según el gasto en salud: los 15 países con 9% o más de gasto en salud sobre el PIB tienen 245 contagiados de media; mientras que los que tienen menos de 9% tienen de media 427 contagiados por cada 100.000 habitantes.

Por otro lado, en la ilustración 37, la situación es parecida pero distinta: la variable que mejor explica el modelo de fallecidos en los países con ingresos altos es la variable de gasto en salud, pero al contrario que en el anterior, cuanto más gasto sobre el PIB tienen mayor es el perjuicio en el número de fallecidos. En cambio, la siguiente rama coincide con la variable migración neta con el anterior modelo, la única diferencia es el punto de corte el cual es superior en este modelo.

Esto puede tener la explicación en la forma de contabilizar, ya que las muertes es más fácil de confirmarlo con un mayor sistema sanitario pero con los contagiados no pasa lo mismo.

Lo que sí se puede confirmar es que la migración neta es un factor que afecta a los países con ingresos altos, tanto en fallecidos como en contagiados, y dichos países son también los que mayor migración neta tienen. Y por otro lado, se puede confirmar que el número de camas sí que afecta al número de contagiados, más que el gasto en salud o en médicos.

A continuación se muestran los modelos de fallecidos y contagiados de los 44 países con ingresos medios:

¿Qué afecta al número de contagiados por Covid-19?

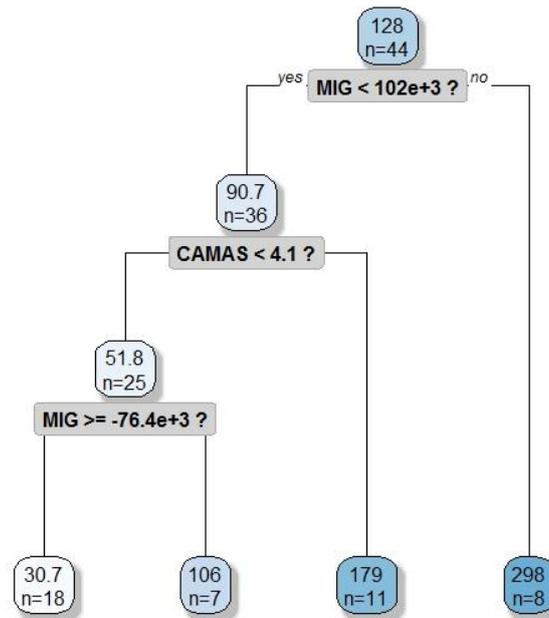


Ilustración 38. Árbol de decisión del modelo de la variable dependiente INFECT en los países con ingresos medios.

¿Qué afecta al número de muertes por Covid-19?

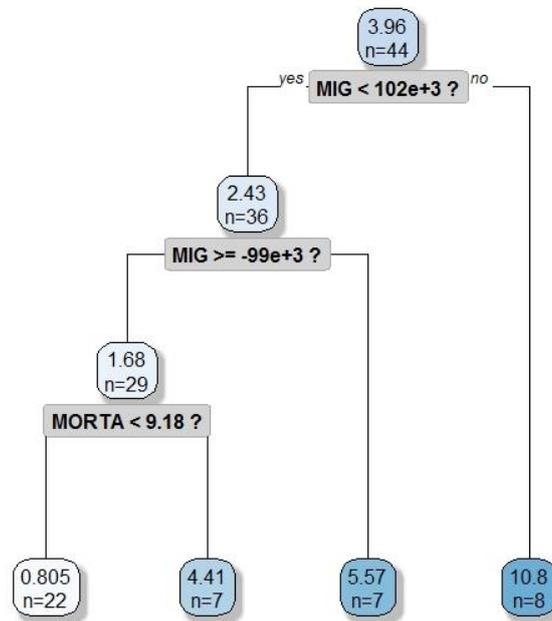


Ilustración 39. Árbol de decisión del modelo de la variable dependiente FALLE en los países con ingresos medios.

En los modelos de las Ilustraciones 38 y 39 se aprecian contradicciones en ambos modelos.

La variable que más afecta en ambos modelos es la migración neta, en la que cuanto mayor es más afecta, con el mismo punto de corte en ambos modelos, 102.000 de migración neta. Aquí coincide con todos los modelos anteriores, se ve la contradicción en las siguientes variables, donde: en el modelo de contagiados de la Ilustración 38, la siguiente rama se divide según el número de camas pero aquí en punto de corte indica que los países que tienen más de 4 camas son los más afectados respecto a los que tienen menos, contradiciéndose con el anterior modelo de los países con ingresos altos; la siguiente rama que continua en el árbol de contagiados vuelve a encontrarse la variable migración neta, pero esta vez con el signo contrario al anterior.

Lo mismo ocurre en el árbol de la Ilustración 39, la segunda rama es la variable migración neta, coincidiendo con la anterior rama pero con el signo distinto, siendo una contradicción. En cambio en la siguiente rama encontramos la variable mortalidad, donde los países que tienen más de 9% de mortalidad están más afectados que los que tienen menos.

Estas contradicciones se pueden deber a la influencia de otros factores en países específicos de ingresos medios. Lo que sí podemos confirmar es que la migración neta sigue siendo el factor más relevante para los modelos ya que siempre aparece, el matiz de verlo contradictorio puede deberse a que la migración neta es la diferencia de los que entran y salen del país, por lo tanto no sabemos exactamente cuanta gente a salido o entrado, solo la diferencia, por lo

que cuanto mayor es la diferencia, y menos se acerca a 0, quiere decir que ha podido haber un movimiento mayor de personas que en países donde no ha habido prácticamente movimiento, es decir, cercano a 0.

Por último, se presentan los árboles obtenidos de los países con ingresos medios, con una muestra de 43 países:

¿Qué afecta al número de contagiados por Covid-19?

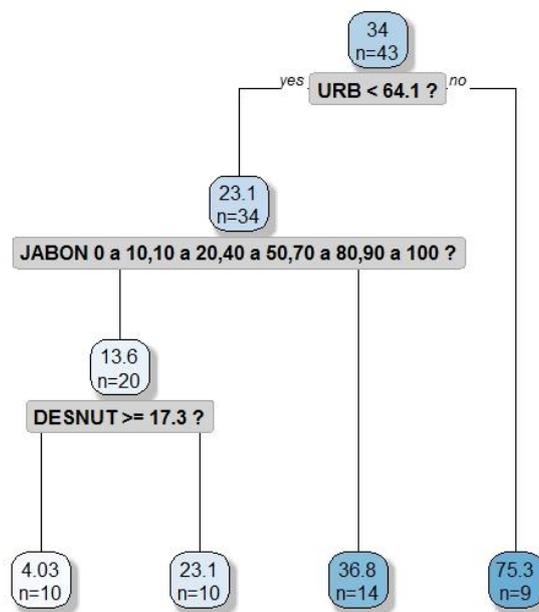


Ilustración 40. Árbol de decisión del modelo de la variable dependiente INFECT en los países con ingresos bajos.

¿Qué afecta al número de muertes por Covid-19?

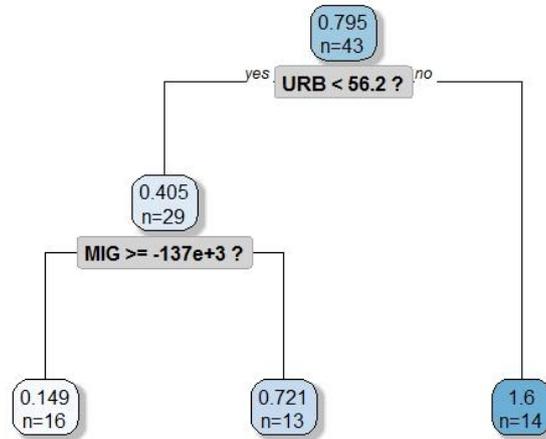


Ilustración 41. Árbol de decisión del modelo de la variable dependiente FALLE en los países con ingresos bajos.

En los modelos de los países pobres se han obtenido resultados totalmente diferentes a los anteriores. En ambos modelos, de infectados y fallecidos, el factor que más influye es el porcentaje de población urbana, con el punto de corte de la variable en 64% y 56% respectivamente, una cifra mucho más baja que en los primeros modelos.

En el árbol de la Ilustración 40, la rama de la izquierda se subdivide en una variable cualitativa, la primera que aparece en los modelos. Ésta indica que de los 34 países, 20 son los que se encuentran en los grupos mostrados, que se dirigen por la rama izquierda, mientras que los 14 restantes que se encuentran en los grupos que no se muestran se dirigen por la derecha. Esto no presenta una coherencia como tal, debido a que los grupos de los extremos se dirigen por la misma rama. Y finalmente, la siguiente variable que subdivide el árbol es la prevalencia de desnutrición, pero ésta tampoco presenta coherencia, pudiendo deberse a un sesgo por la rama anterior.

Por otro lado, como se puede observar en la Ilustración 41, el árbol obtenido de este modelo es muy pequeño. La siguiente variable que sigue a la anteriormente mencionada es la migración neta, pero esta vez con el signo al contrario de lo habitual.

Por lo tanto, podemos concluir que en los países con ingresos bajos lo que más afecta es el porcentaje de población urbana, mientras que en países con ingresos medios y altos lo que más afecta es la migración y factores análogos a la salud.

Lo que si se aprecia comparando los modelos entre los grupos es la enorme diferencia del nivel de contagiados y fallecidos que poseen de media. Con unas muestras similares entre los grupos se observa que: los 50 países con ingresos altos tienen de media 328 contagiados y 13,8



fallecidos por cada 100.000 habitantes; los 44 países con ingresos medios tienen 128 contagiados y 3,96 fallecidos por cada 100.000 habitantes; y los 43 países con ingresos bajos tienen 34 contagiados y 0,79 fallecidos por cada 100.000 habitantes.

4. Conclusiones y futuras líneas de investigación

En esta última sección se exponen las conclusiones más relevantes obtenidas durante la investigación. Además de una breve descripción de utilidad para continuar con la investigación.

4.1. Conclusiones

Como conclusiones, el desarrollo de todo el estudio ha dado lugar a la obtención de un modelo estadístico consistente en conocer los factores que mejor explica el grado de impacto que sufren los países ante una pandemia viral.

Pandemias virales han existido durante toda la historia de la humanidad, pero en esta situación nunca se había encontrado el mundo. Con tanta globalización, y un virus tan silencioso al principio, como con una mortalidad considerable al final, ha provocado esta desestabilidad global, tanto social como económica.

Se han escogido el número de fallecimientos por Covid-19 y el número de contagiados como variables dependientes a explicar. De las 14 variables escogidas para conseguir el modelo mejor representativo, tan solo 6 variables son las que mejor han encajado: migración neta, población urbana, densidad de población, porcentaje de población en aglomerados, tasa de mortalidad y número de camas.

En el modelo que explica el número de muertes la migración neta es la más influyente, apareciendo también en el modelo de infectados. De hecho, los países como EE.UU, Alemania, Turquía, Reino Unido, Canadá, Colombia, Rusia, Australia, Italia y Sudáfrica, tienen la mayor migración neta, pero son también algunos de los países más afectados. Por ello, este es uno de los factores más importantes. Es obvio que cuanto más movimiento de gente entre países mayor riesgo de contagios.

En cambio, de las variables que explican el número de infectados, el porcentaje de población urbana es el más influyente, es decir, cuanto más gente vive junta en ciudades más contagios. También aparece en el modelo de fallecimientos, y países como San Marino y Catar, además de ser de los más afectados tienen también más del 90% de población urbana.

La variable densidad de población y porcentaje de población en aglomerados urbanos van de la mano. San Marino vuelve a ser un país con alta cantidad de estas variables, al igual que Suecia. Aunque no se cumple tanto como las anteriores variables, sí que es representativo para el modelo.

Estas variables ya mencionadas explican el porqué de tantos contagios, y que deriven en fallecimientos, pero no nos muestra una clara variable de por qué mueren más personas en unos países que en otros con un número similar de contagiados. La variable, tasa de mortalidad, muestra un poco de claridad sobre esto, puesto que de los 20 países con más tasa de mortalidad son casualmente de los que más fallecimientos tienen, aunque también existen excepciones de países en el que no se cumple esto.

En cuanto la variable número de camas, es una de las variables que más correlación tiene con la variable número de infectados, sin ser tan significativo, pero forma parte del modelo con un punto de corte que divide casi la mitad de los países. A medida que los países tienen menos camas de hospital, tienen a su vez mayor número de contagiados.

Dos de las variables que aparecieron en el modelo tenían una relación inversa a la que se esperaba. Según los modelos obtenidos en los primeros cálculos, las variables número de

médicos y gasto en salud tienen un efecto negativo en nuestros modelos, a medida que aumentan estas variables causan mayor número de muertes y contagios. Esto también concuerda en el que los países que tienen mayor inversión en salud y mayor número de médicos son mayoritariamente países del norte, del primer mundo, quienes también han sido los más afectados por la pandemia.

En el análisis bivariante se puede observar lo mismo, donde los países con prácticamente el total de la población con instalaciones básicas de agua e higiene son precisamente los países más desarrollados y a su vez más afectados por la pandemia. Por lo que ni la sanidad ni la higiene son capaces de frenar la pandemia, si no que los aspectos sociales son los más perjudiciales para los países.

Eso sí, las variables que tenían que ver con el tipo de confinamiento y el número de días hasta que se aplican no han resultado ser representativas para el modelo, ni han tenido una relación significativa con las variables de interés. Pudiendo deberse a la circunstancia de que el Covid-19 es un virus que no presenta casi síntomas o los genera lentamente, que provoca una reacción tardía de los países ante el impacto que otorga.

Existe una clara evidencia de diferencia entre los países del norte y del sur, diferencias que todos conocen y que no son casualidad tampoco en la propagación del virus. Virus que no ha sido tan mortal como otros que sí que han sufrido países del sur, pero con un nivel de contagio que ha traído por descuido a los países desarrollados, que son los que más riesgo tienen de hacer circular el virus. Los países del sur al estar más aislados del mundo, y al no estar tan concentrados en ciudades, han podido superar de distinta manera la pandemia global.

También se ha visto durante los análisis que, el paro principalmente, y el PIB, se han visto influenciado por virus. Los países, de los que se han podido obtener estos datos, han sufrido el impacto de la pandemia, de media, la variación del PIB y el paro ha sido de -2% y 0,3% respectivamente. Pero no todos los países han tenido perjuicios durante la pandemia global en este sentido. Excepciones como Chile y Suecia, que son países muy afectados por la pandemia, han experimentado variaciones positivas del PIB, y a su vez, Italia, Francia y Bélgica han experimentado variaciones negativas del paro. Por lo tanto, además del impacto generalizado por la pandemia deben influir también las políticas aplicadas en cada país en la variación de estas variables. Esto solo se ha podido comprobar en países con ingresos altos y medios debido a la ausencia de los datos.

Al separar los países en tres grupos según su nivel económico se ha podido comprobar que los países más afectados, que han sido los países con ingresos altos, la relación con las variables se enfocan más en lo sanitario, social, y económico, debido a que las variables de población urbana y aglomeración de población son valores muy aproximados entre estos países, y los resultados se reflejan más en otras variables. Mientras que los países con ingresos medios y bajos la relación con las variables se dirige más al movimiento de personas y social, debido a que tienen valores más dispersos, en cambio en el aspecto sanitario tienen valores más similares. Por lo tanto sabemos que faltan variables que expliquen mejor las diferencias de los países en cuanto al impacto del Covid-19, pero se ha podido hacer una criba de todos los factores analizados y se han encontrado variables que influyen significativamente en los países.



Por último, los modelos obtenidos de esta separación refleja que el impacto de los países, con ingresos medios y altos, se debe más a la migración de personas y sistemas sanitarios. Mientras que el impacto de los países con ingresos bajos se debe más a la población urbana. Habiendo una clara diferencia entre los países de ingresos altos, medios y bajos.

4.2. Futuras líneas de investigación

Este estudio estadístico, junto al modelo, sirve como base para profundizar más sobre futuros modelos que ayuden a las futuras actuaciones de los países. Estos resultados estadísticos son de utilidad para tener en cuenta ciertas variables que están influenciados por otros factores.

Una vez finalizado el virus, se podrá volver a realizar un análisis más extenso y exacto, con premisas para formalizar un nuevo modelo y evitar pasar por los mismos pasos y conclusiones.

5. Bibliografía

Bibliografía introducción

Web

[1] OMS. URL: <https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019/advice-for-public/q-a-coronaviruses>

Bibliografía apartado “antecedentes”

Artículos

[2] RTVE (2020). <<China ¿Qué se sabe del coronavirus de Wuhan?>>. En: *rtve noticias*.

Pillado, A (2014). <<Conoce los 10 virus más letales del mundo>>. En: *sdp noticias*.

Maldita ciencia (2020) <<Por qué el coronavirus no es como una gripe: COVID-19 tiene más mortalidad y transmisibilidad y no tiene tratamiento ni vacuna pero todavía se puede contener, según la OMS>>. En: *Maldita ciencia*.

L.A. (2020). << El coronavirus, comparado con las pandemias más letales de la historia>>. En: *ABC*.

Ángeles. L (2004). <<Historia de la gripe aviar>>. En: *El Mundo*.

Web

[3] Matemáticas y sus fronteras, marzo 2020. URL: <http://www.madrimasd.org/blogs/matematicas/2020/03/28/147534>

Information is beautiful. URL: <https://www.informationisbeautiful.net/visualizations/the-microbescope-infectious-diseases-in-context/>

Information is beautiful. URL: <https://infobeautiful4.s3.amazonaws.com/2014/10/IIB-microbescope-1276x2.png>

Statista. URL: <https://es.statista.com/estadisticas/1095743/infecciones-y-muertes-de-los-mayores-brotos-virales-de-los-ultimos-50-anos/>

El Correo. URL: https://www.elcorreo.com/apoyos/especiales/gripe_aviar/faqs.html

Wikipedia. URL: <https://es.wikipedia.org/wiki/Marburgvirus>

OPS. URL: https://www.paho.org/hq/index.php?option=com_content&view=article&id=14911:hantaviruses&Itemid=40721&lang=es

Bibliografía definiciones

Web

[4] Máxima Formación. URL: <https://www.maximaformacion.es/blog-dat/que-es-r-software/>

[5] Estadística para todos. URL: <https://www.estadisticaparatodos.es/taller/graficas/cajas.html>

Libros

[6] Aamonde, A y Luaces, R. (mayo 2020). << UNIDAD 4. Arboles de Decisión: Cart; Clasificación C5.0>>. En: DATA MINING CON R.

Therneau, T y Atkinson, B. (abril 2019). << Recursive Partitioning and Regression Trees (Package 'rpart')>>

Milborrow, S. (agosto 2019). << Plot 'rpart' Models: An Enhanced Version of 'plot.rpart'>>.

Santana, A. << Introducción al uso de R-commander>>

Bibliografía base de datos

Artículos

Velasco, M (2020) << Islandia: un caso único en el mundo que permitirá desvelar las incógnitas del coronavirus>> En: Huffpost.

Christian, J, (2020) << El Supremo de Malauí vuelve a bloquear la aplicación del confinamiento en el país a causa del coronavirus>>. En: Europa Press.

Infobae (2020) << Malí, empobrecido y en guerra, teme lo peor cuando llegue el coronavirus>> En: Infobae.

Von, B (2020). <<Kenia anuncia el confinamiento durante 15 días en dos zonas de Nairobi y Mombasa a causa del coronavirus>>. En: Europa Press.

Parrock, J (2020). <<Malta inaugura la temporada de caza a pesar del confinamiento>>. En: Euronews.

Martínez, J (2020). <<El dilema del confinamiento en África: salvar vidas o aumentar la pobreza>>. En: Lamarea.

Xinhuanet (2020) <<Kirguistán extiende confinamiento en algunas ciudades>>. En: Xinhuanet.

Milfort, M (2020) <<El hambre acecha a los nuevos desempleados en Haití>>. En: EFE.

Maceda, G (2020). << Confinados en Irán por la pandemia del coronavirus: "Aquí la gente no se queda en casa sin nada que llevarse a la boca">>. En 20Minutos.

Lahoz, R (2020). <<Antonio Sarría: "Solo hay 25 muertos por la covid en Kazajistán">>. En: Heraldó.

Ondacero (2020). <<¿Cómo afecta el confinamiento en Islandia a un entrenador de baloncesto?>>. En: Ondacero.

La Vanguardia (2020). <<Macedonia del Norte declara el estado de emergencia en dos ciudades>>. En: La Vanguardia.

Europa Press (2020). <<Surinam anuncia su primera muerte a consecuencia del coronavirus>>. En: Europa Press.

Gasparini, J (2020). <<Coronavirus: Suiza empieza a limitar el confinamiento el lunes>>. En: Perfil.

Castrillo, B (2020) <<El Gobierno de Suazilandia desmiente que el rey Mswati III esté grave por coronavirus>> En: Hola!

Fresno, D (2020). <<Suiza empieza a salir del confinamiento tras registrar más de 30.000 casos de covid-19>> En: Voz Populi.

La Vanguardia (2020). <<La policía de Sri Lanka alegra el confinamiento con música tradicional>> En: La Vanguardia.

Europa Press (2020). << Yibuti da marcha atrás en el levantamiento a las restricciones al considerarlas prematuras>>. En: Europa Press.

Europa Press (2020). << Zambia extiende durante dos semanas las medidas decretadas para frenar la propagación del coronavirus>>. En: Europa Press.

Infobae (2020). <<Los amantes del khat en Yemen no conocen el confinamiento>>. En: Infobae.

Swissinfo (2020). <<El confinamiento se incumple en Yibuti y la epidemia se agrava>>. En: Swissinfo.

Diaria (2020). <<Rafael Radi: "El confinamiento en forma sostenida es imposible">>. En: Diaria.

Fernández, N (2020). <<Los uruguayos rompen el confinamiento y Lacalle Pou exhorta a mantener los cuidados>>. En: La Nación.

Albertonews (2020). <<Trinidad y Tobago extiende el confinamiento hasta el 15 de mayo: se espera más retorno migratorio al Delta>>. En: Albertonews.

Marca (2020). << Los 19 países sin rastro del Coronavirus >>. En: Marca.

Europa Press (2020). << Macedonia del Norte extiende el toque de queda y anuncia franjas de movimiento por edad por el coronavirus>>. En: Europa Press.

Europa Press (2020). << Malta anuncia el cierre de bares, restaurantes, discotecas y cines a causa del coronavirus>>. En: Europa Press.

Mancilla, S (2020). << Cuarentena en México: Cuánto dura y hasta cuándo se puede alargar el confinamiento>>. En: AS México.

Europa Press (2020). << El ministro del Interior de Moldavia se contagia de coronavirus>>. En: Europa Press.

El País (2020). <<Alberto de Mónaco, positivo en coronavirus>>. En: El País.

Sport (2020). <<Japón inicia hoy su estado de alarma durante un mes completo>>. En: Sport.

Sánchez, D (2020). <<Japón no decreta el confinamiento total pese a que teme que los contagios salten de 4.000 a 80.000>>. En: Libre mercado.

Alamillos, A (2020). << ¿Qué está pasando en San Marino? El país más afectado por coronavirus del mundo>>. En: El Confidencial.

Solís, A (2020). <<Emergencia mundial: ¿En qué países se ha ordenado el confinamiento?>>. En: Economía Digital.

Europa Press (2020). <<El ministro del Interior de Moldavia se contagia de coronavirus>>. En: Europa Press.

Amorós, A (2020). << Burundi lanza la campaña electoral a pesar de la Covid-19 >>. En: ABC.

Swissinfo (2020). << África empieza a decretar toques de queda y confinamientos contra el coronavirus >>. En: Swissinfo.

Bolsa Manía (2020). << Coronavirus.- Belice, último país latinoamericano al que llega el coronavirus>>. En: Bolsa Manía.

El País (2020). << Qué es la cuarentena general obligatoria y qué implicó en otros países >>. En: El País.

Cinco Días (2020). <<Del confinamiento en casa al retraso del contagio: cómo afrontan los países europeos el Covid-19>>. En: Cinco Días.

Malavia, M (2020). <<La medida más sorprendente del presidente de Tanzania contra el coronavirus: tres días de oración nacional>>. En: vida nueva digital.

Europa Press (2020). <<Uzbekistán amplía la cuarentena por el coronavirus hasta el 10 de mayo>>. En: Europa Press.

Europa Press (2020). <<Somalia impondrá desde el miércoles un toque de queda en Mogadiscio por el coronavirus>>. En: Europa Press.

Fresno, D (2020). << El caso de Suecia descoloca a Europa: ¿cómo es posible que no hayan impuesto el confinamiento?>> En: Voz Populi.

Europa Press (2020). <<Sudán impondrá desde el sábado el confinamiento en la capital por la pandemia de coronavirus>>. En: Europa Press.

Statista (2020) (Caído) Orús, A (2020) <<El nuevo coronavirus (COVID-19) – Datos estadísticos>> En: Statista.

Web

Index Mundi. URL: <https://www.indexmundi.com/g/r.aspx?v=2227&l=es>

Wikipedia. URL:
https://es.wikipedia.org/wiki/Cuarentena_por_la_pandemia_de_enfermedad_por_coronavirus_de_2019-2020_-_Cuarentena_por_pa%C3%ADses

Index Mundi. URL: <https://www.indexmundi.com/g/r.aspx?t=0&v=2225&l=es>

Wikipedia. URL:
https://es.wikipedia.org/wiki/Anexo:Cronolog%C3%ADa_de_la_pandemia_de_enfermedad_por_coronavirus_de_2019-2020

Wikipedia. URL:
https://es.wikipedia.org/wiki/Pandemia_de_enfermedad_por_coronavirus_de_2020_en_Guyana

Eurostat. URL:
https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsq_urgan&lang=en

Datosmacro. URL: <https://datosmacro.expansion.com/pib>

Ep.data. URL: <https://www.epdata.es/datos/desempleo-pib-inflacion-union-europea-ue-eurozona/432>

Datosmacro. URL: <https://datosmacro.expansion.com/paro>

Datosmacro. URL: <https://datos.bancomundial.org/indicador/SL.UEM.TOTL.ZS>

OCDE.Stat. URL: <https://stats.oecd.org/index.aspx?queryid=66948>

RTVE. URL: <https://www.rtve.es/noticias/20200517/mapa-mundial-del-coronavirus/1998143.shtml>

RTVE. URL: <https://app.flourish.studio/visualisation/2888792/>

Bibliografía apartado “análisis multivariante”

Artículos

[7] ABC (2020). << El brote del asentamiento irregular de Albacete alcanza los 23 positivos>>. En: en ABC Castilla la Mancha

6. Anexo

En el Anexo se aporta una parte de la base de datos con las variables que, al final, han sido las más significativas, y la cual ha sido construida personalmente y trabajada para el análisis. Además, se incluyen los gráficos no aportados durante en el proyecto.

Base de Datos

Países	MIG	URB	AGLO	CAMAS	MORTA	FALLE	INFECT
Afganistán	-314602,00	25,50	10,792	0,50	6,58	1,47	73,54
Albania	-69998,00	60,32		2,90	7,71	1,33	60,08
Alemania	2719112,00	77,31	9,580	8,30	11,30	10,68	227,49
Andorra		88,06		2,50		67,52	1108,95
Antigua y Barbuda	0,00	24,60		3,80	6,32	3,12	27
Arabia Saudí	674895,00	83,84	47,280	2,70	3,45	3,24	419,09
Argelia	-50002,00	72,63	6,379	1,90	4,72	1,89	26,68
Argentina	24000,00	91,87	43,008	5,00	7,62	2,05	79,9
Armenia	-24989,00	63,15	36,599	4,20	9,86	10,23	610,92
Australia	791229,00	86,01	60,868	3,80	6,50	0,41	29,58
Austria	324998,00	58,30	21,482	7,60	9,50	7,77	194,59
Azerbaiyán	6002,00	55,68	22,990	4,70	5,80	1,34	110,58
Bahamas	4999,00	83,03		2,90	6,72	2,85	26,97
Bangladés	-1847503,00	36,63	15,118	0,80	5,53	0,81	61,04
Barbados	-397,00	31,15		5,80	8,94	2,44	33,84
Bélgica	240000,00	98,00	26,980	6,20	9,60	84,62	526,92
Belize	6000,00	45,72		1,30	4,68	0,52	5,74
Benín	-10000,00	47,31		0,50	9,03	0,08	4,98
Bielorrusia	43648,00	78,60		11,00	12,60	3,42	590,84
Birmania				0,90		0,01	0,49
Bolivia	-47520,00	69,43	41,336	1,10	6,80	5,98	182,2
Bosnia y Herzegovina	-107926,00	48,25		3,50	10,56	5,05	94,5
Botsuana	14999,00	69,45		1,80	5,83	0,04	3,5
Brasil	106000,00	86,57	41,946	2,20	6,38	22,2	456,09
Brunéi	0,00	77,63		2,70	4,35	0,7	32,87
Bulgaria	-24001,00	75,01	18,115	6,80	15,50	2,62	50,42
Burkina Faso	-125000,00	29,36	12,816	0,40	8,34	0,27	4,55
Burundi	10003,00	13,03		0,80	8,11	0,01	0,93
Cabo Verde	-6709,00	65,73		2,10	5,68	1,29	145,65
Camerún	-24000,00	56,37	28,029	1,30	9,50	1,09	39,12
Canadá	1210159,00	81,41	45,933	2,70	7,50	22,43	273,87
Catar	200000,00	99,14		1,20	1,17	2,95	2990,06
Chile	558539,00	87,56	35,668	2,20	6,08	19,3	1177,99
China	-1741996,00	59,15	27,888	4,20	7,11	0,33	5,98
Colombia	1023981,00	80,78	43,897	1,50	5,50	3,8	110,95
Corea del Sur				11,50		0,54	23,75
Costa Rica	21000,00	79,34	27,158	1,10	5,01	0,24	37,42
Croacia	-40004,00	56,95		5,60	13,00	2,62	55,24
Cuba	-72000,00	77,04	18,843	5,20	8,83	0,74	20,11
Dinamarca	75998,00	87,87	22,783	2,50	8,20	10,32	215,65



Ecuador	182000,00	63,82	27,635	1,50	5,10	23,45	283,83
EE. UU.	4774029,00	42,70	46,258	2,90	8,50	36,03	662,19
Egipto	-190164,00	72,02	25,565	1,60	5,86	1,88	50,01
El Salvador	-202694,00	82,26	17,236	1,30	7,01	1,28	63,33
Emiratos Àrabs Unidos	200000,00	86,52	59,972	1,20	1,43	3,06	450,26
Eritrea	-199290,00	40,70		0,70	7,33	0	4,08
Eslovaquia				5,80		0,51	28,66
Eslovenia	9999,00	54,54		4,60	9,90	5,26	72,47
Espanya	200000,00	80,32	25,664	3,00	9,00	57,99	522,87
Estonia	19555,00	68,88		5,00	11,80	5,22	149,55
Etiopía	150002,00	20,76	4,028	0,30	6,69	0,06	3,44
Fiji	-31008,00	56,25		2,30	8,18	0	2,04
Filipinas	-335758,00	46,91	14,278	1,00	5,84	1,04	25,54
Finlandia	70000,00	85,38	23,180	4,40	9,80	5,91	129,04
Francia	182636,00	80,44	22,757	6,50	9,00	44,16	236,16
Gabón	16301,00	89,37		6,30	7,00	1,42	199,55
Gambia	-15436,00	61,27		1,10	8,04	0,04	1,49
Georgia	-50000,00	58,63	28,875	2,60	12,92	0,38	23,96
Ghana	-50000,00	56,06	18,491	0,90	7,41	0,22	42,3
Grecia	-80000,00	79,06	29,416	4,30	11,60	1,74	29,85
Guatemala	-46073,00	51,05	16,530	0,60	4,76	2,5	65,23
Guinea	-20000,00	26,61	14,847	0,30	8,72	0,21	37,6
Guinea Ecuatorial	79998,00	36,14		2,10	9,54	2,44	127,12
Guinea-Bissau	-6996,00	72,14		1,00	9,81	0,8	79,6
Guyana	-30001,00	43,36		1,60	7,37	1,54	21,95
Haití	-175000,00	55,28	23,705	0,70	8,64	0,74	42,15
Honduras	-34000,00	57,10	14,217	0,70	4,43	3,5	107,42
Hungria	29999,00	71,35	18,011	7,00	13,50	5,8	41,72
India	-2663434,00	34,03	15,435	0,70	7,21	0,9	27,13
Indonesia	-494777,00	55,33	13,383	1,20	6,43	0,85	15,48
Irak	39171,00	70,47	25,077	0,20	4,83	2,01	63,11
Irán	-274998,00	74,90	25,989	1,40	4,84	11,23	238,45
Irlanda	118020,00	63,17	24,754	2,80	6,30	35,13	520,64
Islandia	1900,00	93,81		3,40	6,50	2,84	514,57
Israel	50002,00	92,42	57,924	3,10	5,10	3,41	223,96
Italia	744713,00	70,44	18,741	3,40	10,70	57,01	393,61
Jamaica	-56658,00	55,67		1,70	7,56	0,34	21,33
Japón	357800,00	91,62	64,628	13,40	10,80	0,75	14,45
Jordania	51099,00	90,98	20,737	1,40	3,84	0,09	9,91
Kazajistán	-90000,00	57,43	15,852	6,70	7,15	0,53	86,89
Kenia	-50000,00	27,03	10,896	1,40	5,58	0,21	7,87
Kirguistán				4,50		0,49	42,02
Kuwait	197600,00	100,00	72,252	2,00	2,60	7,4	907,18
Letonia	-74186,00	68,14		5,80	14,80	1,56	57,29



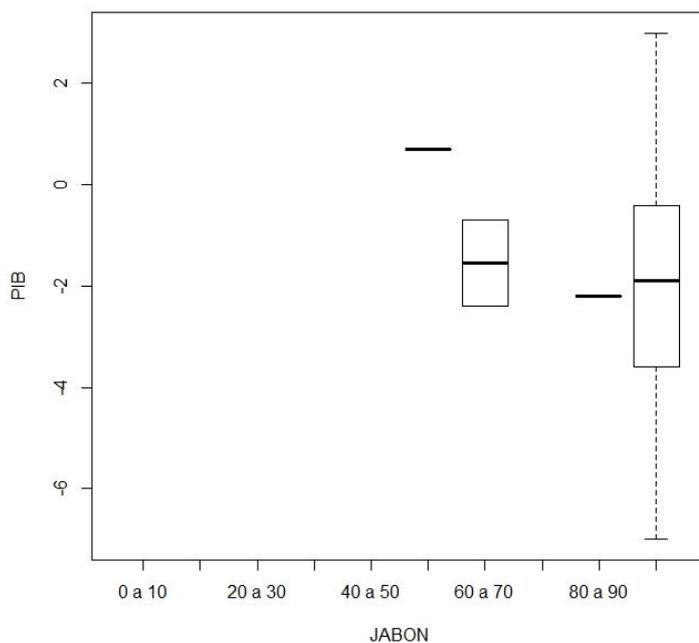
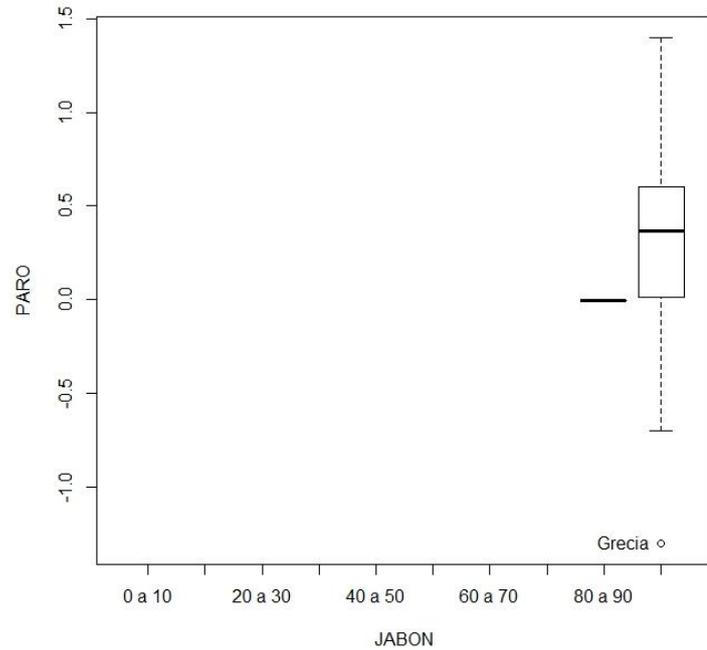
Líbano	-150060,00	88,59	34,827	2,90	4,26	0,47	21,74
Liberia	-25000,00	51,15	29,432	0,80	7,73	0,68	10,71
Libia	-9997,00	80,10	17,335	3,70	5,10	0,15	7,49
Lituania	-163902,00	67,68		7,30	14,20	2,71	63,47
Luxemburgo	48704,00	90,98		4,90	7,10	18,09	671,93
Macedonia del Norte	-4999,00	57,96		4,40	9,97	10,08	215,17
Madagascar	-7500,00	37,19	11,646	0,20	6,19	0,05	5,25
Malasia	249999,00	76,04	23,991	1,90	5,01	0,38	27,01
Malauí	-80263,00	16,94	5,675	1,30	6,81	0,03	3,15
Maldivas	56851,00	39,81		4,30	2,88	1,55	411,09
Mali	-200000,00	42,36	12,825	0,10	9,98	0,56	9,91
Malta	4501,00	94,61		4,70	7,60	1,86	136,6
Marruecos	-257096,00	62,45	21,732	1,10	5,07	0,59	24,97
México	-300000,00	80,16	40,709	1,50	5,93	15,12	126,63
Moldavia	-6935,00	42,63		5,80	11,53	16	470,5
Mónaco		100,00		13,80		10,34	255,95
Montenegro	-2400,00	66,81		4,00	10,59	1,45	53,52
Nepal	208549,00	19,74	4,734	0,30	6,40	0,07	25,55
Nicaragua	-106360,00	58,52	16,208	0,90	5,05	0,99	28,2
Nigeria	-300000,00	50,34	16,514	0,50	12,11	0,24	9,05
Noruega	140000,00	82,25	19,047	3,80	7,70	4,57	163,63
Nueva Zelanda	74403,00	86,54	31,867	2,80	6,95	0,45	31,13
Omán	437000,00	84,54	29,953	1,60	2,46	2,4	540
Países Bajos	80000,00	91,49	12,416	4,70	8,80	35,36	286,75
Pakistán	-1166895,00	36,67	20,072	0,60	6,99	1,46	75,45
Panamá	56000,00	67,71	42,699	2,30	5,05	11,3	541,22
Papua Nueva Guinea	-3999,00	13,17		4,00	7,52	0	0,09
Paraguay	-82780,00	61,59	46,322	1,30	5,47	0,19	18,8
Perú	495345,00	77,91	32,482	1,60	5,44	22,69	753,09
Polonia	-146976,00	60,06	4,655	6,50	10,60	3,39	80,85
Portugal	-30001,00	65,21	41,182	3,40	10,60	14,81	366,32
Reino Unido	1303250,00	83,40	26,819	2,80	9,20	63,43	450,27
Rep. Checa	110057,00	73,79	12,155	6,50	10,50	3,13	95,6
Rep. Dominicana	-150000,00	81,07	29,849	1,60	6,12	5,96	226,82
Ruanda	-44998,00	17,21	8,599	1,60	5,31	0,02	5,19
Rumanía	-369997,00	54,00	9,353	6,30	13,30	7,45	116,92
Rusia	912279,00	74,43	23,191	8,20	12,90	5,3	388,36
Samoa	-14013,00	18,24		1,00	5,30	0	0
San Marino		97,23		3,80		124,33	2060,39
Serbia	20000,00	56,09	19,899	5,70	14,80	3,68	179,33
Singapur	-1000,00	100,00	100,000	2,40	5,00	0,46	730,95
Siria				1,50		0,04	1,05
Somalia	-200002,00	44,97	13,870	8,70	11,06	0,59	17,96
Sri Lanka	-489932,00	18,48		3,60	6,55	0,05	8,88



Suazilandia	-41764,00	23,80		2,10	9,84	0,35	49,55
Sudàfrica	727026,00	66,36	35,308	2,80	9,50	2,9	139,17
Sudán	-250001,00	34,64	13,239	0,80	7,26	1,17	19,19
Suecia	200000,00	87,43	15,545	2,40	9,10	49,54	536,22
Suiza	259999,00	73,80	16,095	4,70	7,90	22,98	366,34
Surinam	-4999,00	66,06		3,10	7,28	1,04	45,31
Tailandia	97222,00	49,95	20,055	2,10	7,56	0,08	4,52
Tanzania	-200381,00	33,78	12,520	0,70	6,62	0,04	0,9
Tonga	-3999,00	23,13		2,60	7,19	0	0
Trinidad y Tobago	-3999,00	53,18		3,00	8,26	0,58	8,85
Túnez	-20000,00	68,95	19,807	2,20	6,27	0,43	9,75
Turquía	1419610,00	75,14	36,855	2,70	5,39	5,91	221,97
Ucrania	50001,00	69,35	12,110	8,80	14,50	2,19	78,02
Uruguay	-15000,00	95,33	50,358	2,80	9,46	0,7	24,61
Uzbekistán	-44314,00	50,48	7,477	4,00	5,00	0,06	17,29
Venezuela	-3266243,00	88,21	32,352	0,80	6,81	0,1	11,73
Yemen	-150000,00	36,64	16,910	0,70	5,98	0,86	3,17
Zambia	-40000,00	43,52	14,545	2,00	6,63	0,06	8,14
Zimbabue	-584288,00	32,21	10,493	1,70	8,04	0,03	2,78

Gràfics no incluídos

Gràfico de Caja y Bigotes de la variable PARO en grupos de la variable JABON.



Gràfico de Caja y Bigotes de la variable PIB en grupos de la variable JABON.

Gráfico XY de las variables
DESNUT y PARO.

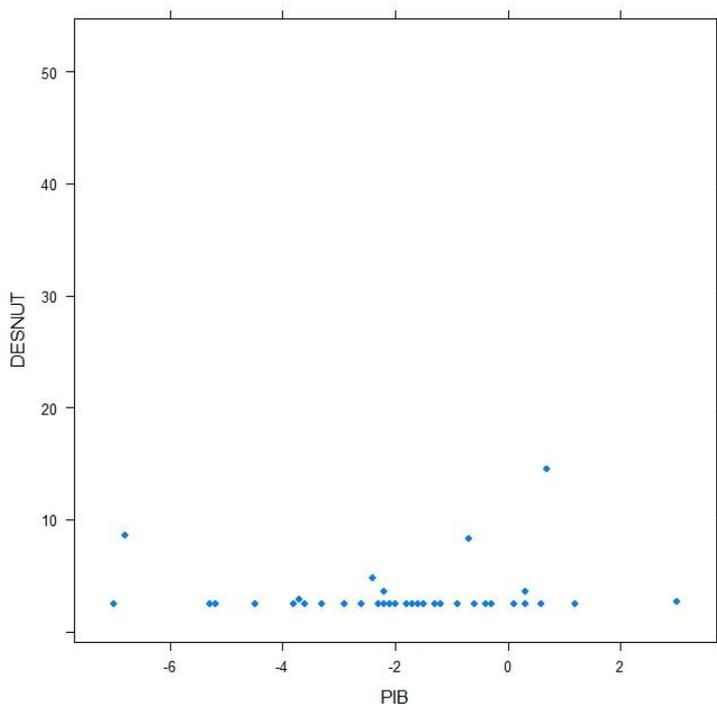
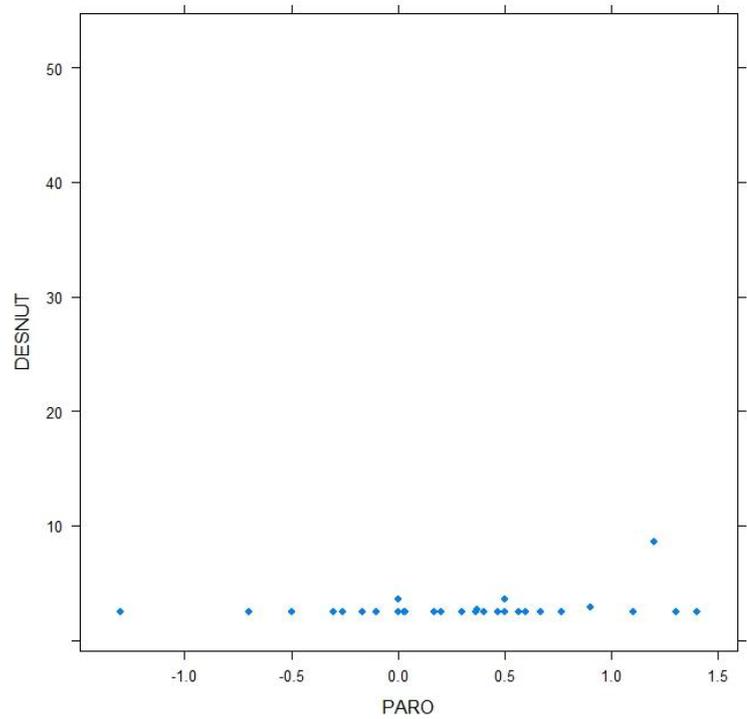


Gráfico XY de las variables DESNUT
y PIB.