



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

CAMPUS D'ALCOI

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Politécnica Superior de Alcoy

Estudio de la COVID-19 por comunidades autónomas españolas. Modelización econométrica mediante series temporales.

Trabajo Fin de Grado

Grado en Administración y Dirección de Empresas

AUTOR/A: Parra Almendros, Carlos

Tutor/a: Carracedo Garnateo, Patricia

Cotutor/a: Hervás Marín, David

CURSO ACADÉMICO: 2022/2023

RESUMEN

La pandemia de coronavirus en el mundo ha provocado una crisis sanitaria sin precedentes donde, a su vez, ha derivado en la paralización de ciertos sectores económicos, como el comercio, la hostelería y el turismo. A día de hoy, se han registrado un total de 13.740.531 casos positivos y 118.712 muertes. Los datos de este trabajo se descargaron del Instituto de Salud Carlos III (<https://cnecovid.isciii.es/covid19/>). La base de datos utilizada tiene en cuenta la evolución temporal de 58.708 brotes de COVID-19 detectados en 19 comunidades autónomas (CCAA) durante el periodo comprendido entre el 1 de enero de 2020 y el 29 de agosto de 2022.

El objetivo de este trabajo es realizar un estudio econométrico sobre el impacto del COVID-19 en las personas mayores de 60 años en las distintas CCAA. Para la metodología de este trabajo se han utilizado series temporales con el fin de estudiar la relación causal entre las diversas variables. El software estadístico utilizado ha sido R, un lenguaje de programación especializado en la visualización, tratamiento y análisis de datos.

Palabras clave: Análisis de datos, COVID-19, Comunidades autónomas, Series temporales, modelos ARIMA

RESUM

La pandèmia del Coronavirus al món ha provocat una crisi sanitària sense precedents on, al seu torn, ha derivat en la paralització d'uns certs sectors econòmics, com el comerç, l'hostaleria i el turisme. Hui dia, s'han registrat un total de 13.740.531 casos positius i 118.712 morts. Les dades d'aquest treball es van descarregar de l'Institut de Salut Carles III (<https://cneccovid.isciii.es/covid19/>). La base de dades utilitzada té en compte l'evolució temporal de 58.708 brots de COVID-19 detectats en 19 comunitats autònomes (CCAA) durant el període comprés entre l'1 de gener de 2020 i el 29 d'agost de 2022.

L'objectiu d'aquest treball és realitzar un estudi economètric sobre l'impacte del COVID-19 en les persones majors de 60 anys en les diferents CCAA. Per a la metodologia d'aquest treball s'han utilitzat sèries temporals amb la finalitat d'estudiar la relació causal entre les diverses variables. El programari estadístic utilitzat ha sigut R, un llenguatge de programació especialitzat en el visualització, tractament i anàlisi de dades.

Paraules clau: Anàlisi de dades, COVID-19, Comunitats autònomes, Sèries temporals, models ARIMA

ABSTRACT

The global coronavirus pandemic has caused an unprecedented health crisis which, in turn, has led to the paralysis of certain economic sectors, such as commerce, hospitality, and tourism. To date, a total of 13,740,531 positive cases and 118,712 deaths have been recorded. The data for this work were downloaded from the Instituto de Salud Carlos III (<https://cnecovid.isciii.es/covid19/>). The database used takes into account the temporal evolution of 58,708 outbreaks of COVID-19 detected in 19 Autonomous Communities (AC) during the period from 1 January 2020 to 29 August 2022.

The aim of this work is to carry out an econometric study on the impact of COVID-19 on people over 60 years of age in different Autonomous Communities. For the methodology of this work, time series have been used to study the causal relationship between the various variables. The statistical software used was R, a programming language specialized in data visualization, processing, and analysis.

Keywords: Data analysis, COVID-19, Autonomous communities, Time Series, ARIMA models

A mi familia, y en especial a mis padres.

Agradecimientos

Este trabajo sólo ha sido posible como fruto y consecuencia de las personas que me han acompañado en mi camino, del mismo modo que existen ciertos materiales que se obtienen únicamente cuando se les aplica cierta presión, frío o calor.

En primer lugar, dar las gracias a mi familia de sangre. A mis padres, a quienes dedico este trabajo, a mi madre por su coraje y a mi padre por su bondad. Agradecer la suerte de tener a mis hermanos quienes son mis mejores compañeros en esta aventura. Gracias a los 5 magníficos y a toda mi familia por su apoyo y su cercanía.

En segundo lugar, agradecer el camino recorrido a mi familia escogida, a Silvia por ser mi guía cuando vagaba sin rumbo, a David por su luz y cariño, a Santos, *Charly*, Rocío, Carlos y Eduardo por ser objeto de inspiración y compañía. A mi grupo de amigos más cercano, José Miguel, Christopher, Luis Miguel y José Luis, por haber disfrutado tantos años a su lado.

Agradecer también a mis compañeras alcoyanas estos intensos años juntos, y a Noelia por la gran aventura que tenemos por delante y que nos queda por recorrer.

Por último, dar las gracias también a mis compañeros castellanomanchegos, Toni, Daniel y Miguel, por su alegría contagiosa y sus ganas de vivir. Su compañía ha sido un privilegio en esta etapa final universitaria.

Para concluir, y siendo el SARS COVID-19 tema central de este trabajo, agradecer por un lado a los profesionales que combatieron contra viento y marea este suceso único que nos llevó a todos al límite; y dedicar este trabajo a todos aquellos que nos dejaron y cuya memoria siempre recordaremos.

Índice General

1. Introducción	8
1.1 Contexto	8
1.2 Motivación y objetivos	15
2. Metodología	15
2.1 Obtención de la información y base de datos utilizada	15
2.2 Modelos ARIMA.....	17
2.3 Ajuste y selección de modelos ARIMA	19
2.4 Series temporales: Análisis y descomposición	20
3. Resultados	23
3.1 Análisis descriptivo	23
4. Análisis y resultados	23
4.1 Análisis de la expansión del COVID-19 por provincias	36
4.2 Impacto económico del COVID-19 en España.....	39
5. Conclusiones	42
6. Referencias y bibliografía	44
7. Anexos	47

Índice de figuras

Figura 1.1 Nuevos casos diarios en España 2020 – 2022 (Fuente: ELDiario).....	8
Figura 1.2 Peso de los sectores más afectados sobre el VAB (Fuente: Banco de España).....	15
Figura 1.3 Sectores con mayor incidencia de ERTES (Fuente: Banco de España)	15
Figura 1.4 Políticas de contención y recesión españolas (Fuente: BBC)	17
Figura 2.1 Modelo ARIMA de Cataluña y sus componentes. Elaboración propia.....	19
Figura 2.2 Modelo ARIMA de la Provincia de Madrid. Elaboración propia.....	20
Figura 2.3 Modelo ARIMA de las Islas Canarias. Elaboración propia	20
Tabla 1. Modelos ARIMA de las Comunidades Autónomas	20
Figura 2.4 Serie temporal sobre el nº casos en Murcia. Elaboración propia	22
Figura 3.1 Serie temporal sobre el nº hospitalizados en Asturias. Elaboración propia	24
Tabla 2. Peso de los sectores más afectados sobre el VAB (Fuente: Banco de España)	25
Figura 3.2 Serie temporal sobre el nº casos en País Vasco. Elaboración propia.....	27
Figura 3.3 Serie temporal sobre el nº de casos en Cataluña. Elaboración propia	29
Figura 3.4 Serie temporal sobre el nº casos en la prov. de Barcelona. Elaboración propia ...	30
Figura 3.5 Serie temporal sobre el nº casos en la C.Valenciana. Elaboración propia	31
Figura 3.6 Serie temporal sobre el nº casos en la C.Madrid. Elaboración propia.....	32
Figura 3.7 Serie temporal sobre el nº hospitalizados en la C.Madrid. Elaboración propia.....	33
Figura 3.8 Serie temporal sobre el nº casos en CLM. Elaboración propia	34
Figura 3.9 Serie temporal sobre el nº casos en I.Canarias. Elaboración propia.....	35
Figura 4.1 Serie temporal sobre el nº casos de la prov. de Madrid. Elaboración propia.....	37
Figura 4.2 Serie temporal sobre el nº casos de la prov. de Barcelona. Elaboración propia ...	37
Figura 4.3 Modelo ARIMA de Cataluña. Elaboración propia.....	38
Figura 4.4 Modelo ARIMA de la Comunidad de Madrid. Elaboración propia	38
Figura 4.5 Modelo ARIMA de País Vasco. Elaboración propia	39
Figura 4.6 Modelo ARIMA de Castilla-La Mancha. Elaboración propia.....	39
Figura 4.7 Modelo ARIMA de Islas Canarias. Elaboración propia	40
Figura 4.8 Modelo ARIMA de Comunidad Valenciana. Elaboración propia.....	40
Figura 4.9 Descenso del PIB en 2020. (Fuente: Banco de España).....	41
Figura 4.10 Descenso del PIB por provincias en 2020. (Fuente: Banco de España).....	42
Figura 4.11 Caída del turismo en 2020 en porcentaje. (Fuente: Banco de España)	43
Figura 4.12 Evolución de la facturación por sectores en 2020. (Fuente: Banco de España) ..	44

1. Introducción

1.1. Contexto

El día 11 de marzo de 2020, la organización mundial de la salud “OMS” (WHO por sus siglas en inglés) declara al SARS COVID-19 (del inglés *severe acute respiratory syndrome coronavirus disease 2019*) como pandemia.

La OMS tuvo noticia por primera vez de la existencia de este nuevo virus el 31 de diciembre de 2019, al ser informada de un grupo de casos de «neumonía vírica» que se habían declarado en Wuhan, Hubei (República Popular China).

Este virus se caracteriza por una transmisión no homogénea, que puede propagarse desde la boca o nariz de una persona infectada en pequeñas partículas líquidas cuando tose, estornuda o habla. La mayoría de las personas infectadas por el virus experimentarán una enfermedad respiratoria de leve a moderada y se recuperarán sin requerir un tratamiento especial. Sin embargo, algunas enfermarán gravemente y requerirán atención médica. Las personas mayores y las que padecen enfermedades subyacentes, como enfermedades cardiovasculares o respiratorias, pueden tener un mayor riesgo de desarrollar una enfermedad grave (OMS, 2020). Diferencias en la densidad de población y comportamiento cultural influyen en las dinámicas de transmisión del virus, por lo tanto, podríamos afirmar que la transmisión no es homogénea, sino heterogénea (López-Mendoza, *et al.*, 2021). Que la transmisión sea heterogénea dificulta su control en cada una de las comunidades autónomas, como consecuencia de la incapacidad del sistema sanitario para poder predecir de qué manera va a contagiarse o expandirse por el territorio nacional.

El primer caso confirmado en España se produjo en las Islas Canarias el 31 de enero de 2020 (Orea, L., & Álvarez, I. 2022). Por otro lado, los primeros casos detectados en la Península se darían en Madrid en el mes de marzo. El día 9 de dicho mes el Gobierno Español y específicamente el ministerio de Sanidad comenzó a fomentar un distanciamiento social para disminuir el contagio del virus. Las comunidades de Madrid, La Rioja, y País Vasco prohibieron la docencia en espacios cerrados, esto es, en universidades y colegios. (Gutiérrez, M. J., et al 2021).

En cuanto a su transmisión interterritorial, varió en las distintas comunidades autónomas; durante los días 9, 10 y 11 de marzo todas las comunidades autónomas comenzaron a imponer medidas para combatir la pandemia.

El 11 de marzo de 2020, había 4400 casos confirmados de contagio, de los cuales más de 2900 pertenecían a Madrid. El resto de los casos se distribuyeron de manera heterogénea a lo largo de las 51 provincias. Sólo tres provincias (Barcelona, La Rioja y Álava), tuvieron una cifra ligeramente superior de 200 casos, y otras veintidós provincias no tenían ningún contagio. El 14 de marzo, todas y cada una de las provincias españolas tenía pacientes infectados (Gutiérrez, M *et al.*, 2021). Es esta fecha cuando el Gobierno español declaró el Estado de Alarma en todo el territorio. (BOE 2020a). Esta legislación permite al gobierno restringir la movilidad de los ciudadanos dentro del país, así mismo también limitar la actividad económica de sectores esenciales.

A lo largo del mes de marzo se produce una transmisión explosiva de casos, desembocando, el 4 de abril de 2020, en 126.859 casos confirmados en territorio español (Orea, L., *et al.*, 2022).

Al comienzo de julio de 2020, cuando la primera oleada era considerada como controlada, España era el quinto país del mundo con mayor ratio de fallecidos por habitante, después de San Marino, Bélgica, Andorra y el Reino Unido (OMS, 7 de julio de 2021, Gutiérrez, M *et al.*, 2021).

Con respecto a la transmisión del virus en las comunidades autónomas españolas, el enfoque llevado a cabo ha consistido en el uso de una estrategia de mitigación, para prevenir el colapso del sistema sanitario, detectando el mayor número de casos posibles pese a la dificultad a consecuencia de que el 33% de los pacientes de SARS-CoV-2 en la primavera de 2020 eran asintomáticos (Pollán, M.; Pérez-Gómez, 2020). Medidas tales como el uso de mascarillas faciales que protegen de la propagación del virus, el control de la movilidad interterritorial para limitar la interacción social entre los ciudadanos de las distintas comunidades, especialmente en espacios cerrados a consecuencia de la alta tasa de propagación en áreas poco ventiladas (López-Mendoza, Montañés, & Moliner-Lahoz, 2021).

Durante el primer mes del Estado de Alarma, ninguna unidad familiar podía abandonar el estricto confinamiento en sus hogares, excepto por actividades esenciales como trabajar, ir al hospital o comprar alimentos. Pese a que las restricciones se declararon de manera reactiva a medida los casos desbordaron el sistema sanitario español, las medidas se

aplicaron con rapidez, impidiendo la circulación de personas tanto en espacios cerrados como en áreas ventiladas. Siendo el fin principal evitar la aglomeración de personas que pudiera favorecer el aumento de la transmisión del virus.

Restricciones sociales estrictas han resultado efectivas curvando la incidencia del COVID-19 (López-Mendoza, *et al.*, 2021). A medida que la efectividad de las medidas tomadas reducía el impacto de la transmisión, comenzaban a eliminarse o relajar su uso. Es en verano del mismo año, cuando se produce una segunda ola de transmisión en España; esta segunda ola se produce a consecuencia de la aparición de una nueva variante de SARS-COVID-19 que emerge y se expande por toda Europa con la mayor tasa de incidencia entre jóvenes adultos.

España es uno de los países más afectados por la pandemia, a fecha de 7 de enero de 2021, con 1.893.502 casos confirmados, y 50.442 muertes, en una población de sobre 47 millones (Gutiérrez, M. J., Inguanzo, B., & Orbe, S. 2021).

El Sistema sanitario español está basado en una cobertura universal, financiado con dinero público, y con productos farmacéuticos de pago dependiendo de la edad y la renta (Mentzakis *et al.* 2019). Las 17 comunidades autónomas tienen autoridad para proveer a sus ciudadanos de servicios sanitarios haciendo uso de esta cobertura social perteneciente al Estado. Con la declaración del Estado de Alarma, se llevó a cabo una gestión conjunta tanto de las políticas restrictivas como de los servicios sanitarios regionales. El Gobierno Central establecía directrices generales que cada comunidad autónoma debía seguir, y eran éstas últimas quienes las organizaban y llevaban a cabo en las correspondientes provincias.

Llevando a cabo una gestión descentralizada de la pandemia a nivel regional, permitió que en aquellas zonas con una mayor densidad de población estuvieran sujetas a políticas de aislamiento y prevención mucho más restrictivas, para poder evitar una mayor propagación.

La expansión del COVID-19 en provincias con una gran presencia del sector servicios fue considerablemente mayor, en contraste con regiones de gran número de trabajadores pertenecientes al sector agrícola y/o construcción que presentaron tasas menores de contagio. El riesgo de contagio en el sector servicios es mayor que en los dos sectores citados previos, a causa de que estos últimos desarrollan su actividad laboral en espacios abiertos, mientras que el primero lo realiza principalmente en espacios interiores (Orea, L., & Álvarez, I. C., 2022).

Esta situación pandémica se prolongaría a lo largo de 2020 y 2021, con las 6 sucesivas olas de variantes que azotarían a la península, aumentando o descendiendo las restricciones ciudadanas en función de la influencia e intensidad de las restricciones decretadas de forma territorial bajo la influencia y los estragos que este virus ha ocasionado en el territorio español.

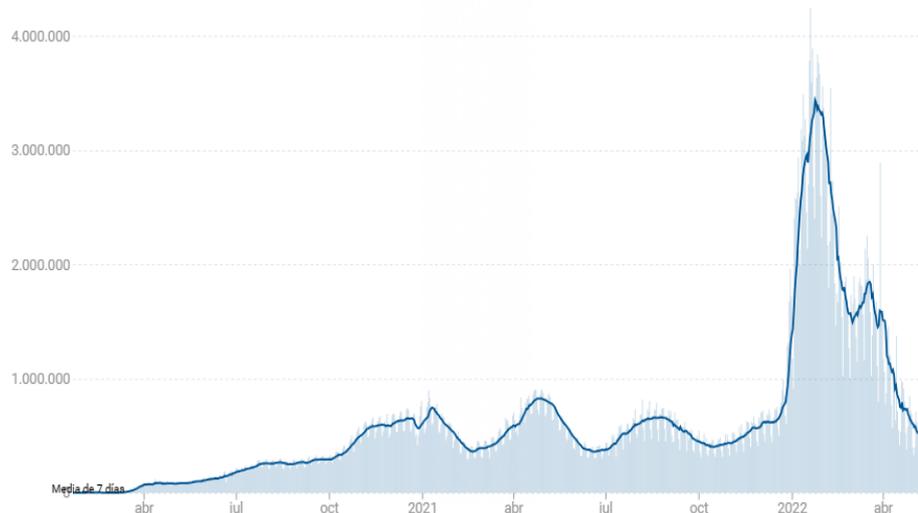


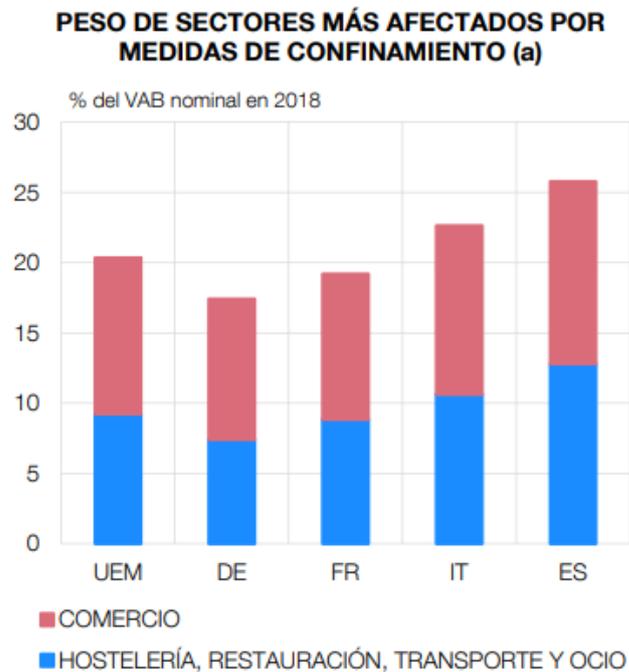
Figura 1.1 Nuevos casos diarios en España 2020 - 2022 (Fuente: *ElDiario.es*, 2022)

El día 9 de mayo el Gobierno español elimina el Estado de Alarma, a través del Real Decreto 926/2020, descentralizando las medidas para que sean las propias Comunidades Autónomas aquellas que decidan la intensidad de las medidas en función de la transmisión de esta patología.

Asegurar la coordinación del gobierno nacional junto a los gobiernos regionales ha sido uno de los puntos clave respecto a la resiliencia del sistema sanitario español durante las primeras semanas de la pandemia (Legido-Quigley *et al.* 2020). Esta coordinación mutua ha permitido que el desarrollo de la propagación y transmisión de esta patología haya sido heterogéneo como hemos citado a principio de este documento, y haya ocurrido a lo largo de varias olas de contagio durante los años 2020 y 2021 bajo restricciones que reducían la movilidad social de los ciudadanos.

Posteriormente, en referencia a los efectos económicos consecuencia del COVID-19, España durante la crisis sanitaria ha experimentado ciclos de escalada y desescalada a nivel regional y a nivel local, con medidas que debilitan la economía para poder mejorar la cobertura sanitaria, mientras esperaba la vacunación de grupos vulnerables. La estrategia española de mitigación llevada a cabo por el Gobierno se ha llevado a cabo con cierta heterogeneidad territorial (López-Mendoza *et al.*, 2021).

Estas medidas incluye el cese de la actividad de empresas no esenciales, esto es, se suspende la apertura al público de los locales y establecimientos minoristas, a excepción de los establecimientos comerciales de alimentación, bebidas, productos y bienes de primera necesidad, así como farmacéuticos o de combustible para la automoción además de una reducción de la oferta de transporte público en un 50%.

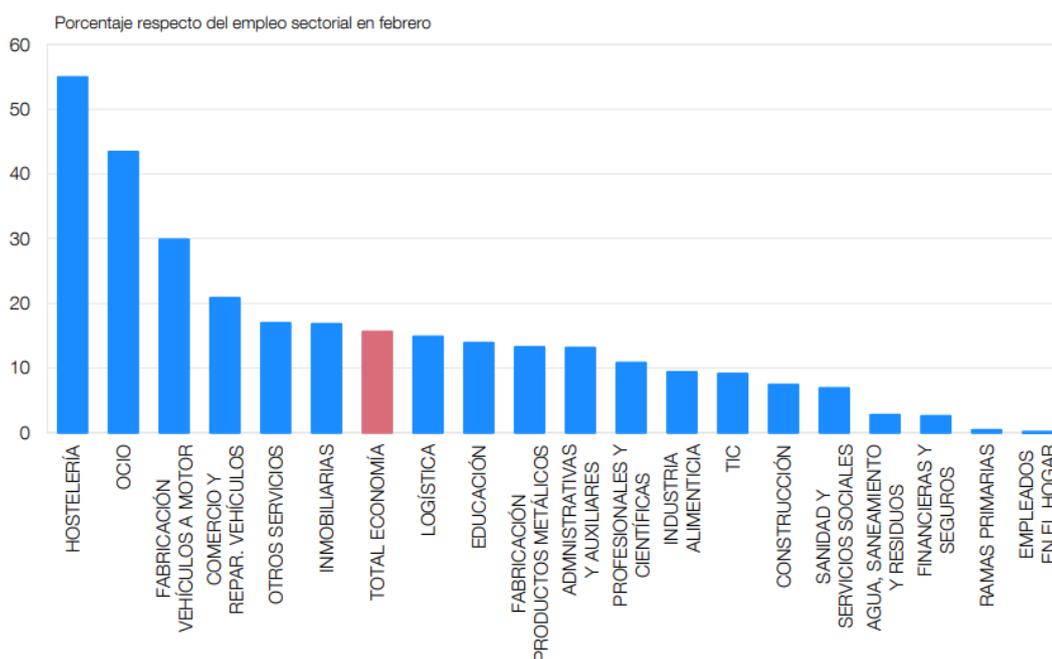


Fuentes: Eurostat, IHS Markit y Universidad de Oxford.

Figura 1.1. Peso de los sectores más afectados sobre el VAB (BDE 2020)

Desde el Gobierno español con la finalidad de paliar la contracción significativa de la demanda y del consumo doméstico, se impulsaron medidas como la creación de ERTES, expedientes de regulación temporal de empleo, que permitían la suspensión temporal de empleados con garantía de prestación por desempleo financiada por el Estado, incluidos aquellos que no hubieran cotizado el periodo mínimo exigido para ello.

INCIDENCIA DE ERTES EN MAYO POR SECTORES DE ACTIVIDAD



Fuente: Ministerio de Inclusión, Seguridad Social y Migraciones.

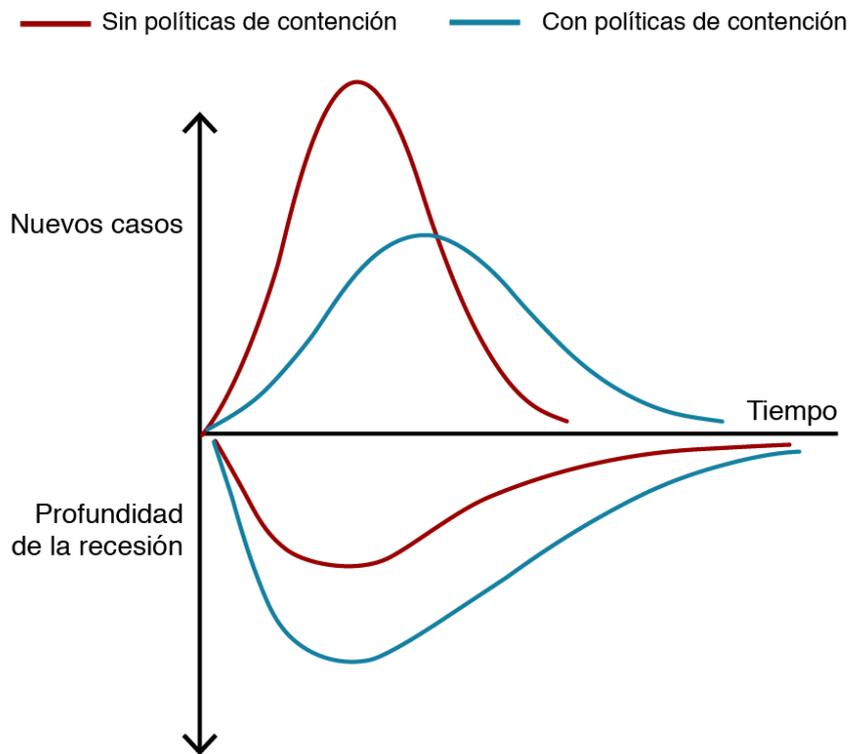
Figura 1.2. Sectores con mayor incidencia de ERTES (Banco de España, 2020).

Entre los factores de demanda destaca la contribución negativa de las perturbaciones de consumo privado y de la demanda de viviendas en el segundo trimestre de 2020, detrayendo 12 puntos porcentuales al PIB. Conviene subrayar la importante contribución del crecimiento del crédito bancario durante el segundo trimestre de 2020 (equivalente a 4,5 puntos porcentuales de PIB), como consecuencia de la política monetaria expansiva adoptada por el BCE, permitiendo un alivio económico que ha facilitado que el sistema bancario haga frente al enorme aumento de las necesidades de financiación de empresas durante la crisis (José E. Boscá *et al.* 2020).

El uso del Estado de Alarma en España, que restringía la actividad a todos los sectores excepto aquellos considerados esenciales, a saber, alimentación y sanitario (farmacéutica y hospitales), tuvo resultados positivos aplanando la curva de casos contagiados desde el 29 de marzo al 13 de abril. (Saez *et al.* 2020; Tobías 2020).

Mientras los casos descendían a medida que aumentaban la intensidad de las restricciones impuestas desde el Gobierno Central, la economía padecía los efectos de una grave contracción del consumo. Se estima que el PIB español puede haber sufrido una contracción de 26.900 millones de euros desde que se declaró la epidemia el 7 de marzo de 2020. Una respuesta rápida a nivel institucional puede salvar un gran número de vidas, y atenuar el impacto económico que esta pandemia ha ocasionado. (Orea, L., & Álvarez, I. C,

COVID-19: la doble curva médica y económica



Fuente: Richard Baldwin, inspirado en ilustraciones de Pierre-Olivier Gourinchas. **BBC** 2022).

Figura 1.3. Políticas de contención y recesión española(www.bbc.com/mundo/noticias-52158532)

1.2 Motivación y objetivos

La crisis del COVID-19 ha supuesto un suceso histórico provocando como consecuencia cambios en el comportamiento sociológico entre ciudadanos, incorporando conceptos como la distancia de seguridad, y acelerando cambios en los procesos productivos y empresariales, como el impulso y aumento del teletrabajo de manera generalizada. La distribución del SARS COVID-19 por todo el territorio nacional supuso un esfuerzo económico y sanitario que afectó de distinta manera a las comunidades autónomas.

En este trabajo se estudian los patrones y rasgos del contagio del coronavirus, cómo afectó a la población de cada una de las regiones que constituyen el Estado Español a lo largo de los años 2020 – 2023.

El objetivo principal de este Trabajo Final de Grado consiste en el análisis del contagio del COVID-19 en cada una de las comunidades autónomas españolas, así como la introducción en un amplio campo de estudio como es el estadístico, y específicamente econométrico, aprendiendo y practicando considerablemente el lenguaje estadístico R, y sus distintas “libraries” para poder desarrollar modelos y series temporales que permitan explicar de qué manera afectó a las comunidades autónomas, y qué patrones o características presentó el COVID-19 en todo el territorio español.

Este trabajo resultará interesante para poder predecir de qué manera se produce la transmisión del SARS COVID-19, como es su expansión y contagio, y qué se puede aprender para poder afrontar situaciones sanitarias similares. Del mismo modo se refleja cómo afectan crisis sanitarias a la demanda y el consumo de un país, pudiendo observarse las contracciones económicas consecuentes de las medidas y decisiones que se toman a nivel político-económico y sanitario.

2. Metodología

2.1 Obtención de la información y base de datos utilizada

Se ha realizado un estudio y análisis econométricos centrado específicamente en las personas mayores de 60 años, con respecto a todas y cada una de las provincias españolas.

La base de datos utilizada pertenece al Instituto de Salud Carlos III donde es posible observar que los datos recogidos están divididos en variables, que posteriormente explicaremos, y que corresponden a la población de 60 o más años. Dichas variables las d: número de contagios, número de hospitalizados, número de ingresos en UCI, y número de defunciones; del mismo modo se han clasificado por sexo, y provincia de residencia.

Del mismo modo se ha utilizado la información elaborada por el Banco de España (BDE) en el año 2021, con la finalidad de reflejar cómo la crisis sanitaria afectó a la actividad económica del país, cómo afecto a los distintos sectores que constituyen el mercado laboral español, y de qué manera se vio afectada la oferta y demanda, pudiendo observarse modificaciones y concretamente contracciones, como consecuencia de las medidas tomadas para paliar los efectos provocados por la expansión del COVID-19 a lo largo de los años 2020 y 2023 durante las comunidades autónomas que conforman el país.

Haciendo uso del citado previamente RStudio, se expone un análisis descriptivo que permita reflejar el impacto ocasionado por esta pandemia mundial en las comunidades autónomas españolas, a través de una muestra de más de 60.000 observaciones, entre las 17 comunidades autónomas que representan el territorio español, a su vez divididas en 52 provincias y dos ciudades autónomas.

Las variables utilizadas en los análisis descriptivos de cada provincia y la correspondiente Comunidad Autónoma las podemos dividir en dos tipos, numéricas y categóricas.

Las variables numéricas y categóricas son dos tipos de variables que se utilizan en la estadística y en el análisis de datos, en la base de datos utilizada en este trabajo se clasifican según:

Variables numéricas:

Nº Casos: Habitantes mayores de 60 años contagiados por COVID-19 en cada provincia española.

Nº Hospitalizados: Habitantes contagiados que han sido trasladados a un hospital a consecuencia de la gravedad de su situación.

Nº Hospitalizados UCI: Unidad de cuidados intensivos, hace referencia al área específica de un hospital donde se integran pacientes en estado crítico o muy grave de salud.

Nº Defunciones: Fallecimientos provocados como causa y consecuencia del SARS COVID-19.

Longitud: Es la distancia en grados, minutos y segundos que hay con respecto al meridiano principal, que es el meridiano de Greenwich (0º).

Latitud: La latitud proporciona la localización de un lugar, en dirección Norte o Sur desde el ecuador y se expresa en medidas angulares que varían desde los 0º del Ecuador hasta los 90º N del polo Norte o los 90º S del Polo Sur.

Variables categóricas:

Provincia: Demarcación administrativa española de cuyos habitantes extraemos los datos.

2.2 Modelos ARIMA

Con la finalidad de explicar y reflejar la incidencia y distribución del COVID-19 en España, se aplicarán series temporales a los datos recogidos en nuestra base de datos para cada una de las Comunidades Autónomas, y se explicará la evolución del contagio, así como patrones y rasgos observables a través de un extenso análisis descriptivo de los datos.

Para poder comprender y describir adecuadamente los datos observados en función del tiempo, se va a realizar el ajuste y la elaboración de un modelo econométrico donde incluiremos las variables numéricas y categóricas descritas previamente.

Posteriormente se ajustará el modelo ARIMA de acuerdo con la base de datos extraída con respecto a las distintas comunidades autónomas, para poder observar cual se ajusta más, haciendo uso de la función "auto.arima" en RStudio.

Estos modelos se componen de tres parámetros, números enteros no negativos; el orden del componente autorregresivo "p", "d" orden de integración refleja el número de veces que se ha aplicado diferenciación para que la serie temporal sea estacionaria, y "q" es el orden del componente de media móvil.

```
ARIMA(2,1,4)

Coefficients:
      ar1      ar2      ma1      ma2      ma3      ma4
      1.2467 -0.8256 -1.7648  1.1682 -0.0261 -0.1519
s.e.  0.0303  0.0833  0.0522  0.1747  0.1512  0.0465

sigma^2 = 263.7: log likelihood = -4860.45
AIC=9734.9  AICc=9735  BIC=9770.27
```

Figura 2.1 Modelo ARIMA de Cataluña y sus componentes. Elaboración propia.

Por un lado, observamos que estos modelos presentan varios componentes autorregresivos de distinta orden; "ar1" hace referencia a que sólo se utiliza el valor inmediatamente anterior en la serie de tiempo para hacer predicciones.

Por otro lado, los componentes contiguos hacen referencia a medias móviles. Estos componentes tras el uso de la función auto.arima, indican el uso de errores pasados en la serie temporal para predecir valores futuros. Del mismo modo que "ar1", "ma1" significa que está utilizando tan sólo el error inmediatamente anterior en la serie de tiempo para predecir el siguiente valor.

La función `auto.arima` es una función estadística empleada para ajustar modelos ARIMA (Auto Regressive Integrated Moving Average), esto es, Modelo autorregresivo integrado de media móvil, que haciendo uso de variaciones y regresiones de datos estadísticos, nos permitirá finalmente predecir valores futuros, basándose en patrones de datos históricos previos. Esta función ajusta un modelo a los datos de la serie de tiempo y, una vez ajustado, se puede utilizar de forma óptima para realizar predicciones sobre los valores futuros de la serie.

Por otro lado, el AIC (Akaike Information Criterion o Criterio de Información de Akaike) es un criterio estadístico utilizado para seleccionar modelos en el análisis de series de tiempo. El AIC representa una herramienta objetiva que permite de manera cuantificar cómo de idóneo es un modelo en base a los componentes del propio modelo. El objetivo es seleccionar aquel modelo ARIMA que posea un valor inferior de AIC. Cuanto más bajo sea el valor del AIC, mejor será el modelo en términos de ajuste y simplicidad. El AIC se calcula utilizando la siguiente fórmula:

$$\text{AIC} = 2k - 2\ln(L)$$

Donde:

k es el número de parámetros estimados en el modelo.

ln(L) es el logaritmo natural con respecto a la función de verosimilitud del modelo, que mide la probabilidad de que los datos observados sean generados por el modelo.

La interpretación del AIC es que modelos con un valor más bajo son preferibles, lo que significa que tienen un buen equilibrio entre ajuste y complejidad.

2.3 Ajuste y selección de modelos ARIMA

Como se ha especificado previamente, para poder comprender y describir adecuadamente los datos observados en función del tiempo, se realiza un ajuste ARIMA, mediante el uso del comando `auto.arima`, que permite observar qué modelo de cada comunidad autónoma presenta una mayor complejidad que otras de tipo más simple.

Las series temporales con respecto a las comunidades autónomas más complejas serían las de la Comunidad de Madrid e Islas Canarias. Las 15 comunidades autónomas (y ciudades

autónomas) restantes presentan un comportamiento mucho más sencillo. La ciudad autónoma de Ceuta presenta un modelo complejo de similar manera, pero a consecuencia de la significativa diferencia en cifras de las dos comunidades previas se ha decidido por estas dos comunidades autónomas.

En estos modelos previos podemos observar hasta 5 componentes autorregresivos.

```
Series: TSMadrid
ARIMA(5,1,5)

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3      ma4      ma5
    0.2534 -0.9532 -0.0064 -0.4850 -0.5226 -0.9515  1.1085 -0.7010  0.4314  0.0629
s.e.  0.0501  0.0414  0.0678  0.0391  0.0444  0.0542  0.0702  0.0784  0.0625  0.0440

sigma^2 = 39960: log likelihood = -7762.52
AIC=15547.04  AICc=15547.27  BIC=15602.62
```

Figura 2.2 Modelo ARIMA de la Provincia de Madrid. Elaboración propia.

La complejidad derivada de estos modelos es consecuencia de su gran número de componentes autorregresivos, y cuya existencia puede verse a todo tipo de factores externos al propio modelo y serie temporal.

Para poder seleccionar adecuadamente el mejor modelo de entre los más complejos, y como consecuencia de que posean el mismo número de componentes en su modelo ARIMA, observaremos cuál de ellos minimiza su AIC, lo que supondrá un mejor equilibrio entre el ajuste y la complejidad del modelo.

```
Series: TSCanarias
ARIMA(5,1,5)

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3      ma4      ma5
    0.3009 -1.0212  0.0811 -0.5716 -0.4564 -0.8699  1.1458 -0.7405  0.6203 -0.0779
s.e.  0.0529  0.0441  0.0742  0.0423  0.0481  0.0548  0.0656  0.0807  0.0615  0.0429

sigma^2 = 545: log likelihood = -5279.52
AIC=10581.05  AICc=10581.28  BIC=10636.63
```

Figura 2.3 Modelo ARIMA de las Islas Canarias. Elaboración propia.

A continuación, la Tabla 1 es un resumen de todos los modelos obtenidos de las Comunidades Autónomas de España.

CCAA	Modelo ARIMA	AIC	AICC
Galicia	(3,1,2)	11039.22	11039.29
Asturias	(2,1,2)	11200.93	11200.99
Cantabria	(2,1,2)	10692.12	10692.17
País Vasco	(5,1,2)	9249.59	9249.72
Navarra	(2,1,3)	11267.03	11267.11
La Rioja	(5,1,2)	9856.58	9856.71
Castilla y León	(2,1,5)	10599.88	10600.1
Aragón	(5,1,4)	11388.7	11388.89
Cataluña	(2,1,4)	9734.9	9735
C.Valenciana	(5,1,4)	10105.17	10105.36
Islas Baleares	(2,1,2)	10532.42	10532.47
Extremadura	(5,1,2)	10175.79	10175.92
Madrid	(5,1,5)	15547.04	15547.27
CLM	(2,1,3)	10978.02	10978.09
Murcia	(5,1,2)	12122.84	12122.97
Andalucía	(2,1,2)	8986.64	8986.69
Islas Canarias	(5,1,5)	10581.05	10581.28
Ceuta	(5,1,5)	5772.36	5772.59
Melilla	(5,1,3)	5957.7	5957.85

Tabla 1. Modelos ARIMA de cada Comunidad Autónoma. Elaboración propia.

En esta tabla se puede observar tanto los componentes de los modelos ARIMA de cada una de las comunidades autónomas como su Criterio de Información de Akaike.

2.4 Series temporales: Análisis y descomposición

Para poder analizar las series temporales incluidas en este trabajo, se tendrá en cuenta las características y variaciones observables gráfica y visualmente. Una serie temporal es una secuencia de datos recopilados en un intervalo de tiempo específico, y a partir de una muestra o base de datos previa. Las series temporales se utilizan en gran variedad de campos, como la economía, ingeniería, la medicina o la econometría. Esta técnica de análisis de datos nos permite observar la evolución de los datos a lo largo del tiempo. Las unidades de tiempo sobre las cuales se sustentan las series temporales, pueden ser días, semanas, meses o años.

Observando la tendencia o evolución de los datos con el paso del tiempo, permite analizar si presenta una tendencia propia, su estacionalidad, e incluso poder predecir en base a la evolución previa de patrones y tendencias, cómo se desarrollarán los valores futuros. Estas series temporales pueden ser tanto univariantes o multivariantes dependiendo del número de variables observadas.

Es posible descomponer las series temporales en las siguientes características:

- Tendencia: De qué manera han evolucionado los datos, puede ser creciente o decreciente.
- Variación estacional: Patrones o fluctuaciones recurrentes en los datos a lo largo del tiempo, y que se repiten de manera regular en los distintos periodos de tiempo.
- Variación cíclica: Patrones o fluctuaciones que se observan igualmente a lo largo del tiempo, pero que se producen de manera irregular, y no periódica.
- Variación irregular: Son fluctuaciones producidas por sucesos que no tienen ningún tipo de periodicidad ni regularidad, y que aparecen como observaciones inusuales o valores atípicos.

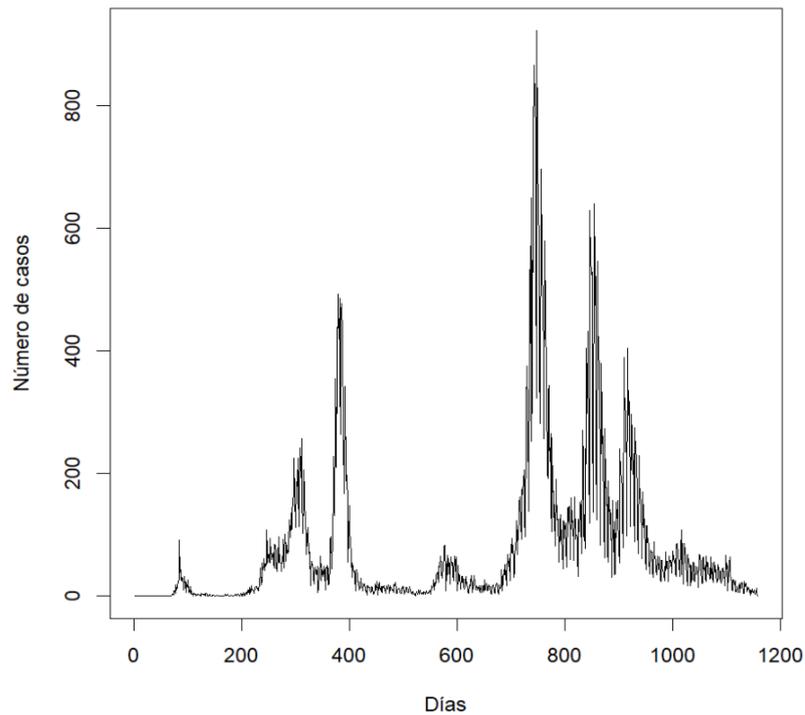


Figura 2.4 Serie temporal del nº casos positivos en Murcia. Elaboración propia.

Por otro lado, se podrá observar que en las distintas series temporales habrá dispersión irregular en el conjunto de datos que utilizamos para analizar la distribución del contagio de COVID-19 entre las distintas comunidades autónomas. Esto hace referencia a la varianza, cuánto se aleja cada valor individual de la media, la dispersión entre los propios datos. Esta distribución de la varianza entre las observaciones utilizadas en nuestra muestra se puede clasificar en dos tipos de distribuciones:

- Homocedasticidad: Distribución de la varianza de forma uniforme o constante, esto es, la dispersión que hay entre los datos se mantiene constante a lo largo de la serie temporal.
- Heterocedasticidad: La distribución de la varianza de manera no uniforme o constante, esto es, los datos se sitúan en la serie temporal sin seguir ningún tipo de patrón en la dispersión de estos.

A menudo, estas distribuciones de la varianza pueden apreciarse a simple vista. Para poder comprobar la asunción de homocedasticidad o de heterocedasticidad, es posible hacer uso de pruebas estadísticas con la finalidad de comprobar, objetivamente y con exactitud, si existe una distribución u otra.

En este trabajo se considera de gran importancia el uso de series temporales, a consecuencia de la posibilidad de explicar en este trabajo cómo se produjo la expansión de los contagios de COVID-19 en el territorio español, y de qué manera se mantuvo en el tiempo. Este lenguaje de programación permite el uso de las denominadas “librerías”, esto es, extensiones o herramientas creadas por distintos usuarios y que pueden aplicarse a bases de datos para elaborar todo tipo de gráficas o figuras, con el fin de poder explicar o analizar rasgos, características o fenómenos estadísticos de la base de datos utilizada.

Las figuras representadas en todas y cada una de las partes de este trabajo se han llevado a cabo mediante el lenguaje de programación R, y el software RStudio, de uso específico para el campo de la ciencia de datos y la estadística.

3. RESULTADOS

3.1 Análisis descriptivo

A continuación, se llevará a cabo un análisis de comunidades autónomas españolas, una breve descripción de éstas y de su grado de contagio por COVID-19, y posteriormente, y de acuerdo con la metodología explicada en el punto previo de este trabajo, se analizarán las series temporales creadas a partir de los datos extraídos de la base de datos de este trabajo.

Este análisis se llevará a cabo teniendo en cuenta las distintas variaciones citadas previamente en la metodología, tales como la ciclicidad, estacionalidad, tendencia o las posibles irregularidades presentes en estas series temporales correspondientes a las comunidades autónomas.

Tanto la tablas como las series temporales se han elaborado en el lenguaje R, y hacen referencia a casos de contagio en mayores de 60 años, creadas a partir de la base de datos extraída como hemos explicado detalladamente en la metodología.

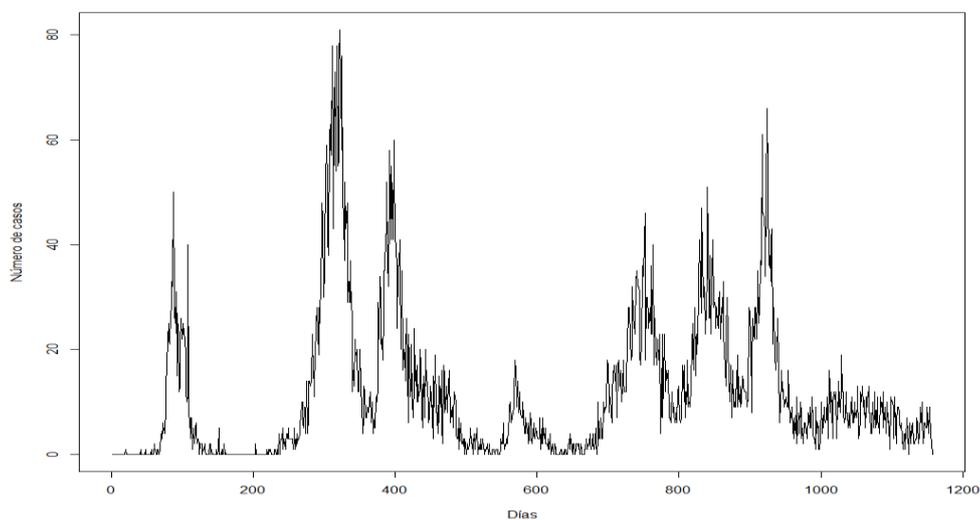


Figura 3.1 Serie temporal sobre el nº de hospitalizados en Asturias. Elaboración propia.

Se ha seguido un criterio geográfico en las comunidades autónomas analizadas comenzando por el área norte peninsular, para poder dotar de un orden objetivo y lógico en la explicación del análisis llevado a cabo de estas. En la explicación inicial de las distintas comunidades, se han redondeado y aproximado las cifras de contagio y para poder dar números enteros, de acuerdo con que los números que aquí se incluyen hacen referencia a personas.

En las series temporales con respecto al número de casos positivos por contagio de COVID-19 diarios, se observará que la expansión del COVID-19 por el territorio español no seguirá una distribución uniforme ni tendencias crecientes o decrecientes de forma general en todas las comunidades autónomas, sino que presentará en su mayoría variaciones irregulares fruto de la naturaleza misma de la distribución de los contagios, presentando valores atípicos de manera puntual, pese a que el mayor número de casos de contagio diarios en España se diese a comienzos de año, entre los meses de febrero y marzo de 2021.

Provincias	Nº Casos	Nº Hosp	Nº UCI	Nº Def
Madrid	383,5	83,75	6,24	17
Barcelona	336,1	33,5	2,35	13,32
Valencia	147,7	20,21	1,43	4,27
Alicante	92,59	15,42	1,15	3,17
Asturias	69,94	11,87	0,71	2,89
Sevilla	55,63	10,52	0,77	2,62
Zaragoza	66,18	10,42	0,91	3,22
Toledo	46,27	10,38	6,24	2,55
Vizcaya	87,85	10,16	0,75	3,39
La Coruña	72,12	9,7	1,16	1,14

Tabla 2 Datos de las CCAA a partir del comando “summary” en Rstudio. Elaboración propia.

Las comunidades del norte de España, a saber, desde Galicia hasta País Vasco, presentan una distribución relativamente homogénea en la concentración de casos desde comienzo del Estado de Alarma en el 2020, hasta el fin de las medidas contra el COVID-19 comenzado el año 2022.

La distribución de contagios en España se concentró en periodos específicos a lo largo de la línea temporal establecida en este trabajo, a saber, desde el año 2020 hasta 2023. Estos periodos específicos fueron denominados como “olas de contagio” por parte tanto del Gobierno español como de los distintos medios de comunicación.

En primer lugar, la primera ola de contagios se da en los meses de febrero y marzo de 2020, con el comienzo de la expansión de casos inicial en España. La segunda ola se dio a comienzos de mayo, cuando empiezan a suavizarse las medidas restrictivas de movilidad dentro del país, y comienza la reapertura de comercios. La tercera comenzaría a finales de 2020, en el mes de diciembre cuando comienzan las vacunaciones el día 27 de diciembre pese a que no descendería el número de contagios hasta el 28 de enero de 2021.

La cuarta ola se produce entre los meses de marzo y junio de este mismo año. La quinta oleada, consecuencia de la relajación en las restricciones establecidas a nivel autonómico y nacional, que fomenta una mayor movilidad intracomunitaria; comienza en junio y finaliza en octubre.

La sexta y última ola de contagios comienza con el fin de la previa, cuando aparece una nueva variante denominada “Ómicron”, que presentaba rasgos característicos tales como una mayor transmisión y contagio, y que alcanzaría las mayores cifras de contagio registradas, tal y como se observará en la parte posterior de este trabajo a través de la visualización de series temporales.

País Vasco

La primera serie temporal analizada será la comunidad autónoma de País Vasco, compuesta por las provincias de Álava, Guipúzcoa y Vizcaya.

Esta comunidad tuvo un número de 671.185 casos de contagio de COVID, y un número de 6.395 fallecidos.

La mayor parte de los contagios se dieron en la provincia de Vizcaya, consecuencia de que aquí se sitúa el núcleo poblacional de Bilbao, que representa la décima ciudad más grande de España, y la primera del norte peninsular.

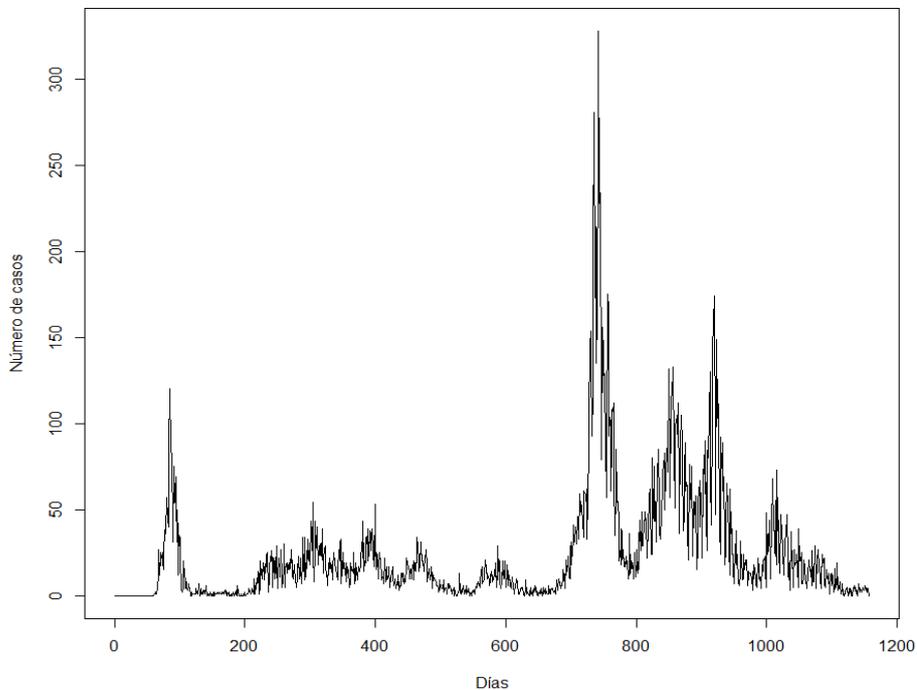


Figura 3.2 Serie temporal sobre el número de casos en País Vasco. Elaboración propia.

En esta serie temporal se destaca la tendencia en el número de casos diarios estacionaria, entre el rango de los 35 y 50 casos diarios durante todo el periodo 2020 – 2021. Sin embargo, si se observa el tramo final del año 2021, destaca el aumento de la incidencia provocada por la ola de contagio consecuencia de la variable Ómicron, que provocará una tendencia creciente en el número de casos diarios, superando los 300 casos diarios, consecuencia del valor atípico provocado por el crecimiento exponencial de casos a nivel nacional en la parte final del verano y el comienzo de otoño, tendencia que descenderá posteriormente.

Cataluña

La comunidad autónoma de Cataluña, compuesta por las provincias de Girona, Barcelona, Tarragona y Lleida. Posee 7.8 millones de habitantes, y por lo tanto la segunda comunidad autónoma más poblada de España.

Cabe destacar que Barcelona, capital de esta comunidad y segunda ciudad más grande de España, con 1 636 193 habitantes (2022), acumula estadísticamente la mayor parte de los casos de contagio en esta región. Tal y como se ha reflejado en la tabla posterior, Barcelona acumula el 75% de los casos positivos por SARS COVID-19, y el 84% de hospitalizados en mayores de 60 años de Cataluña.

El número total de casos en Cataluña alcanzó los 2.358.191 casos, con 18.773 fallecimientos reportados, convirtiéndose así en la comunidad autónoma con más casos de contagio por COVID-19 de todo el territorio español.

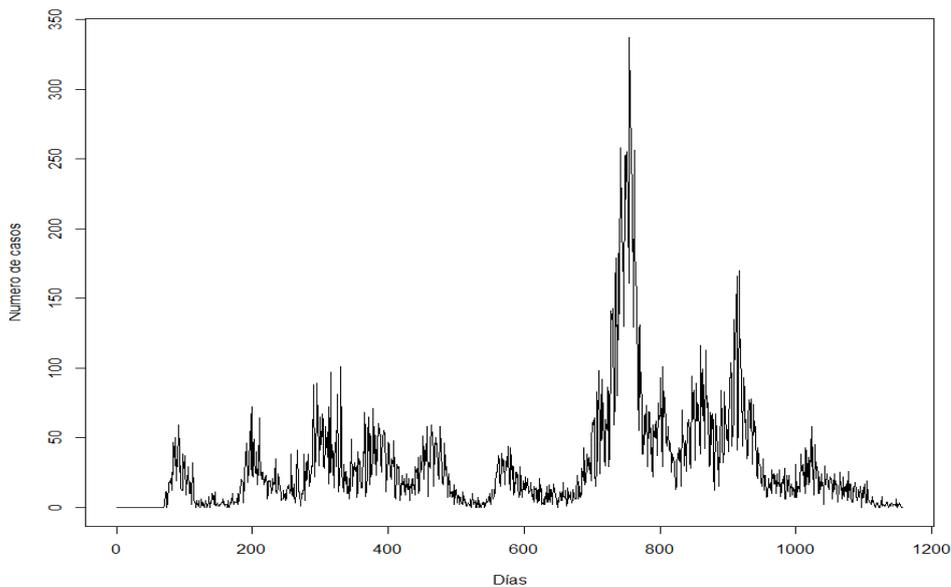


Figura 3.3 Serie temporal sobre el nº de casos en Cataluña. Elaboración propia.

El primer año, con el comienzo del contagio generalizado de COVID-19 por todo el territorio nacional supuso, lógicamente, una tendencia creciente en el número de casos de contagio en esta comunidad. Paradójicamente, y pese a las puntuales subidas, se puede observar una tendencia en el número de casos de mayores de 60 años estacionaria hasta finales de 2021 y comienzos de 2022.

Sería en el comienzo de 2022, cuando se produciría un aumento explosivo en el número de casos, sobrepasando los 300 positivos diarios en esta comunidad; rápidamente del mismo modo que la tendencia ascendió alcanzando máximos de contagio, descendería posteriormente.

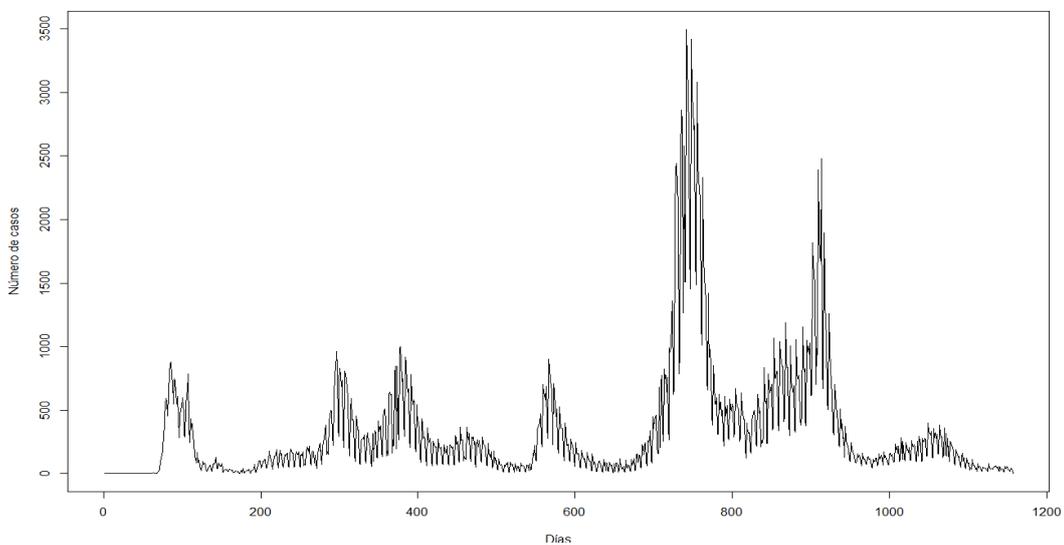


Figura 3.4 Serie temporal sobre el nº casos en la provincia de Barcelona. Elaboración propia.

Se puede apreciar similitud entre las concentraciones en el número de casos a lo largo de la serie temporal entre la comunidad autónoma de Cataluña y la provincia de Barcelona, consecuencia de cómo influye la concentración de casos en la capital sobre las cifras de la comunidad autónoma.

Comunidad Valenciana

La comunidad autónoma de la Comunidad Valenciana, compuesta por las provincias de Castellón, Valencia y Alicante; tanto en la provincia de Valencia, cuya capital de esta comunidad autónoma da nombre, como en la provincia de Alicante es donde se concentrará la mayor parte de los casos positivos de contagio por COVID-19, en el caso de los mayores de 60 años, el 87% de los casos positivos diarios declarados se dieron en estas dos provincias de entre toda la Comunidad Valenciana.

El número total de positivos por contagio ascendió a 1.348.138 casos, con una cifra de 9.135 fallecimientos.

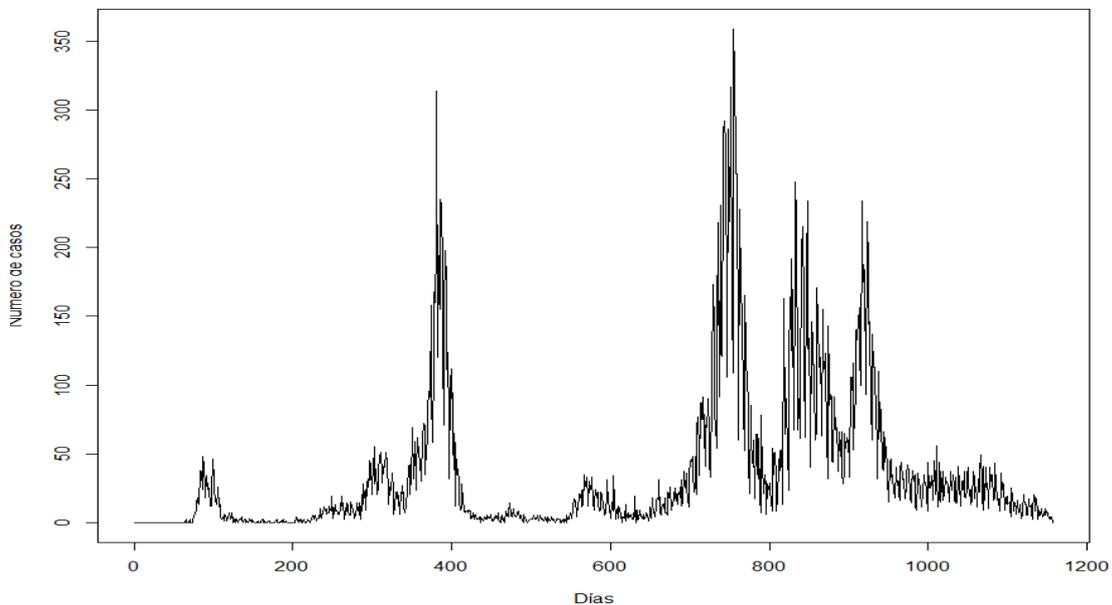


Figura 3.5 Serie temporal sobre el nº casos en Comunidad Valenciana. Elaboración propia.

La Comunidad Valenciana posee una distribución única tanto del número de casos tal y como se observa a simple vista en la serie temporal elaborada. En primer lugar, con la primera oleada el número de contagiados es reducido, inferior a los 50 casos positivos por coronavirus, para descender y mantener una tendencia estacionaria tras la implantación del Estado de Alarma a nivel nacional. A finales de 2020 y comienzos de 2021 se produce una explosión de casos de contagio, que descenderá posteriormente de manera significativa, manteniendo una tendencia estacionaria durante todo el 2021, hasta con la llegada de la variante Ómicron, producirse de nuevo un nuevo máximo en el número de casos.

Comunidad de Madrid

La comunidad autónoma de Madrid, compuesta por la provincia de Madrid, cuya ciudad es la capital de España. Situada en el centro, es la ciudad más grande del territorio español, con más de 3 millones de habitantes, y una población de 6.825.000 personas distribuidas por esta comunidad.

Esta ciudad representa el centro financiero y económico del país, además de una importante industria del turismo. 1.636.314 casos positivos por COVID-19 fueron declarados en esta comunidad, y un número de defunciones que alcanzó los 17.952 fallecimientos.

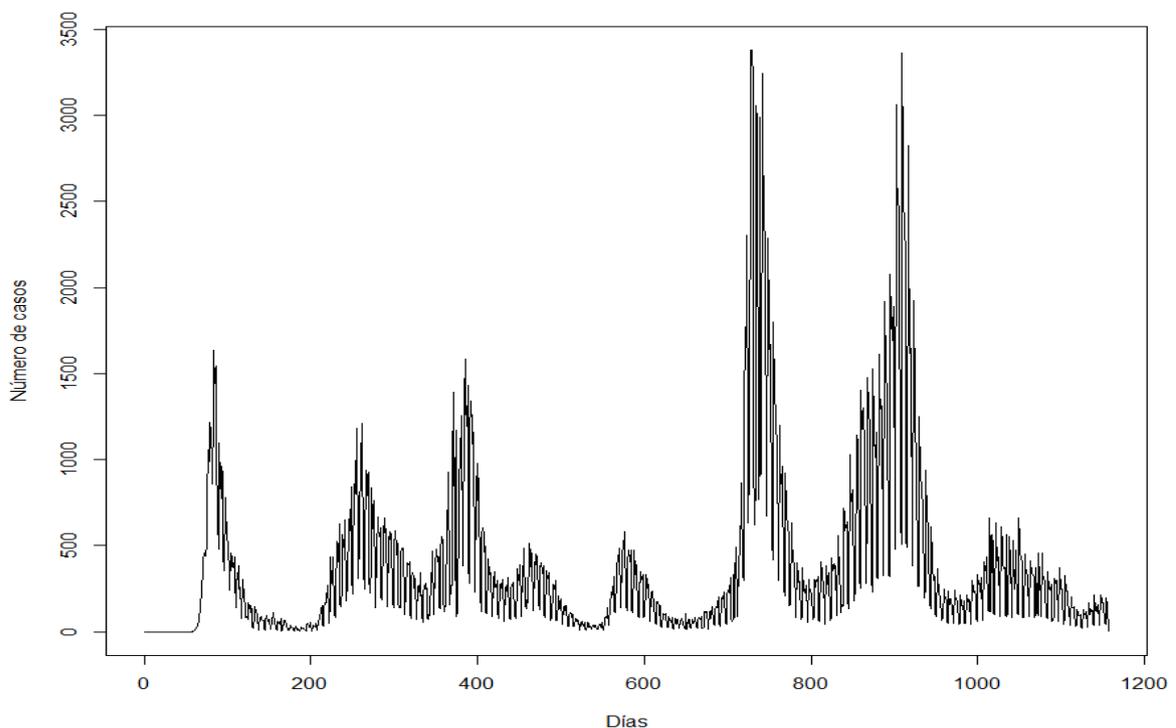


Figura 3.6 Serie temporal sobre el nº casos en la Comunidad de Madrid. Elaboración propia.

La Comunidad de Madrid ha presenciado el pico de mayor número de casos de contagio de todo el territorio nacional, con una visiblemente estacionalidad con respecto a la tendencia del número de casos.

4 de las 6 olas de contagio ocurridas a lo largo de los años incluidos en esta serie temporal, han supuesto un máximo en el número de casos positivos, sobrepasando los 1500 casos diarios durante los dos primeros años, y alcanzando cifras de más de 3000 contagios positivos por COVID-19 durante el final del 2021, y posteriormente en 2022.

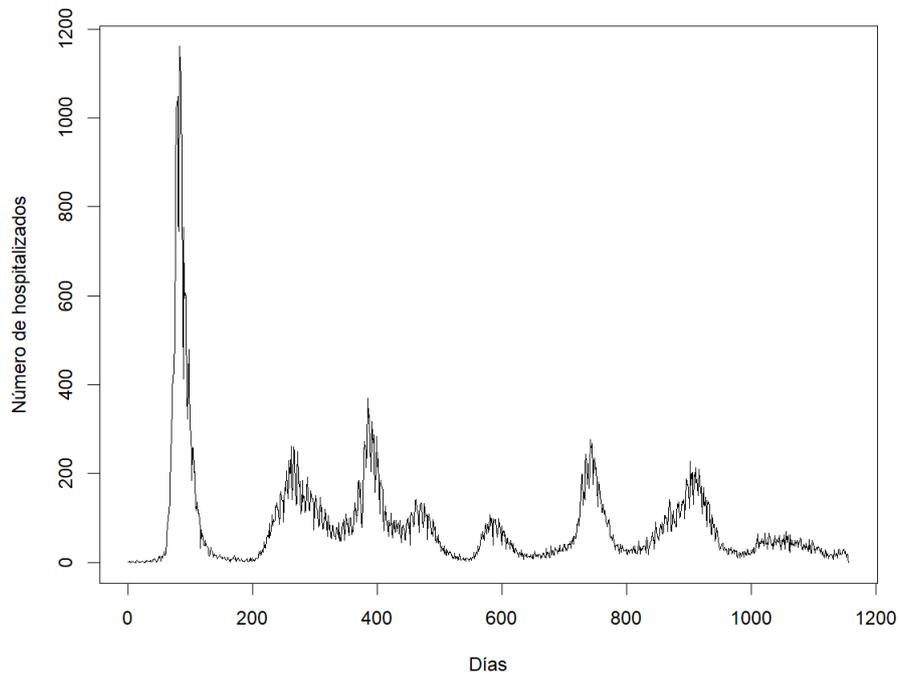


Figura 3.7 Serie temporal sobre el nº de hospitalizados en la C. de Madrid. Elaboración propia.

La Comunidad de Madrid también ha tenido el mayor número de hospitalizados. Cabe destacar que la mayor parte de los pacientes hospitalizados por COVID-19 se dieron en el comienzo de la primera ola de contagio, sobrepasando los 1000 hospitalizados diarios en mayores de 60 años, cifra tan elevada que no se alcanzará en el resto de los años posteriores ninguna cifra que alcance la mitad de esta.

Castilla-La Mancha

La comunidad autónoma de Castilla-La Mancha, compuesta por las provincias de Toledo, Guadalajara, Cuenca, Ciudad Real y Albacete. Castilla-La Mancha posee una población de 2,4 millones de habitantes.

La mayor concentración poblacional se encuentra en las 5 principales ciudades homónimas. En el norte, están las dos ciudades próximas a Madrid, Guadalajara y Toledo, ambas superando los 80.000 habitantes; en el sur la mayor parte de la población se sitúa alrededor de la ciudad de Albacete, con más de 170.000 habitantes, siendo la ciudad más poblada de toda Castilla-La Mancha. Esta comunidad declaró 486.970 casos positivos por COVID-19, y 7.137 fallecimientos.

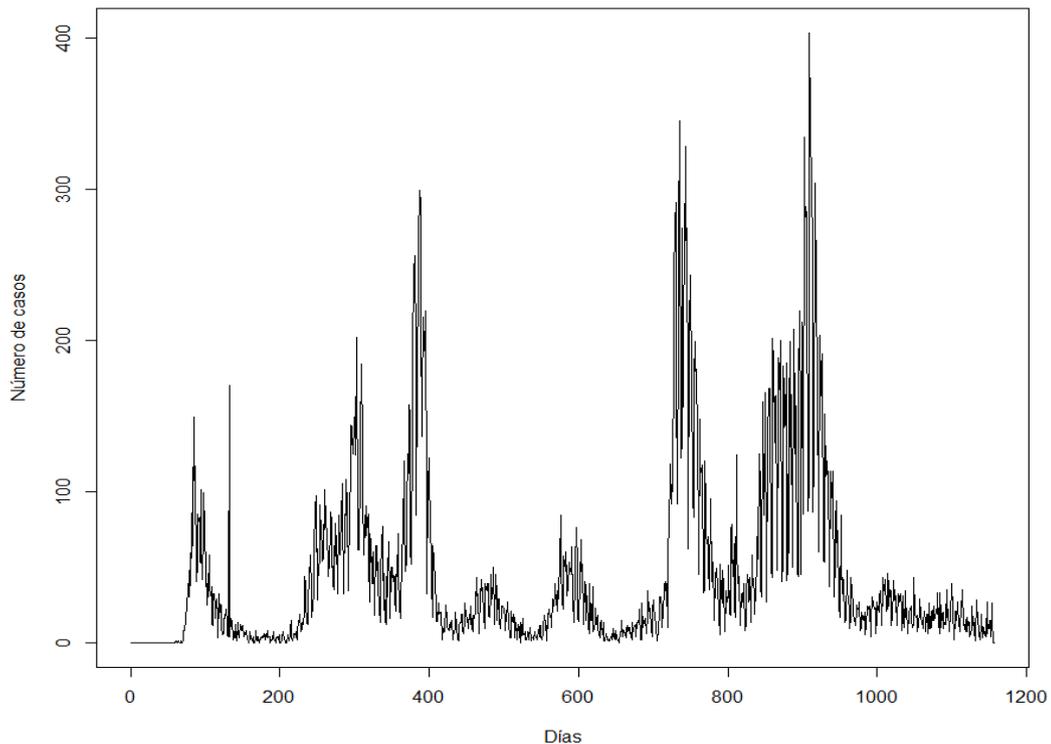


Figura 3.8 Serie temporal sobre el nº casos en Castilla-La Mancha. Elaboración propia.

Castilla-La Mancha posee una distribución del número de casos positivos por SARS COVID-19 bastante distinta y peculiar a todas las comunidades autónomas.

En esta comunidad se alcanzan máximos de manera consecutiva durante todo el año 2020, un suceso que no se repite en otras comunidades autónomas. Del mismo modo que se produce en las comunidades autónomas del interior peninsular como Castilla y León y Extremadura, el número de casos es inferior a otras comunidades con una elevada

actividad del sector servicios, tales como la Comunidad de Madrid o las Islas Baleares. Durante la mayor parte de 2021 mantendrá una tendencia estacionaria adquiriendo un crecimiento explosivo en el tramo final de este último año.

Islas Canarias

La última comunidad autónoma de este trabajo es la comunidad autónoma de las Islas Canarias, que posee 7 islas mayores divididas administrativamente en dos provincias, Las Palmas de Gran Canaria y Santa Cruz de Tenerife.

Las Islas Canarias poseen una población de alrededor de 2,2 millones de habitantes, y una superficie total de 7.500 km². Su economía está basada en la agricultura con producción de alimentos propios de climas tropicales, y principalmente una gran importancia del sector servicios, consecuencia de la gran importancia del turismo en esta comunidad autónoma. Los casos positivos por contagio de COVID-19 declarados ascendieron a 332.186 casos, y 1.634 fallecimientos.

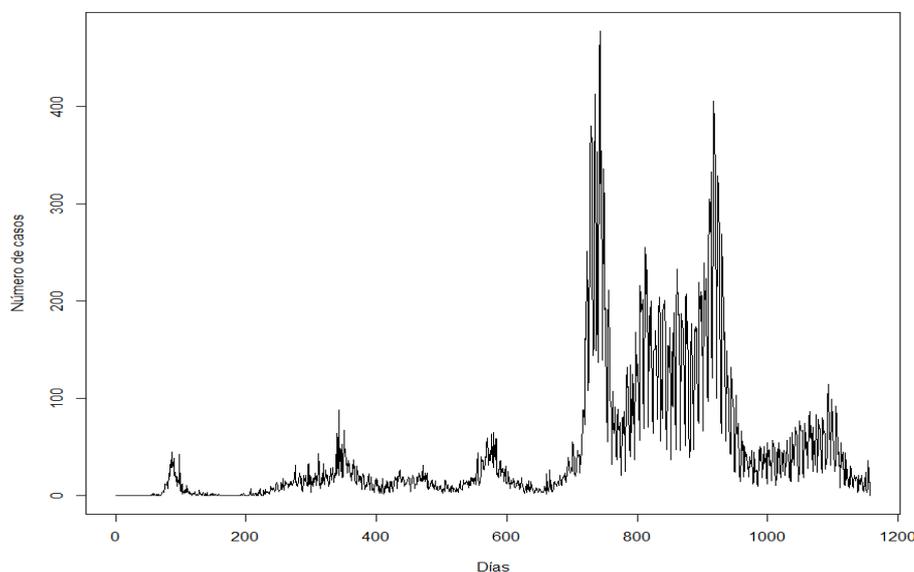


Figura 3.9 Serie temporal sobre el nº de casos en Islas Canarias. Elaboración propia.

Exclusivamente en esta comunidad autónoma, la distribución del COVID-19 en personas mayores de 60 años presentará una tendencia estacionaria durante todo el periodo en los años 2020 y 2021, con ligeros despuntes de manera muy puntual. Se observa como la tendencia no será creciente hasta el final de 2021 y comienzo del año 2022, donde se alcanzarán cifras máximas de contagio sobrepasando los 400 casos positivos por coronavirus diarios.

4. Análisis y resultados

A continuación, se expondrá y analizará de qué manera se expandió el SARS COVID-19 por todo el territorio nacional con respecto al número de contagios en mayores de 60 años. Con la finalidad de poder observar cuál fue la distribución del contagio, de qué manera ocurrió y qué zonas se vieron más afectadas, se llevará a cabo un análisis más minucioso del contagio a nivel provincial, pudiendo explicar posibles características o rasgos haciendo uso de series temporales y modelos econométricos que nos ayuden en el análisis final de este trabajo.

4.1 Análisis de la expansión del COVID-19 por Comunidades Autónomas

En primer lugar, se puede observar cómo los casos de contagio se concentran en las grandes urbes españolas. Pese a que un mayor número de habitantes puede provocar un aumento en la transmisión de este virus, no podemos hablar exclusivamente de una correlación entre población y número de casos en las principales provincias españolas, pues no se cumple que las 10 provincias con mayor número de contagios sean las más pobladas.

La provincia de Asturias, que ocupa el 5º lugar en número de casos positivos por coronavirus, representa la 14ª provincia más poblada de España. Del mismo modo ocurre con la ciudad de Zaragoza o Toledo, que se sitúan entre las 10 provincias más afectadas, y sin embargo ocupan la provincia 15ª y 23ª en número de habitantes.

Como se ha especificado al comienzo de este trabajo, el COVID-19 posee una distribución heterogénea, lo que supone un reto para poder describir los posibles patrones de contagio de este virus.

Madrid y Barcelona, con más de 12 millones de habitantes (INE, 2022) suponen más del 25% de la población española, y son las dos provincias con mayor número de contagiados diarios, sin embargo, la distribución del contagio no ha sido homogénea entre ambas provincias. Madrid ha presentado un número de casos diarios relativamente similar al de la provincia de Barcelona, pese a que la primera casi triplica el número de hospitalizados a la segunda. Debido al gran tamaño de ambas ciudades, capitales de sus respectivas comunidades autónomas, se refleja como el comportamiento del COVID-19 en estas dos provincias influencia las cifras de contagio de ambas comunidades.

Podemos observar como la distribución del número de casos positivos ha sido relativamente similar en ambas ciudades, pese a que ha habido una tendencia más estacionaria en la provincia de Barcelona durante los dos primeros años de la línea temporal, a diferencia de la provincia de Madrid que ha presentado cifras mayores en el número de casos diarios en cada una de las oleadas de contagio que azotaron al territorio nacional durante los años 2020 y 2023.

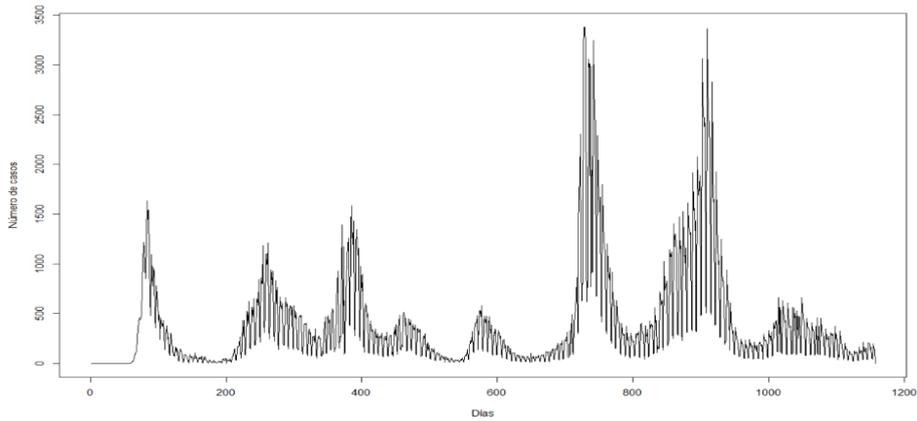


Figura 4.1 Serie temporal de la provincia de Madrid. Elaboración propia.

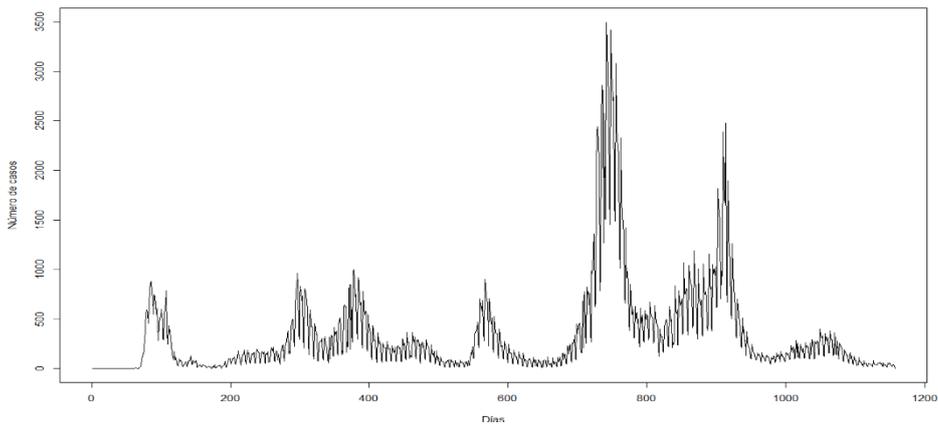


Figura 4.2 Serie temporal de la provincia de Barcelona. Elaboración propia.

```

Series: TSCataluna
ARIMA(2,1,4)

Coefficients:
      ar1      ar2      ma1      ma2      ma3      ma4
s.e.  1.2467 -0.8256 -1.7648  1.1682 -0.0261 -0.1519
      0.0303  0.0833  0.0522  0.1747  0.1512  0.0465

sigma^2 = 263.7: log likelihood = -4860.45
AIC=9734.9  AICc=9735  BIC=9770.27

```

Figura 4.3 Modelo ARIMA de la comunidad autónoma de Cataluña. Elaboración propia.

```

Series: TSMadrid
ARIMA(5,1,5)

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3      ma4      ma5
s.e.  0.2534 -0.9532 -0.0064 -0.4850 -0.5226 -0.9515  1.1085 -0.7010  0.4314  0.0629
      0.0501  0.0414  0.0678  0.0391  0.0444  0.0542  0.0702  0.0784  0.0625  0.0440

sigma^2 = 39960: log likelihood = -7762.52
AIC=15547.04  AICc=15547.27  BIC=15602.62

```

Figura 4.4 Modelo ARIMA de la Comunidad de Madrid. Elaboración propia.

Del mismo modo que se ha explicado previamente en numerosas ocasiones dentro del contenido de este trabajo, como consecuencia de la transmisión heterogénea del COVID-19 se pueden observar patrones y características únicas respecto a la distribución y contagio del mismo.

La Comunidad de Madrid ha acumulado los mayores picos de contagio de todo el país durante la crisis sanitaria, con cifras que sobrepasan los 3000 casos de mayores de 60 años al día; sólo en Cataluña, y como consecuencia de la ciudad de Barcelona, se alcanzarían cifras similares de contagio. Además, la Comunidad de Madrid presenta uno de los modelos ARIMA más complejos de los incluidos en este trabajo, un gran número de factores son los que influyen en el contagio del COVID-19 dentro de esta comunidad autónoma. De distinta forma está compuesto el modelo de la Comunidad autónoma de Cataluña, de menor complejidad o mayor simplicidad, tal y como se puede observar a simple vista en las imágenes de ambos modelos.

País Vasco presenta un modelo más sencillo que el de la Comunidad de Madrid, aunque incluye un componente o coeficiente adicional al de Cataluña. Si se observa el Criterio de Información de Akaike de las dos comunidades mencionadas previamente junto a País Vasco, es este último quien posee una cifra inferior de acuerdo al Criterio de Akaike.

```
Series: TSPaisVasco
ARIMA(5,1,2)

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ma1      ma2
    0.1423  -0.6615  -0.3240  -0.2664  -0.4386  -0.6164  0.6136
s.e.  0.0498   0.0305   0.0324   0.0256   0.0328   0.0563   0.0293

sigma^2 = 173.1:  log likelihood = -4616.8
AIC=9249.59  AICc=9249.72  BIC=9290.01
```

Figura 4.5 Modelo ARIMA de la comunidad autónoma de País Vasco. Elaboración propia.

En el caso de Castilla-La Mancha, durante el primer año se alcanzan máximos de contagio sucesivos, este patrón de contagio se dará sólo en esta comunidad autónoma, que presenta un modelo ARIMA sencillo, de 5 componentes o coeficientes que simbolizan una reducida complejidad en referencia a los factores que influyen en la transmisión del COVID-19 en esta comunidad autónoma.

```
Series: TSCLM
ARIMA(2,1,3)

Coefficients:
      ar1      ar2      ma1      ma2      ma3
    1.2442  -0.8951  -1.8483  1.4434  -0.3678
s.e.  0.0175   0.0204   0.0309  0.0510   0.0329

sigma^2 = 773.3:  log likelihood = -5483.01
AIC=10978.02  AICc=10978.09  BIC=11008.33
```

Figura 4.6 Modelo ARIMA de Castilla-La Mancha. Elaboración propia.

Aunque con cifras mucho más inferiores en el número de contagiados, Islas Canarias presenta un modelo de elevada complejidad, con parámetros (5,1,5) y hasta 10 coeficientes en el propio modelo.

```

Series: TSCanarias
ARIMA(5,1,5)

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3      ma4      ma5
0.3009 -1.0212  0.0811 -0.5716 -0.4564 -0.8699  1.1458 -0.7405  0.6203 -0.0779
s.e.  0.0529  0.0441  0.0742  0.0423  0.0481  0.0548  0.0656  0.0807  0.0615  0.0429

sigma^2 = 545:  log likelihood = -5279.52
AIC=10581.05  AICc=10581.28  BIC=10636.63

```

Figura 4.7 Modelo ARIMA de Islas Canarias. Elaboración propia.

Por último, y de similar complejidad a los modelos previos, la Comunidad Valenciana presenta 9 coeficientes en su modelo ARIMA, con un elevado Criterio de Información de Akaike en comparación al resto de modelos incluidos en este apartado.

```

Series: TSComValenciana
ARIMA(5,1,4)

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3      ma4
0.3877 -1.0309  0.1671 -0.5786 -0.3789 -0.9488  1.0900 -0.6311  0.5333
s.e.  0.0509  0.0380  0.0623  0.0300  0.0401  0.0514  0.0543  0.0524  0.0382

sigma^2 = 361.7:  log likelihood = -5042.58
AIC=10105.17  AICc=10105.36  BIC=10155.7

```

Figura 4.8 Modelo ARIMA de la Comunidad Valenciana. Elaboración propia.

Como se ha comentado al principio de este trabajo, los modelos más complejos de los incluidos en este trabajo hacen referencia a la Comunidad de Madrid y a Islas Canarias. De acuerdo al elevado número de componentes, se puede observar la complejidad del modelo, además del inferior AIC de la comunidad autónoma de Islas Canarias comparado al AIC de la Comunidad de Madrid.

El modelo con el número más reducido de componentes o coeficientes en su modelo ARIMA es el de la comunidad autónoma de Castilla-La Mancha, que representa el modelo más sencillo de los expuestos en este trabajo. Por otro lado, el Criterio de Información de Akaike inferior de los modelos anteriores es el perteneciente a la comunidad autónoma de País Vasco; a consecuencia de ello el modelo de ésta comunidad es el más idóneo en términos de ajuste y simplicidad.

4.2 Impacto económico del COVID-19 en España

Los efectos económicos del COVID-19 supusieron un grave receso en el segundo trimestre de 2020, con una caída del PIB de casi el 12%. La contracción en España ha sido más pronunciada debido, entre otros, al mayor peso del sector servicios, principalmente hostelería, y como consecuencia del peso de las PYMES en el tejido empresarial español.

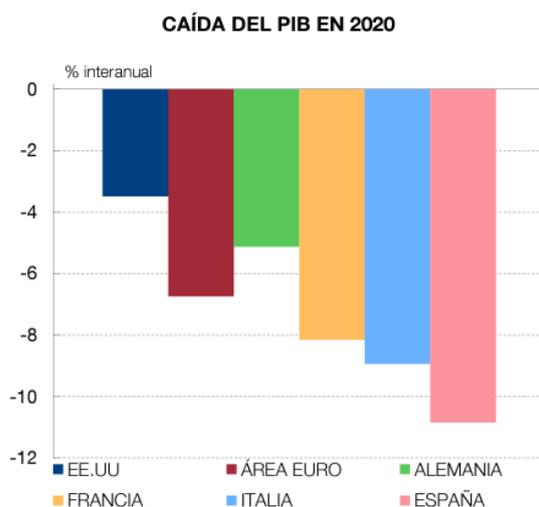


Figura 4.9 Descenso del PIB en el año 2020. (Fuente: Banco de España, 2021)

Las comunidades autónomas que se han visto más afectadas son aquellas con un gran peso del turismo y hostelería. Comunidades como las Islas Baleares y Canarias, con una elevada relevancia turística se han visto gravemente afectadas con caídas superiores al 14% en su producto interior bruto, como consecuencia de la restricción en la movilidad tanto nacional como internacional.

Por otro lado, las comunidades autónomas del centro español, mucho más rurales y con gran peso del sector primario tales como Extremadura o Castilla-La Mancha se han visto menos afectadas económicamente con caídas inferiores al 8%.

ESTIMACIÓN DE LA CAÍDA DEL PIB EN 2020 POR PROVINCIAS

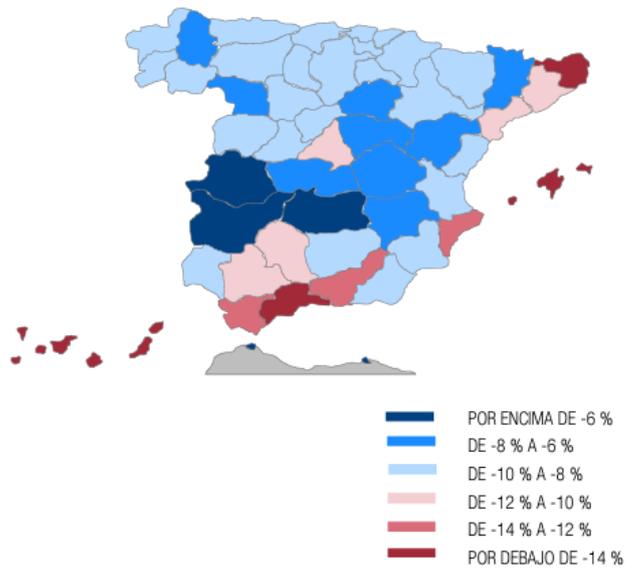


Figura 4.10 Caída del PIB en el año 2020 por provincias. (Fuente: Banco de España, 2021).

Se podrá observar como la costa mediterránea representa, junto a Madrid, las provincias y comunidades autónomas más afectadas como consecuencia de la contracción en el número de turistas que visitan el país, concretamente Cataluña, las Islas Baleares y Canarias, comunidades que registraron el mayor número de turistas en el año 2021 (INE, 2022); Baleares acumuló en 2021 el 22,8% de todos los ingresos turísticos del país, seguida de Canarias (19%), Cataluña (15,8%), Andalucía (13,8%) y la Comunidad Valenciana (12,6%).

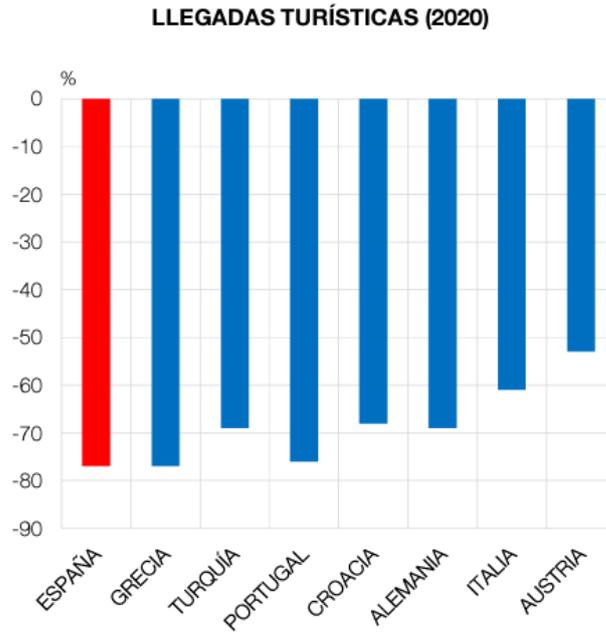


Figura 4.11 Caída del turismo porcentualmente en el año 2020. (Fuente: Banco de España, 2021)

Del mismo modo que se ha explicado previamente, el sector más afectado es el sector servicios, siendo la hostelería y la restauración los sectores más afectados como consecuencia de las restricciones en la movilidad intercomunitaria en España, y la limitación en el número de turistas que visita el país.

Paradójicamente, se produciría una grave contracción del consumo de combustible fósiles con la imposibilidad del desplazamiento tal y como se ha especificado previamente. Ambos sectores presentan una elevada correlación entre sí, pues son los periodos y temporadas con mayor afluencia turística cuando se lleva a cabo un mayor consumo de combustibles fósiles como consecuencia del desplazamiento entre las distintas comunidades autónomas.

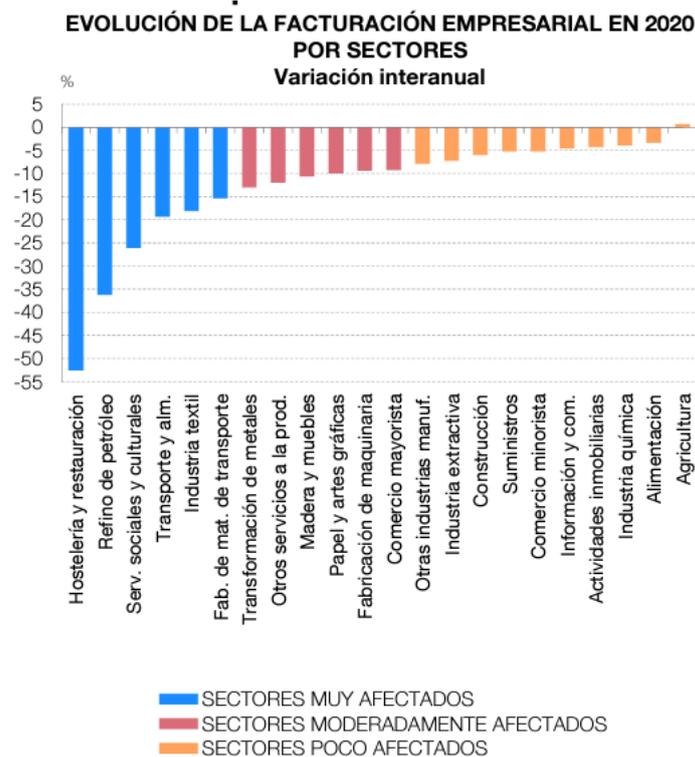


Figura 4.12 Evolución de la facturación por sectores en 2020 (Fuente: BDE, 2021)

De diferente manera afectaría al sector primario, característico de las provincias del interior español, cuya actividad no se vería afectada significativamente como consecuencia de la propia naturaleza de esta actividad económica.

En conclusión, las comunidades más afectadas han sido aquellas con mayor presencia del sector turístico y restauración, a saber, Canarias, Baleares, Cataluña y la Comunidad de Madrid; descendiendo el producto interior bruto de estas comunidades autónomas en más de un 14% para ambas islas, y entre un 10% y 12% en Cataluña y Madrid, siendo éstas las dos comunidades autónomas con mayor número de casos positivos por COVID-19.

Con una caída de casi el 12% del PIB en el primer semestre de 2020, España representa uno de los países europeos más afectados de la con un descenso del PIB de aproximadamente el doble que la media de los países de la Zona Euro. La intensidad de las medidas de restricción de movilidad implantadas desde las Administraciones Públicas, ha favorecido la reducción y el aplanamiento de la curva de contagios en todas y cada una de las comunidades autónomas; con la implantación del Estado de Alarma en España el 13 de marzo de 2020 se frenó el contagio generalizado de COVID-19 en todo el territorio nacional, sacrificando la actividad económica que padeció una grave contracción del consumo generalizado.

5. Conclusiones

A lo largo de este Trabajo de Fin de Grado se ha estudiado el SARS COVID-19 y su impacto en personas mayores de 60 años. Este virus ha provocado una pandemia mundial alterando las vidas y las economías de todos los países del mundo. A través de las distintas olas de contagio acontecidas entre los años 2020 y 2022, el coronavirus ha afectado a más de 13 millones de españoles, y provocado más de 100 mil fallecidos.

Este trabajo ha consistido, en primer lugar, en explicar los rasgos, características y la breve historia conocida hasta ahora sobre el COVID-19. Desde su aparición el 31 de diciembre de 2019 en la ciudad china de Wuhan, alrededor de 6.3 millones de personas (junio, 2022) han fallecido a nivel mundial como consecuencia de este virus. Se han comentado su distribución heterogénea en todo el territorio nacional, y se ha explicado de qué manera se ha combatido desde las instituciones políticas y desde el propio sistema sanitario español.

En las primeras oleadas de contagio, el SARS COVID-19 se caracterizaba por expandirse rápidamente, provocando el fallecimiento de miles de personas, especialmente en aquellas que padecían de otras enfermedades cardio-respiratorias.

A comienzos de 2022 la variante Ómicron dispararía de manera explosiva el número de contagios por todo el territorio; esta variante tenía la particularidad de distribuirse y expandirse rápidamente entre individuos, pero su letalidad era reducida con respecto a las variantes previas. Como consecuencia de la baja letalidad de Ómicron, y los grandes esfuerzos del sistema sanitario español por inmunizar a la población española con varias campañas generalizadas de vacunación, entre otros, lo que provocó que la tendencia en el número de fallecidos descendiera bruscamente desde las más de 90.000 defunciones ocurridas entre el periodo 2020 – 2021, a las menos de 40.000 ocurridas entre 2022 y 2023.

Los modelos elaborados en este trabajo representan una manera adecuada estadísticamente de analizar los datos. Es a través de estos modelos econométricos que se puede observar si los factores y hechos causales que provocan los sucesos tienen un impacto en las comunidades autónomas y la expansión del COVID-19 en mayores de 60 años, causas que se analizan en este trabajo. Se recomienda realizar otro análisis donde se profundice en los hechos y características de estos factores, cómo afectan a los modelos de las series temporales, cual es su origen y su importancia relativa en cada una de las comunidades autónomas españolas.

Con bases de datos adicionales que analicen distintos factores que afecten la transmisión del COVID-19, tales como cifras turísticas o desplazamientos entre comunidades, afluencia de actividad económica e importancia relativa de los sectores con mayor peso en cada provincia, se podría realizar un estudio más incidente en las causas de la transmisión del coronavirus dentro de un país, así como la correlación que presentan estas variables entre sí. Especialmente importante serían variables que tengan que ver con los sectores que más influyen en la transmisión del COVID-19 como el sector turístico, y más concretamente hostelería y restauración, sectores que conllevan una actividad social inherente.

Este trabajo representa una manera útil de poder analizar el impacto y expansión de un virus altamente contagioso como el que se distribuyó de manera progresiva a comienzos del año 2020. Sirva este documento como un método econométrico de medir la evolución de las distintas áreas y regiones españolas, así como su impacto económico en los sectores que componen el PIB español.

Tanto las series temporales como los modelos ARIMA presentados en este trabajo, permiten predecir en base a información previa recogida en una base de datos, de qué manera aproximada podría expandirse enfermedades patógenas por todo el territorio nacional. Estos modelos permiten visualizar y monitorizar la evolución que pueda producirse en las distintas provincias, qué factores influyen en el contagio, y qué variables pueden tener una mayor correlación entre sí, de esta manera se podría tratar más adecuadamente una crisis sanitaria futura analizando cuáles son los factores que influyen significativamente en la expansión de las enfermedades o las próximas pandemias que puedan suceder, así como de qué depende su mayor letalidad o incidencia.

6. Referencias y bibliografía

Orea, L., & Álvarez, I. C. (2022). How effective has the Spanish lockdown been to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces. *Health Economics (United Kingdom)*, 31(1), 154–173.

Gutiérrez, M. J., Inguanzo, B., & Orbe, S. (2021). Distributional impact of COVID-19: regional inequalities in cases and deaths in Spain during the first wave. *Applied Economics*, 53(31), 3636–3657. <https://doi.org/10.1080/00036846.2021.1884838>

Ponce-de-Leon, M., del Valle, J., Fernandez, J. M., Bernardo, M., Cirillo, D., Sanchez-Valle, J., Smith, M., Capella-Gutierrez, S., Gullón, T., & Valencia, A. (2021). COVID-19 Flow-Maps an open geographic information system on COVID-19 and human mobility for Spain. *Scientific Data*, 8(1). <https://doi.org/10.1038/s41597-021-01093-5>

López-Mendoza, H., Montañés, A., & Moliner-Lahoz, F. J. (2021). Disparities in the evolution of the covid-19 pandemic between spanish provinces. *International Journal of Environmental Research and Public Health*, 18(10). <https://doi.org/10.3390/ijerph18105085>

Pollán, M.; Pérez-Gómez, B.; Pastor-Barriuso, R.; Oteo, J.; Hernán, M.A.; Pérez-Olmeda, M.; Sanmartín, J.L.; Fernández-García, A.; Cruz, I.; Fernández de Larrea, N.; et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. *Lancet* 2020, 396, 535–544.

Mentzakis, E., M. García-Goñi, A. R. Sequeira, and F. Paolucci. 2019. “Equity and Efficiency Priorities within the Spanish Health System: A Discrete Choice Experiment Eliciting Stakeholders Preferences.” *Health Policy and Technology* 8 (1): 30–41.

Legido-Quigley, H., J. T. Mateos-García, V. R. Campos, M. Gea-Sánchez, C. Muntaner, and M. McKee. 2020. “The Resilience of the Spanish Health System against the COVID-19 Pandemic.” *The Lancet Public Health* 5 (5): e251– e252.

Raúl Sánchez, A. O. (2022, March 29). *El coronavirus en España: Mapas y últimos Datos de los Casos y su evolución*. elDiario.es https://www.eldiario.es/sociedad/mapa-datos-coronavirus-espana-comunidades-autonomas-marzo-29_1_1039633.html

Instituto Nacional de Estadística - <https://www.ine.es/>

CARLOS E. CUÉ. El Gobierno informa de que es la única autoridad en toda España, limita los desplazamientos y cierra comercios. *El País*. <https://elpais.com/espana/2020-03-14/el-gobierno-prohibe-todos-los-viajes-que-no-sean-de-fuerza-mayor.html>

El Periódico. (2020). <https://www.elperiodico.com/es/economia/20200317/medidas-economicas-coronavirus-gobierno-7893836>

Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeen F (2023). *_forecast: Forecasting functions for time series and linear models_*. R package version 8.21, <<https://pkg.robjhyndman.com/forecast/>>.

Hyndman RJ, Khandakar Y (2008). "Automatic time series forecasting: the forecast package for R." *_Journal of Statistical Software_*, *26*(3), 1-22. doi:10.18637/jss.v027.i03 <<https://doi.org/10.18637/jss.v027.i03>>.

Boscá, J. E., Doménech, R., Ferri, J., & Ulloa, C. (2020). <https://www.bbvaesearch.com/publicaciones/espana-la-situacion-ciclica-de-la-economia-espanola-durante-la-crisis-del-covid-19/>

<https://www.bde.es/f/webbde/GAP/Secciones/SalaPrensa/IntervencionesPublicas/DirectoresGenerales/economia/Arc/Fic/arce180221.pdf>

Box, Jenkins and Reinsel; *Time Series Analysis: Forecasting and Control*; John Wiley & SONS.; 4th edition(Jun 30, 2008), ISBN: 470272848

Walter Enders; *Applied Econometric Time Series*; Wiley; 4th edition(Nov 03, 2014), ISBN: 1118808568

Martínez, Diego & Albín, Julio & Cabaleiro, José & Pena, Tomas & Rivera, Francisco & Blanco, Vicente. (2009). El criterio de información de Akaike en la obtención de modelos estadísticos de Rendimiento. 439-444.

7. Anexos

Anexo 1 – Series temporales de todas las Comunidades Autónomas

Anexo 2 – Modelos ARIMA de las Comunidades Autónomas.

Anexo 3 – Líneas de código del lenguaje de programación estadístico R.