**[EAE - Business School](http://www.eae.es/)**

**Master en finance data science**

**Trabajo Final de Master**

“Crisis sanitaria y económica: Cómo ha afectado el COVID a las economías europeas y cómo minimizar el impacto de posibles futuras pandemias”

Autores:

María Cruz Fernández, (maria.cf6@gmail.com)

José Miguel González Domínguez, (jose.miguelgonzalezdominguez@gmail.com)

Vanessa López García dirección de email

Cristina Madrid Crespo, dirección de email

(Grupo 4)

Tutor:

Prof. D. Serge Chávez Feria, dirección de email

Madrid, 2021

**Índice General**

**Índices de figuras** …………………………………………………………………… p. X.

**[times new roman, 12, versálitas, neg.]**

**Abreviaturas** ………………………………………………………………………… p. X.

**Resumen/Abstract** ………………………………………………………………….. p. X.

**I.- Parte Introductoria** …………………………………………………………… p. X.

1.- Interés del estudio …

2.- Fines y objetivos …

3.- Metodología y estado de la cuestión …

**II.- Parte General** …

4.- Análisis estratégico; externo e interno …

5.- Análisis DAFO…

5.1.- X …

5.2.- X …

**III.- Parte Especial** …

6.- Objetivos del plan de marketing

7.- Estrategias operativas del plan

7.1.- X …

7.2.- X …

8.- Plan de acción/Tácticas

(…)

**IV.- Conclusiones** …

**V.- Anexos** …

14.- Fuentes de consulta …

14.1.- Bibliografía …

14.2.- Recursos electrónicos …

14.3.- Otros Recursos …

15.- Regulación …

15.1.- Regulación imperativa …

15.2.- Regulación dispositiva …

16.- Documentación de trabajo

**Índices de figuras**

Figura 1.- [Rótulo] …

Figura 2.- [Rótulo] …

(…)

**Abreviaturas**

DAFO: Debilidades, Amenazas, Fortalezas y Oportunidades.

PMK: Plan de marketing.

(…)

**Resumen**

El trabajo que tenéis entre manos no es sino un análisis sosegado de una de las crisis más importantes a la que nos hemos enfrentado como sociedad, una crisis sanitaria de efectos devastadores en el plano humano que ha traído también consigo importantes secuelas negativas para la economía, algunas inmediatas y palpables y otras que observaremos a largo plazo. Con el fin de determinar el impacto de esta pandemia en los países que componen la Unión Europea – y así poder predecir o al menos enfrentarnos con más garantías a un posible suceso de igual magnitud – nuestro objetivo es el de analizar minuciosamente los datos de fuentes públicas y privadas para obtener las correlaciones que nos den una imagen fiel de la relación entre variables sanitarias y económicas, es decir, qué efectos ha tenido la pandemia en sectores tan importantes como el del turismo, la automoción o las ventas en línea, por ejemplo, comparando los resultados de los distintos países que conforman la Unión Europea, y todo ello utilizando herramientas de análisis de datos y machine learning para alcanzar el resultado deseado, que no es otro que el de reflejar la realidad de una manera veraz, científica y fidedigna.

**Palabras clave:** correlaciones; análisis de datos; machine learning

**Abstract**

The work you have in hand is nothing but a safe analysis of one of the most important crises that we have faced as a society, a health crisis with devastating effects on the human level that has also led to significant negative consequences for the economy. some immediate and palpable and others that we will observe in the long term. In order to determine the impact of this pandemic in the countries that make up the European Union - and thus to be able to predict or at least confront with more guarantees a possible event of the same magnitude - our objective is to carefully analyze data from public and private sources to obtain the correlations that give us a true picture of the relationship between health and economic variables, that is, what effects the pandemic has had on such important sectors as tourism, automotive or online sales, for example, comparing the results of the different countries that make up the European Union, and all this using data analysis tools and machine learning to achieve the desired result, which is none other than reflecting reality in a truthful, scientific and reliable way.

**Key-words:** correlations; data analysis; machine learning

**I.- Parte Introductoria**

**1.- Interés del estudio**

No cabe duda de que la crisis sanitaria a la que nos enfrentamos ha sido uno de los mayores retos que hemos vivido como ciudadanos y así mismo ha supuesto y supone un reto diario para las instituciones que nos gobiernan. No es descabellado decir que no estábamos del todo preparados para la magnitud de tal crisis sanitaria y de sus consecuencias económicas derivadas y que los efectos han sido desiguales en las economías de nuestro continente, algo que la Unión Europea ha tratado de mitigar. Siempre se dice que es importante conocer el pasado para no repetir los errores en el futuro, y este caso nuestro trabajo busca reflejar con fidelidad el impacto que la vertiente epidemiológica ha tenido en la economía, para así, gracias a los datos poder enfrentarnos de una manera más consciente y poder mitigar los efectos de una posible nueva pandemia u otra incidencia de una magnitud similar que pudiéramos afrontar en un futuro.

Para la obtención de unas conclusiones válidas y eficaces las tareas a realizar comienzan con la búsqueda y obtención de fuentes de datos de la mayor fiabilidad para después realizar el tratamiento de los mismos con las herramientas de análisis adecuadas y las técnicas de machine learning que nos permitan alcanzar nuestro objetivo. La finalidad es la representación de las conclusiones obtenidas no solamente de una manera numérica o compleja, sino de una manera gráfica y amigable para que el lector pueda de un simple vistazo obtener conclusiones, poniendo también por supuesto los datos más pormenorizados al servicio de las personas que se acerquen a este trabajo con intención de profundizar más o incluso con fines de estudio.

Es indudable la necesidad de un trabajo de este tipo que aborde una problemática que no es que este de actualidad, sino que es la actualidad misma, y que tiene consecuencias directas en la vida de millones de personas. Que existe una correlación entre COVID y crisis económica es algo evidente, pero un análisis pormenorizado de sus efectos en las diferentes economías de la unión, y en diferentes sectores, poniendo el foco en las diferencias entre países, podría llevar a una comprensión que haga que en un futuro las redes de seguridad, las políticas públicas y las propias políticas de las empresas se adecuen mejor y estén más preparadas si se enfrentan a un desafío de magnitudes similares.

En cuanto a los autores del presente trabajo, son los adecuados para realizar los análisis de la cuestión al estar poniendo en práctica los conocimientos adquiridos en el Máster dentro del que se enmarca este trabajo, que aúna los conocimientos de técnicas avanzadas de análisis de datos con materias centradas en la estadística y la econometría y otras disciplinas ya puramente del ramo de la economía. La conjunción de estos conocimientos nos lleva a la comprensión y al conocimiento de cómo manejar grandes volúmenes de datos, como presentar los resultados de una manera atractiva y además no a quedarnos simplemente en los números o gráficos, sino también a obtener unas conclusiones razonadas.

**2.- Fines y objetivos**

Como ya avanzábamos en el punto anterior de introducción, y cómo se deduce del título de este trabajo, el fin del estudio es el de analizar minuciosamente como ha afectado la crisis sanitaria del COVID en las economías de la zona europea, comparando los efectos en los diferentes países y en diferentes sectores de la economía, como pueden ser el turismo, el ocio o las ventas en línea, y para ello emplearemos herramientas para el análisis de datos y su representación gráfica, estableciendo unas conclusiones destacadas que extraeremos de dichos análisis.

Es por ello que en el trabajo, siguiendo los métodos científicos, se hará una exposición de las fuentes que se han utilizado para la obtención de los datos y se detallaran de una manera pormenorizada las operaciones analíticas y estadísticas que realizaremos con los mismos, para después mostrar los resultados de manera que sean fáciles de comprender. Dividiremos dichos datos por países y por sectores económicos estudiados.

Una vez mostrados los datos, y digamos, la parte más analítica, será el turno de presentar las conclusiones obtenidas a la vista de los mismos.

**3.- Estado del arte**

No cabe duda de que a fecha de publicación de este estudio -diciembre de 2021- se habrá hablado y se habrán escrito páginas y páginas sobre los efectos económicos de la pandemia de la COVID-19. Un fenómeno de tan gran escala con consecuencias humanas tan devastadoras ha tenido por supuesto una repercusión en las economías mundiales, y, como ya hemos indicado, ello es objeto de estas páginas, en las que trataremos de hacer un análisis pormenorizado desde la estadística y la aplicación de técnicas de machine learning, para tratar de alcanzar unas conclusiones científicas y veraces.

Es objeto de este punto el ver en qué momento se encuentran los estudios sobre la cuestión, y hacer mención a otras personas o entidades que hayan abordado la misma, o cuestiones similares. Por supuesto el Banco Central Europeo en su informe del año 2020 hace referencia a la cuestión, de la que podemos ver un avance en su resumen del año (“El año en síntesis”), y ya adelanta que *“La actividad económica se contrajo notablemente durante el primer semestre del año como consecuencia de las medidas de confinamiento y del aumento de la aversión al riesgo.”* y cómo las políticas públicas y la respuesta favorable a la vacunación de los ciudadanos “ayudaron a estabilizar la actividad en el segundo semestre”, pero en conjunto *“el PIB de la zona del euro se contrajo un 6,6 % en 2020. La inflación general anual se redujo hasta el 0,3 %, desde el 1,2 % de 2019, en gran parte como resultado del descenso de los precios de la energía, aunque también de otros factores relacionados con la pandemia”*. Desde las instituciones europeas, y también a nivel estatal se han fomentado políticas para paliar los efectos de unos confinamientos y toques de queda que podrían haber sido mucho peores de no haber recibido una respuesta por parte de nuestros gobernantes, y así “*Estas medidas incluyeron un nuevo programa de compras de emergencia frente a la pandemia de carácter temporal, la relajación de los criterios de admisión de los activos de garantía, y nuevas operaciones de financiación a plazo más largo. La respuesta de política monetaria fue crucial para estabilizar los mercados y contribuyó a hacer frente a los importantes riesgos derivados de la pandemia”*.

El Centro Internacional de Investigación e Información sobre la Economía Pública, Social y Cooperativa, en el número 63 de su revista -año 2020- señala la importancia de la economía social, que ha servido y funcionado como un colchón para los ciudadanos. Hace mención esta publicación a la Carta Abierta de Social Economy Europe a la Comisión Europea, al Parlamento Europeo y al Consejo Europeo titulada “Una crisis sin precedentes que requiere una respuesta sin precedentes de la UE para restablecer el progreso económico y social” en la que se detalla que la economía social ha proporcionado una asistencia sanitaria para todos y ha hecho hincapié en los colectivos de más vulnerabilidad como ancianos o migrantes y ha asegurado la prestación de servicios básicos para proteger a las familias en riesgo de exclusión. También han promovido políticas de favorecimiento del teletrabajo como forma de mantener de la mejor manera posible la productividad y mantener los empleos.

Las medidas de confinamiento y otras como las limitaciones de aforo o la prohibición de reuniones han afectado a diversas industrias, pero una de las que más lo ha sufrido han sido la industria de la cultura. La consultora Ernst & Young ha publicado en 21 de enero de 2021 el estudio “Rebuilding Europe: 2ème Panorama européen des industries culturelles et créatives”, encargado por **GESAC** (Agrupación Europea de Sociedades de Autores y Compositores), en el que se trata esta cuestión y se analiza el impacto del covid-19 en las Industrias Culturales y Creativas (ICC), y su potencial como un motor de crecimiento y de reactivación de la economía en Europa. En el estudio se llega a la conclusión de que estas industrias se han visto más afectadas que la industria del turismo y casi tanto como la industria del transporte aéreo, llegando a cifras alarmantes, como que las industrias culturales han asumido unas pérdidas de más del 30% de su volumen de negocios en el año 2020, es decir, una pérdida acumulada de 199.000 millones de euros. Recomienda este estudio en sus conclusiones que se apliquen políticas basadas en financiación pública, la promoción de la inversión privada y a su vez la creación de marcos legales que favorezcan o faciliten el funcionamiento de estas industrias para reactivarlas, y a la vez para que estas sean un paradigma de reactivación cultural.

El Istututo de la Juventud de España ha publicado el análisis “*JUVENTUD EN RIESGO: análisis de las consecuencias socioeconómicas de la COVID-19 sobre la población joven en España*” en el que, por supuesto, se centran en este espectro de la población , y del que se extraen, como podemos imaginar, unas consecuencias para los jóvenes que no han sido muy favorables. Así señalan que *“La tasa de paro de la población joven en España se sitúa en el 25,2% durante las primeras semanas del confinamiento, registrando un incremento trimestral más de dos veces superior al que se ha dado entre la población de 30 a 64 años.”* y asimismo que *“una vez más, las personas jóvenes son las que han experimentado con mayor intensidad los efectos del parón económico derivado del confinamiento. Así se refleja, por ejemplo, en la caída de la tasa de actividad (que mide la proporción de personas jóvenes que tienen un empleo o lo buscan activamente) y en la caída de la tasa de empleo”*. Este estudio se circunscribe a un ámbito estatal español, pero no es descabellado pensar que el reflejo sea similar en todas las economías de la zona euro.

En el ámbito español se encuadra también el inorme “La digitalización de la economía 1/2021”, actualización del mismo informe de 2017, que “, *la crisis ha puesto de manifiesto la necesidad de acelerar el proceso de digitalización de manera convergente, es decir, prestando especial atención a esas brechas digitales. En primer lugar, en el ámbito laboral, donde la irrupción del COVID-19 ha acelerado algunos procesos, como muestra el peso que ha cobrado el teletrabajo*” y se centra sobre todo en señalar algo que por obvio merece también ser objeto de estudio y pormenorización, y es la cuestión de que la crisis económica derivada del Covid ha afectado más a las regiones y zonas en las que los porcentajes de digitalización son más bajos, datos, que al igual que en el estudio que mencionábamos con anterioridad – del INJUVE- , se centran en España, pero es lógico que se puedan extrapolar a los diferentes países, pues una digitalización mayor permite paliar en cierta medida consecuencias económicas, ya que permite el teletrabajo y mantener un porcentaje de las ventas de determinadas industrias gracias al comercio en línea.

**II.- Parte General**

**4.- DESCRIPCIÓN Y TRATAMINETO DE LOS DATOS**

En este capítulo se procede a explicar el proceso de extracción de los datos así como la utilización de los mismos en diferentes modelos de Machine Learning. Todo lo mencionado se ha implementado en Python.

**4.1.- OBTENCIÓN DE LOS DATOS, LIMPIEZA DE LOS MISMOS Y UNA PRIMERA APROXIMACIÓN AL MODELO**

Para esta primera aproximación a lo que se convertirá finalmente en el trabajo de fin de Máster presentamos un acercamiento a la obtención de los datos y los modelos que vamos a utilizar, a modo de demostración.

El análisis de los datos por ejemplo se reduce solamente a un periodo de seis meses, pero ya están presentes los seis países que van a ser objeto de nuestro estudio, tales son España, Portugal, Alemania, Suecia, Ucrania y Letonia, que hemos escogido en un intento de representar un escenario que recoja la diversidad existente en Europa, ya que el Covid 19 ha afectado de manera diferente a cada país, y la respuesta de los estados a la pandemia ha sido también diferente.

Por tanto, el primer paso fue elegir los países mencionados, y la siguiente tarea es la de la obtención de los datos correspondientes a los mismos para nutrir las variables que compondrán el estudio. En este caso hemos utilizado la plataforma “kaggle”, un repositorio de sets de datos muy popular entre los científicos de datos. El dataset escogido para esta primera aproximación ha sido el siguiente:

<https://www.kaggle.com/imdevskp/corona-virus-report?select=full_grouped.csv>

El motivo de la elección es que teniendo ya los países que, como hemos dicho, se eligieron por motivos de diversidad, distancia geográfica y su diferente acercamiento a cómo afrontar la pandemia, necesitábamos un set de datos que en el que estuvieran incluidos, y que al mismo tiempo contara también con las variables de Covid requeridas, tales como casos confirmados, muertes, etc. y ambos requisitos se cumplen en el set de datos que finalmente hemos elegido.

El dataset cuenta con 604 observaciones y 11 columnas. La variable dependiente Bolsa representa el valor diario de bolsa para cada uno de los países. Las variables independientes que usaremos para predecir la variación en el valor de la bolsa son los números de casos confirmados, las muertes, las recuperaciones, los casos activos, los nuevos casos y las nuevas recuperaciones

Una vez cargados los datos comienza el tratamiento de los mismos, que ha consistido -siempre hablando de esta demostración preliminar- en una primera limpieza para eliminar del dataset los países que no nos interesaban en el estudio y en un segundo paso que consiste en eliminar los apuntes referentes a los meses anteriores al mes de marzo de 2020, ya que para los países elegidos no existen datos antes de esa fecha.

Después transformamos la variable fecha para establecer un formato en el que se separan el día, mes y año, formato que es mucho más conveniente a la hora de trabajar después con estos datos. La variable fecha se elimina al no ser ya necesaria. Hubo que transformar también la variable “Bolsa”, ya que al tratarse de una variable continua sus entradas de datos deberían de ser del tipo “float”.

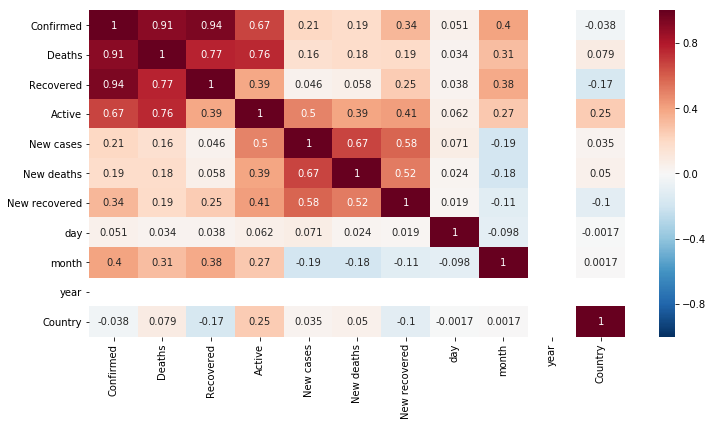


Figura 1. Correlaciones entre variables

Se observa que existe una correlación importante entre algunas de las variables. Destacamos que la variable 'Confirmed' tiene una correlación de 0,91 con la variable 'Deaths' y una correlación de 0.94 con la variable 'Recovered'. Si bien hay varias correlaciones relativamente altas, no se eliminará ninguna de ellas, ya que con el número de variables que presenta el dataset no es del todo necesario. Además en este caso, la correlación de estas variables pueden demostrar el impacto que la enfermedad COVID-19 ha tenido en los países en cuanto a la relación de confirmados y muertes. De hecho, la primera conclusión que podemos sacar de esta matriz de correlaciones, es que hay una mayor correlación entre el número de confirmados y recuperados que entre el número de confirmados y muertes, de forma que hay un mayor porcentaje de recuperaciones frente a muertes.

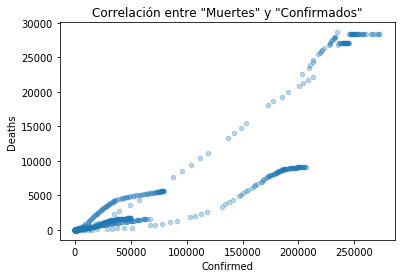
 

Figura 2. Correlación existente entre “Muertes”-“Confirmados”   
y “Recuperados”-“Confirmados”

Como comentábamos en un inicio esta es solamente una primera aproximación a lo que será un estudio de los datos mucho más pormenorizado, pero que en el punto actual ya sienta las bases de lo que será nuestro trabajo.

**4.2.- VISUALIZACIÓN DE LOS DATOS**

Para una mejor comprensión del escenario que se está estudiando, se va a hacer una visualización de los datos.

Se va a dibujar la evolución temporal de cada una de las variables principales: valor de la bolsa, número de fallecidos, casos confirmados, población contagiada y población recuperada por COVID-19 en cada país y, para cada día, así como la de la tasa de mortalidad. Para ello, comenzaremos guardando la información de cada país en una variable, se controlará el tamaño del gráfico total, que corresponderá al tamaño (16, 45).

El gráfico total, estará compuesto por un total de doce subgráficas, dos por cada uno de los seis países que componen el análisis, uno indicando la evolución de los casos y otro la evolución de la tasa de mortalidad de cada país. Relacionado con la evolución de los casos, será una gráfica de líneas que relacionará las distintas variables mencionadas con la fecha, para así poder hacer una comparación entre los parámetros.

En cuanto a la evolución de la mortalidad, será una gráfica de dispersión de relación entre tasa de mortalidad y casos confirmados.

A continuación, se puede observar el gráfico resultante.

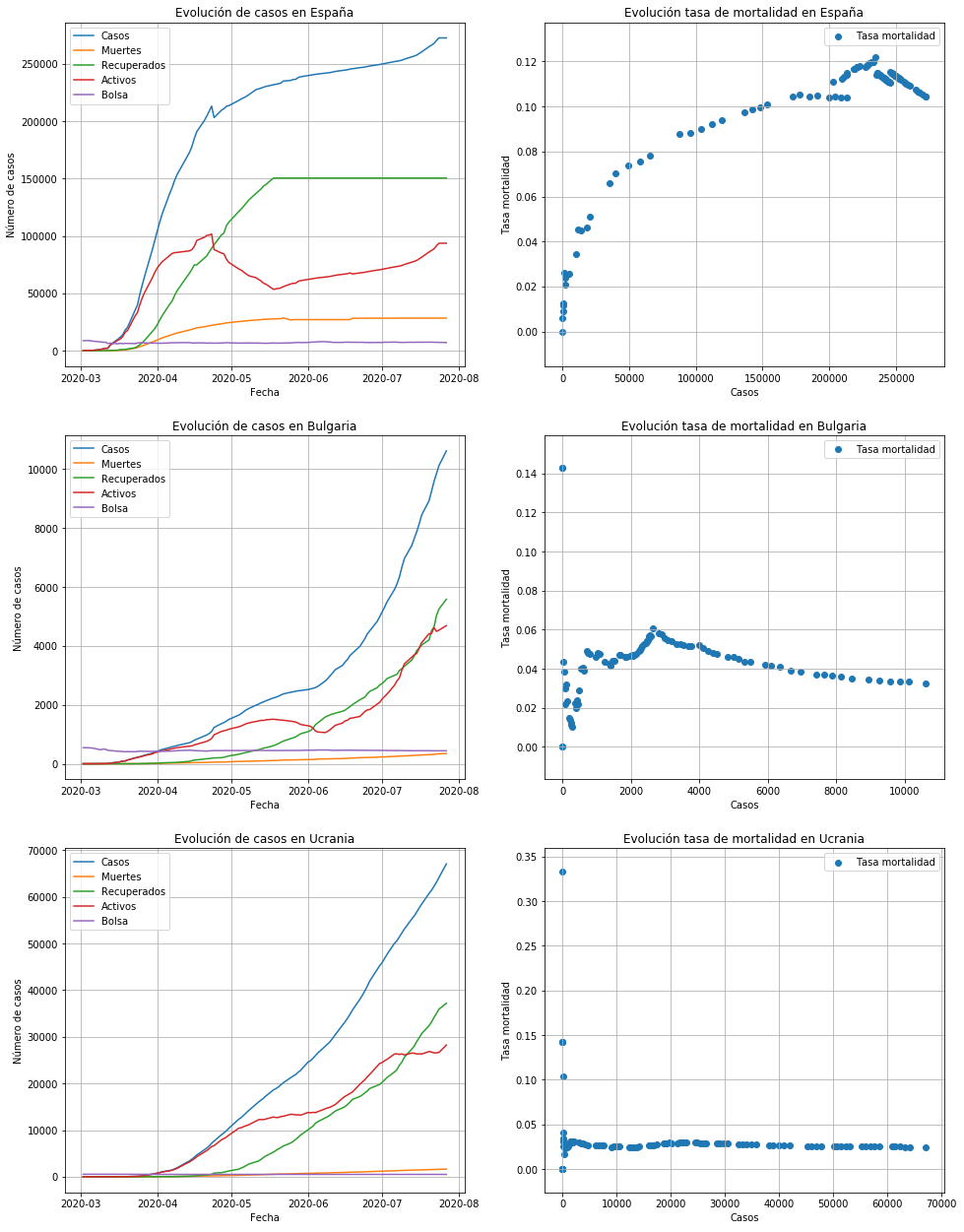


Figura 3.1. Evolución de casos y tasa de mortalidad

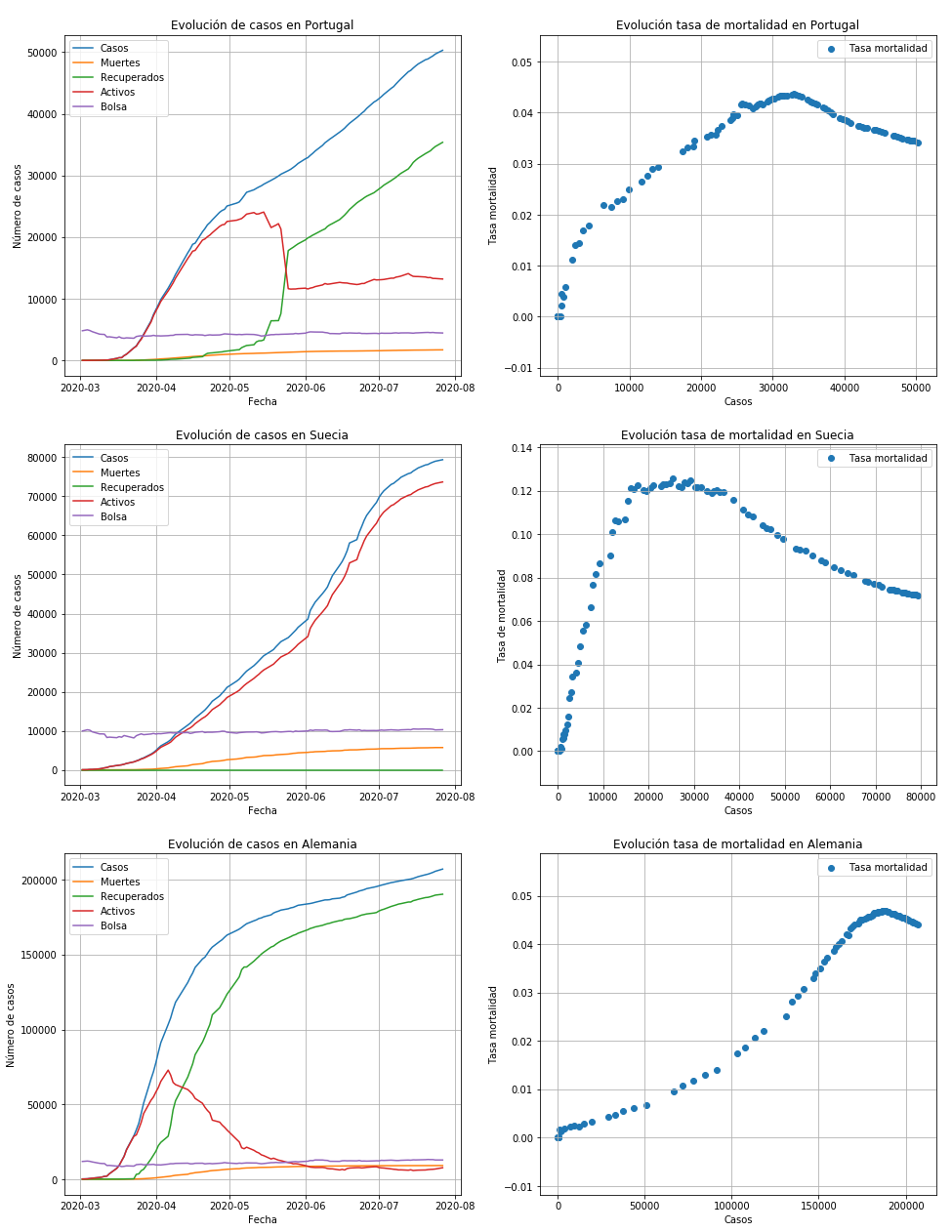


Figura 3.2. Evolución de casos y tasa de mortalidad

Algunas de las conclusiones que podemos sacar son que: (i) el mayor número de muertes lo alcanzó España, (ii) y que tanto Bulgaria como Ucrania muestran un crecimiento de contagios exponencial, lo que puede traducirse en una toma de medidas sanitarias más tardías que en los demás países. También podemos observar que la tasa de mortalidad se dispara con el número de casos confirmados, especialmente en el caso de España y Alemania. Sin embargo, la mayor tasa de mortalidad la alcanza Suecia, seguida de España, y el país con menor tasa de mortalidad de nuestro estudio es Ucrania, el cual se mantiene más o menos constante.

Para continuar, se van a visualizar dos gráficos de barras. El primero con datos sobre el número de fallecidos y casos confirmados por país, y el segundo mostrará la media de fallecidos y casos confirmados por mes, englobado para todos los países. Este último mostrará en color rojo la media mensual de fallecidos y en color azul los casos promediados.

Con estos gráficos podemos observar cómo España se coloca en primer lugar respecto al número de muertes y contagios y cómo asciende la tendencia a medida que pasan los meses, siendo Julio el mes más “problemático” para la crisis sanitaria.

Gráfico

Descripción generada automáticamente

Figura 4. Gráfico de barras fallecidos – casos

A continuación, se hará una comparación de distribuciones entre pares de variables, mediante la función de Python del paquete Seaborn: pairplot( ).

Imagen en blanco y negro

Descripción generada automáticamente con confianza media

Figura 6. Distribuciones entre pares de variables.

**4.3.- PREDICCIONES. TRAIN Y TEST**

Los datos de entrenamiento o *training data,* son los datos que usamos para entrenar un modelo. La calidad de nuestro modelo de aprendizaje automático va a ser directamente proporcional a la calidad de los datos. Por ello las labores de limpieza, depuración o *data wrangling* consumen un porcentaje importante del tiempo de los científicos de datos.[1]

Los datos de prueba, validación o *testing data* son los datos que se reservan para comprobar si el modelo que hemos generado a partir de los datos de entrenamiento puede funcionar. Es decir, si las respuestas predichas por el modelo para un caso totalmente nuevo son acertadas o no.[1]

Además, es importante que el conjunto de datos de prueba tenga un volumen suficiente como para generar resultados estadísticamente significativos, y a la vez, que sea representativo del conjunto de datos global. El conjunto de datos se suele repartir en un 70% de datos de train y un 30% de datos de test. [1]

Por tanto, en cuanto a predicciones de la bolsa de valores, es importante que no se elijan los datos de train y test aleatoriamente. En este caso, se tendrá que usar los datos más recientes para el conjunto de test, ya que proporcionará unos mejores resultados para nuestras predicciones.

Comenzamos separando el conjunto de test, mediante la función train\_test\_split( ), la cual no se modificará en ningún momento hasta el final. En cuanto al conjunto train, del cual iremos sacando folds de validación cruzada, solo se va a cuando se haga Grid Search.

Antes de hacer los modelos y validaciones cruzadas, se deben plantear una serie de decisiones sobre los datos, si se deben descartar o seleccionar variables, si se deben quitar outliers o si se debe estandarizar las features.

Al hacer validación (cross validation), así como Grid Search, es recomendable aplicar estos pre-procesados, exclusivamente sobre los conjuntos de train, y no sobre los de validación. Los modelos, así como todos los posibles pre-procesados que los acompañan, solo deben ser entrenados sobre el conjunto de train, con el que el modelo entrena. Si no lo hacemos así, filtramos información a los modelos más allá de la de los datos con los que entrenan, y esto ocasiona sobreajuste.

La respuesta a la primera pregunta sobre el descarte/ selección de variables es que no sería necesario. Esencialmente todos los modelos tanto de clasificación como de regresión modernos pueden lidiar con unas cuantas decenas de features sin que resulte un problema.

Con respecto a la segunda pregunta sobre el preprocesamiento de outliers, dado que el problema que se analiza parte de una situación extraordinaria en la economía sería posible encontrar outliers en aquellos puntos más críticos de la pandemia. Sin embargo, estos outliers son necesarios para el completo análisis y entendimiento de los datos, con el fin de entender la magnitud del impacto que este suceso ha tenido en la economía. Prescindir de estos datos, sería perder fiabilidad en nuestro análisis.

La estandarización de features será necesario para que el modelo pueda tratar las variables de forma homogénea y no darles mayor peso a aquellas features aparentemente más grandes debido a diferentes magnitudes, unidades y rango. En todos los modelos de sklearn debemos estandarizar las features excepto en los árboles de decisión.

Tras la separación del dataset en el conjunto de tren y de test comenzaremos a aplicar los modelos de machine learning a los conjuntos de datos. El primer modelo elegido será un arbol de decisión regresor. En general, si el modelo a tratar es uno basado en árboles podemos ahorrar la parte de pre-procesado de selección de variables ya que los árboles son capaces de seleccionar bien por motu proprio.

**4.4.- ÁRBOL DE DECISIÓN**

Para la generación del árbol de decisión, en primer lugar, se va a hacer una medición sobre el conjunto de train, mediante los métodos fit( ) y predict( ).

Después, se seleccionarán los hiperparámetros. A la hora de seleccionar los parámetros para el árbol regresor, probaremos diferentes combinaciones para saber cuál es la combinación que mejor funciona.

**5.- VISUALIZACIÓN DE RESULTADOS**

Para una rápida visualización de los datos de Covid analizados y del modelo que se ha considerado que más se ajusta a lo que se está estudiando, se ha desarrollado una página web con esta información contenida.

**5.2.- HERRAMIENTAS UTILIZADAS**

**start bootstrap**

Start bootstrap es una biblioteca multiplataforma o conjunto de herramientas de código abierto para diseño de sitios y aplicaciones web.

En primer lugar, para comenzar con el diseño y desarrollo de la página web se ha utilizado una plantilla de StartBootstrap [3] a partir de la cual se han hecho las modificaciones según los gustos y preferencias del proyecto.

Para la modificación del código dado por la plantilla se ha utilizado un entorno de desarrollo llamado IntelliJ.

**intelli-j idea**

IntelliJ IDEA es un entorno de desarrollo integrado (IDE) para el desarrollo de programas informáticos. [2]

Realmente no hubiese sido necesario utilizar un entorno de desarrollo tan potente, con un simple editor de texto podría haberse desarrollado el código, pero se ha decidido utilizar este por comodidad a la hora de visualizar la estructura del proyecto y la capacidad de edición del mismo.

**github**

Github es un portal creado para alojar el código de las aplicaciones de cualquier desarrollador. Como su nombre indica, la web utiliza el sistema de control de versiones Git. [4]

Además, permite que los desarrolladores creen repositorios, que contienen las distintas versiones del proyecto, de forma gratuita. [4]

Se ha utilizado GitHub para así poder disponer todos los miembros del equipo la última versión del proyecto. Se ha alojado en el repositorio <https://github.com/crismadrid28/tfm.git> que contiene la estructura del proyecto.

**5.3.- ESTRUCTURA DEL PROYECTO**

El proyecto tiene la siguiente estructura, cuenta con cuatro directorios y el archivo *index.html*.

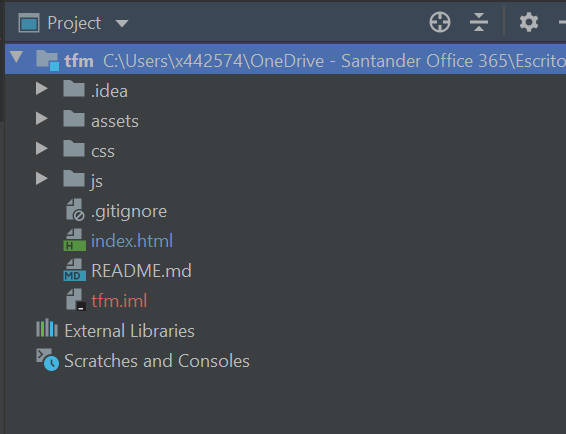


Figura 7. Estructura del proyecto

El directorio .*idea* contiene las configuraciones generales del programa, por tanto nada de lo contenido en este apartado ha necesitado ser modificado.

En el directorio *assets* podemos encontrar todas las imágenes e iconos que se utilizan en el diseño de la página web.

Si seguimos con el *css*, tendremos el fichero *styles.css*. CSS significa ‘hojas de estilo en cascada’, por tanto en este fichero .css encontraremos los estilos y las orientaciones de todos los elementos añadidos.

El directorio *js* contiene los controles de todos los botones y elementos que se pueden pulsar y es donde se indica las redirecciones de cada uno de ellos, a priori, no hemos necesitado añadir más botones a parte de los ya incluidos en la plantilla inicial, por tanto este apartado tampoco ha sido necesario modificarlo.

Por último, y más importante, el *index.html*, es el fichero principal, donde se han añadido todas las modificaciones tanto de texto como de imágenes. Contiene código escrito de lenguaje HTML (lenguaje de marcado de hipertexto).

Ya que se trata de datos que no me actualizan cada día, ya que las fechas examinadas son pasadas, se ha desarrollado una página web hardcodeada, es decir, se han utilizado valores de comportamiento fijo.

**5.3.- ANALISIS FUNCIONAL**

En este apartado, se van a mostrar todas las funciones que podrá realizar el sitio web mediante un diagrama de acciones.

**Vista inicial**



**NOMBRE APLICACIÓN**

**MENÚ**

Figura 8. Vista inicial

En el menú superior derecho se encuentran las secciones de la página web a las cuales podemos acceder. Todo el margen superior es fijo, por tanto se podrá acceder a él desde cualquier vista.

El menú cuenta con una sección de ‘países’, ‘evolución’, ‘predicciones’ y …

En la parte central de la vista se muestra el nombre que se ha escogido representativo de la página. Se ha escogido este, ya que se van a analizar los datos durante el Covid en los países escogidos y mediante modelos de machine learning se han hecho predicciones.

Para acceder a las siguientes vistas, tenemos dos opciones, utilizar el menú ya explicado o simplemente haciendo scroll-down.

**Vista países**

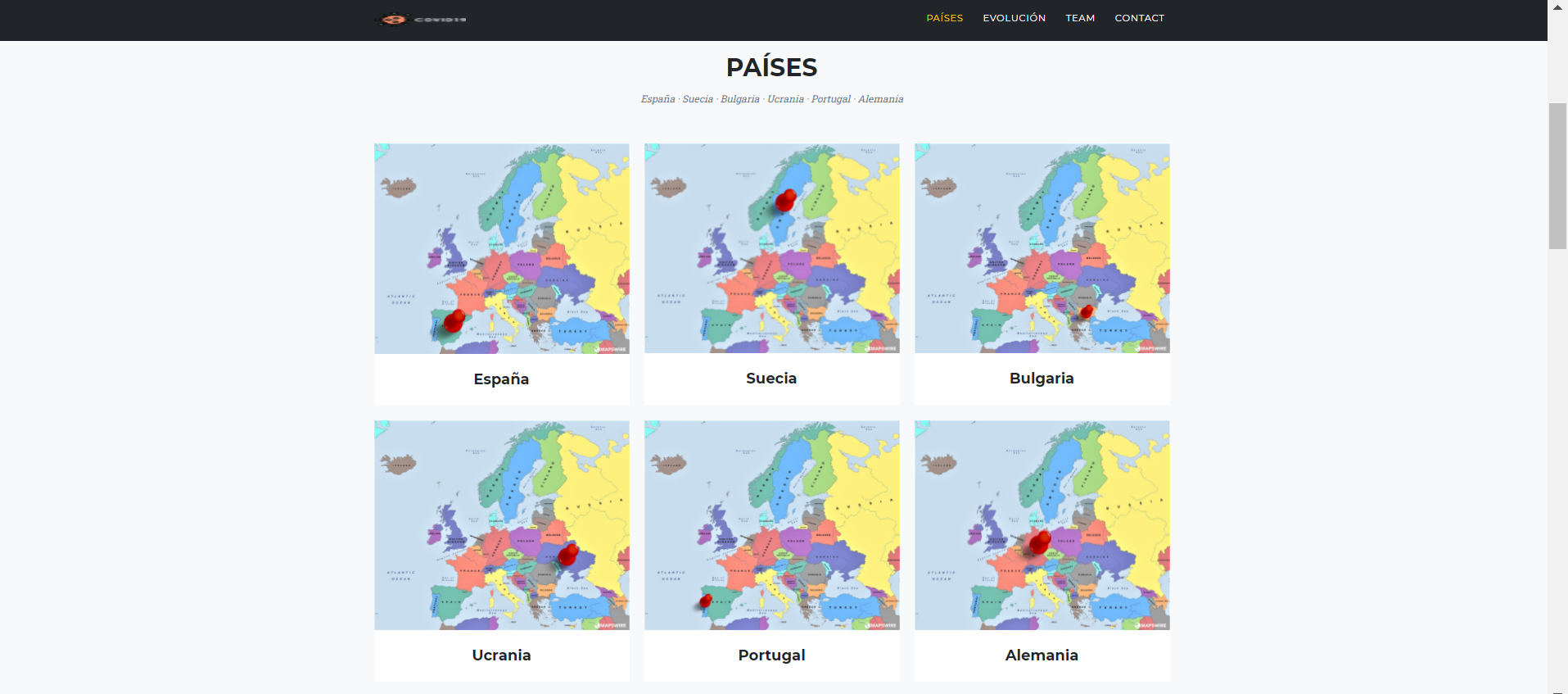


Figura 9. Vista países

En esta vista se encuentran los países que se han tenido en cuenta para el estudio. Estos son ‘España’, ‘Suecia’, ‘Bulgaria’, ‘Ucrania’, ‘Portugal’ y ‘Alemania’.

Junto al nombre de cada país se encuentra un mapa con una chincheta indicando la localización de cada uno de ellos.

A su vez pulsando sobre el mapa de cada país encontramos información concreta sobre los casos y fallecidos entre las fechas seleccionadas, se mostrarán en forma de pop-up.

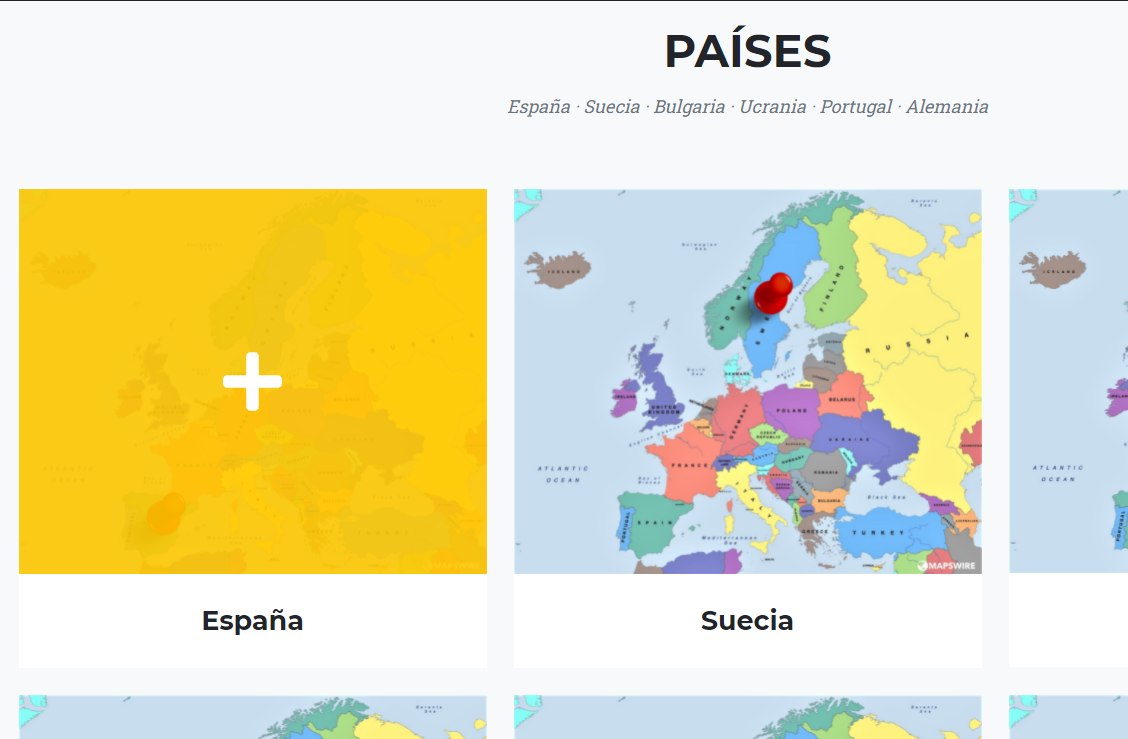


Figura 10. Vista acceso pop-up



Figura 11. Vista pop-up

Cada pop-up, cuenta con información relacionada con los casos, fallecidos, recuperados, activos y el mercado de la bolsa. Estos gráficos se han obtenido según lo explicado en el apartado ‘Visualización de los datos’.

Además se encuentra una breve explicación del gráfico, un resumen de los casos totales, fallecidos totales, casos de los últimos 28 días y fallecidos de los últimos 28 días. Esta información se ha obtenido de la página de la universidad de Johns Hopkins [5], que desde los inicios del Covid han desarrollado una página web con datos actualizados cada día y considerándose de las páginas con la información más fiable de la pandemia.

Por último, podremos cerrar el pop-up, con el botón ‘cerrar’ que se encuentra en la parte inferior central.

De esta misma forma, se podrá acceder a esta información sobre el resto de países.

**Vista evolución**

En la pantalla que se muestra a continuación, se sigue poniendo en contexto del entorno que se está estudiando. Para esto, se ha decidido, hacer un eje cronológico con los principales hechos de la pandemia, la fecha del primer caso confirmado de cada país, la fecha del inicio del confinamiento de cada país y el porcentaje de vacunas administradas a la población de cada país.



Figura 12. Vista evolución

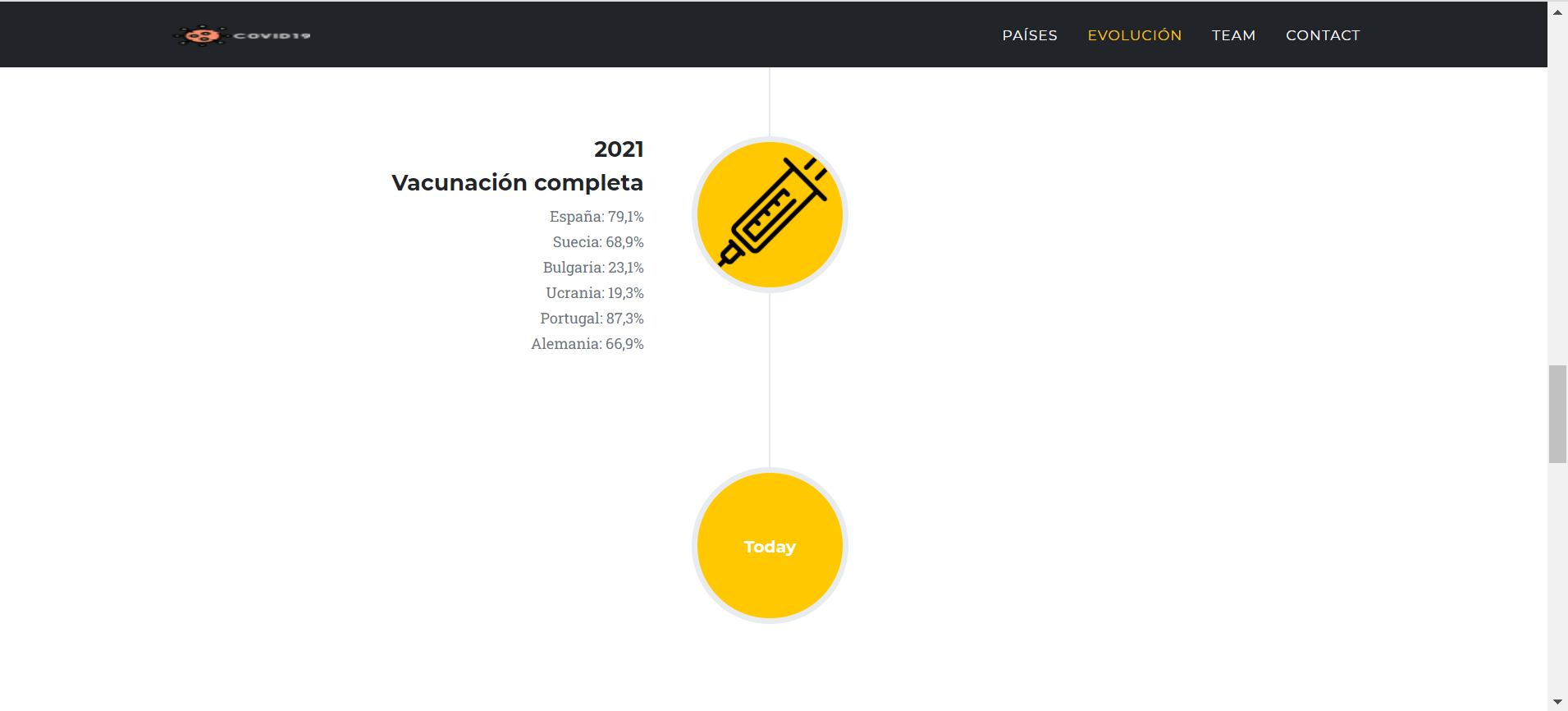


Figura 13. Vista evolución 2.0.

**5.4.- SERVIDOR**

**5.5.- RESULTADOS**

**3.- Metodología y estado de la cuestión**

[Párrafo de enfoque; párrafo de técnicas de investigación; párrafo estado de la cuestión; e.g. El presente estudio posee una vocación interdisciplinaria, procurando conectar diversas ciencias, y las mismas con la realidad social y sus demandas. Por tanto, se parte de un enfoque propio de las Ciencias Empresariales, para conectar con otras Ciencias Sociales (e.g. Derecho, Ciencia Política, Sociología, Psicología), y Humanidades (e.g. Historia, Filosofía, Comunicación) (…) Se han empleado diversas técnicas de investigación social, tanto cuantitativas (e.g. rastreo de ítems, encuestas), como cualitativas (e.g. análisis de discurso y de contenidos, estudio de caso). En concreto, se ha recurrido especialmente a aquellas técnicas propias de un plan de negocio: DAFO, etc. (…) En lo tocante al estado de la cuestión, que fija dónde está la frontera del conocimiento, se destacan la siguientes publicaciones de los últimos cinco años, que han influido inicialmente en la formulación e implementación de este trabajo. Se citan por tanto los siguientes trabajos (relegándose al apartado de Bibliografía, vid. Anexos, aquellas otras publicaciones significativas, pero que exceden el estado de la cuestión): - (artículo 1: cita y comentario). – (artículo 2: cita y comentario) (…)]

***A partir de esta página, respetando los aspectos mínimos indicados en este documento el grupo puede realizar un diseño particular de su trabajo, incluyendo aquellos elementos gráficos que considere adecuados para la comunicación de su plan***

**II.- Parte General**

**4.- Análisis estratégico ; Externo e Interno**

[Aclarar coordenadas del estudio: espacio (e.g. país, región), tiempo (corto y medio plazo), y materia (sector del negocio)]

**5.- Análisis DAFO**

[punto de partida u origen del desarrollo de la propuesta]

**III.- Parte Especial**

**6.- Objetivos del plan**

[Resumen los principales objetivos o prioridades que aborda el plan de marketing y fragmento que se imprimirá a los miembros del tribunal, luego prepárese al final del mismo un cuadro con nombre y fotos que identifiquen a los autores]

**7.- estrategias del plan**

**8.- Plan de acción / tacticas**

**IV.- Conclusiones**

[Recapitulación, conclusiones generales y particulares, más propuesta de recomendaciones para una mejor puesta en práctica del negocio…]

**V.- Anexos**

**14.- Fuentes de consulta**

**14.1.- Bibliografía**

[1] “Datos de entrenamiento vs datos de test”. Telefónica Tech. [Online]. Disponible: <https://empresas.blogthinkbig.com/datos-entrenamiento-vs-datos-de-test/>. [Accedido: 01-11-2021].

[2] “¿Qué es IntelliJ IDEA?”. Wikipedia. [Online]. Disponible: <https://es.wikipedia.org/wiki/IntelliJ_IDEA>. [Accedido: 13/11/2021]

[3] “Plantillas StartBootstrap”. Start Bootstrap. [Online]. Disponible: <https://startbootstrap.com/previews/agency>. [Accedido: 10/11/2021]

[4] “¿Qué ofrece GitHub a los desarrolladores?”. Xataka. [Online]. Disponible: <https://www.xataka.com/basics/que-github-que-que-le-ofrece-a-desarrolladores>. [Accedido: 13/11/2021]

[5] “Map Coronavirus”. Universidad de Johns Hopkins. [Online]. Disponible: <https://coronavirus.jhu.edu/map.html>. [Accedido: 16/11/2021]

[Sistema APA, vid. diapositivas y documentación en Blackboard]