

THE BRIDGE - DATA SCIENCE - AGOSTO, 2020

COVID 19 - PROYECTO EDA

GRUPO D



José Luis Manjón
Filipa Teles
Mónica Villasuso

Presentación

COVID EN PORTUGAL, ESPAÑA, VENEZUELA, GRAN BRETAÑA Y TURQUÍA

Visión general

Objetivo

Contexto

EDA

Tendencias - Hallazgos

Conclusión

Visión general - Objetivos

Reforzar y practicar lo aprendido hasta ahora
EDA - Librerías de Gráficos - API



Generar sinergias a partir del trabajo en equipo
Roles - eficiencia en el proceso



Importancia de los datos
tratamiento y calidad para la toma de decisiones



CONTEXTO - FUENTES DE DATOS

- **COVID 19**

Fuente de datos (Data Set)

[<https://covid.ourworldindata.org/data/owid-covid-data.csv>]

- **Estado de alarma**

Período y condiciones en el que cada país del grupo decretó el estado de alarma.

- **Índice de mortalidad**

Nos ayuda a comparar la incidencia de la enfermedad en diferentes lugares.. Es la relación entre los fallecidos y el número de casos

- **Grupo D**

Portugal, Venezuela, Turquía, Reino Unido y España

Revisión de ajustes a las cifras realizadas por los países para explicar comportamientos "raros" de los datos.

- **Stringency Index**

Una medida de lo estricto que ha sido cada país en cuanto a las medidas tomadas (en términos de distanciamiento, rigurosidad de normas, etc)

Proceso EDA



GRUPO D

Información de principales indicadores entre países del grupo a analizar

POR PAÍS

Cifras de indicadores contrastadas con la situación específica de cada país

DATOS MUNDIALES

Visión global - mundial y por continente - de las cifras

WRANGLING - MINING - CLEANING

Distribución y análisis de valores tipo Nan. Se observa que la mayoría de ellos se concentran en determinadas filas y columnas.

Eliminación de datos totales mundiales o sin país o sin continente (que afecten los análisis globales). Creación de un dataframe `world_df`

Eliminación de registros con información previa al inicio de la pandemia que no agreguen valor al análisis.

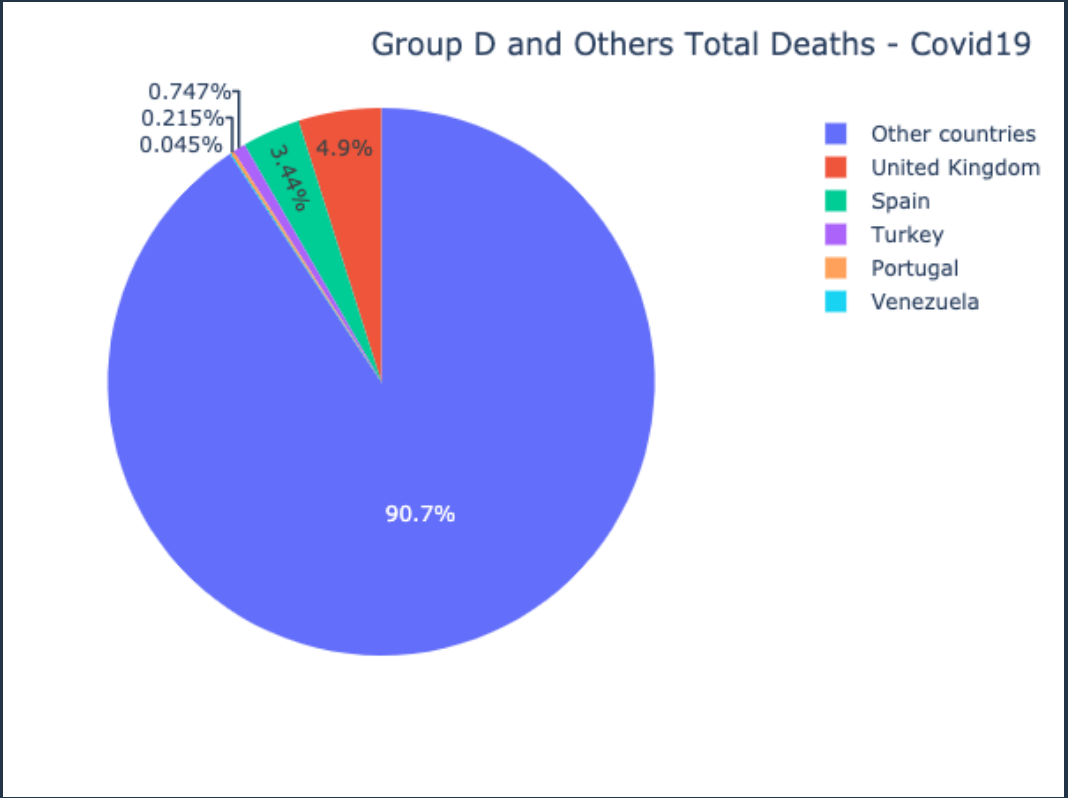
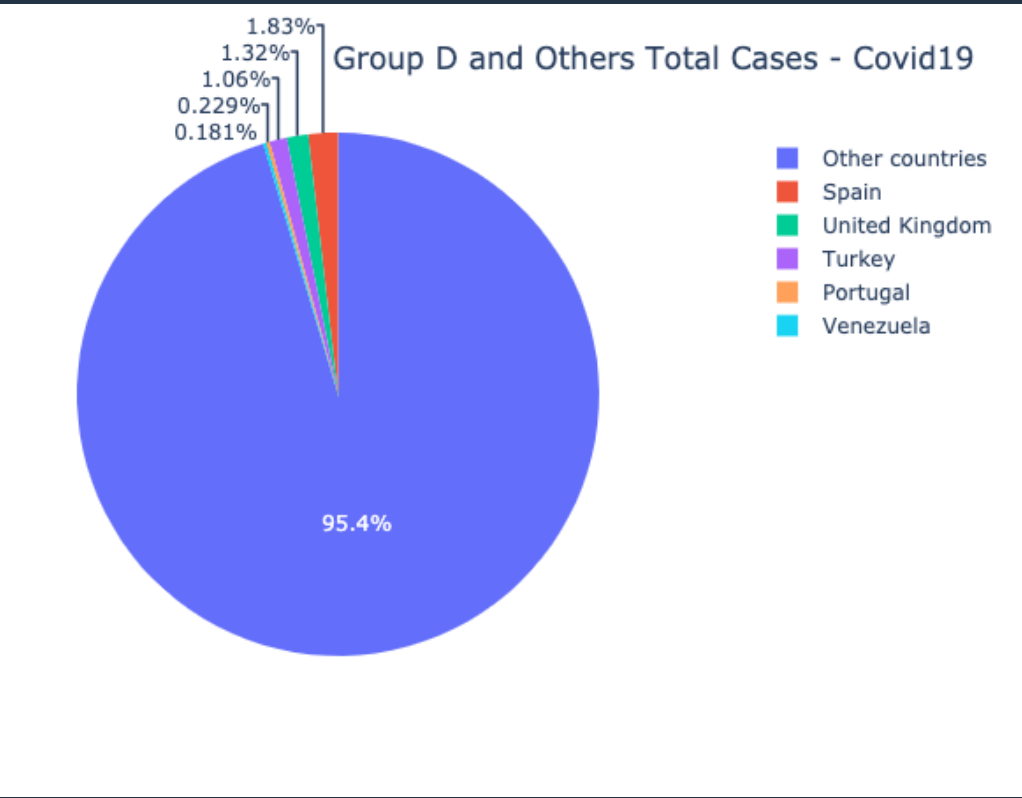
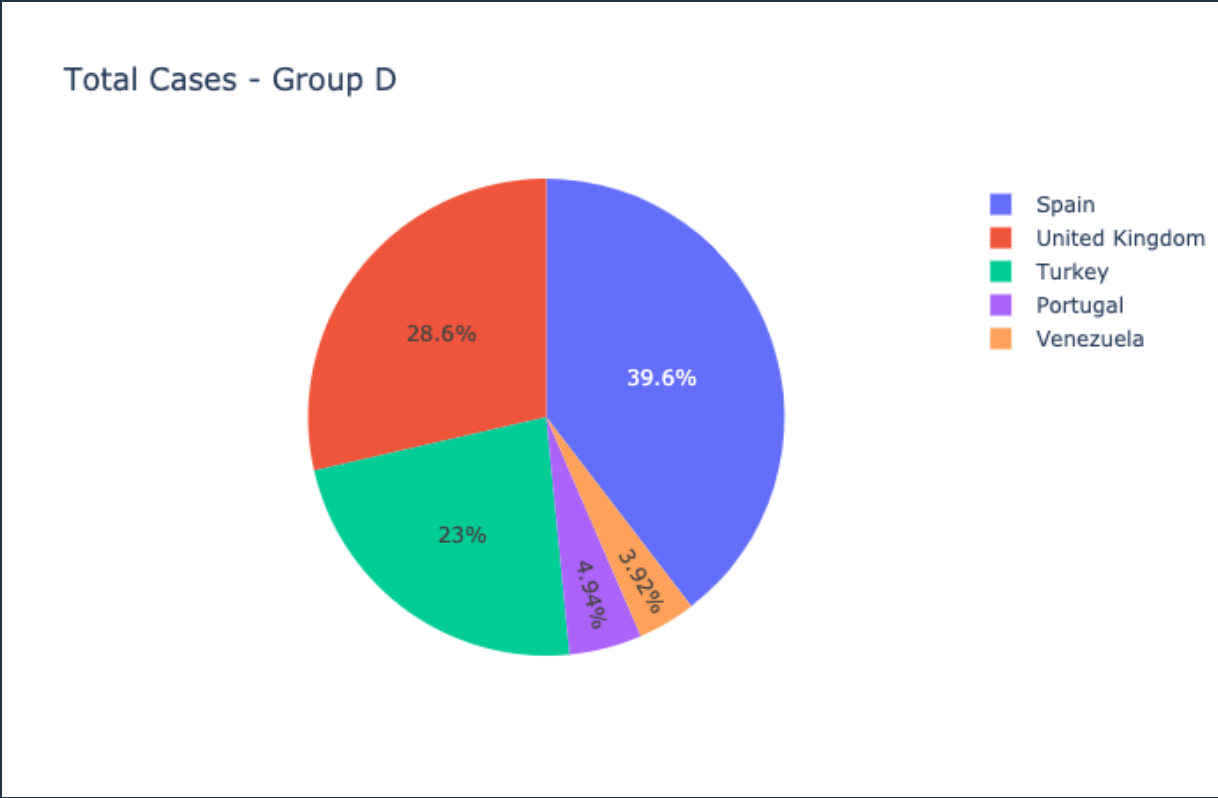
Eliminación de columnas que no contengan información (`handwashing_facilities`).

Creación de un dataframe específico para el análisis de los países del grupo D (`covid_grupoD`) y de variables útiles para el análisis (`países_grupoD`)

GRUPO D

Casos y Muertes Totales

Visión general de los casos y de los fallecimientos en el mundo., junto con la comparativa de casos totales en los países de nuestro estudio.

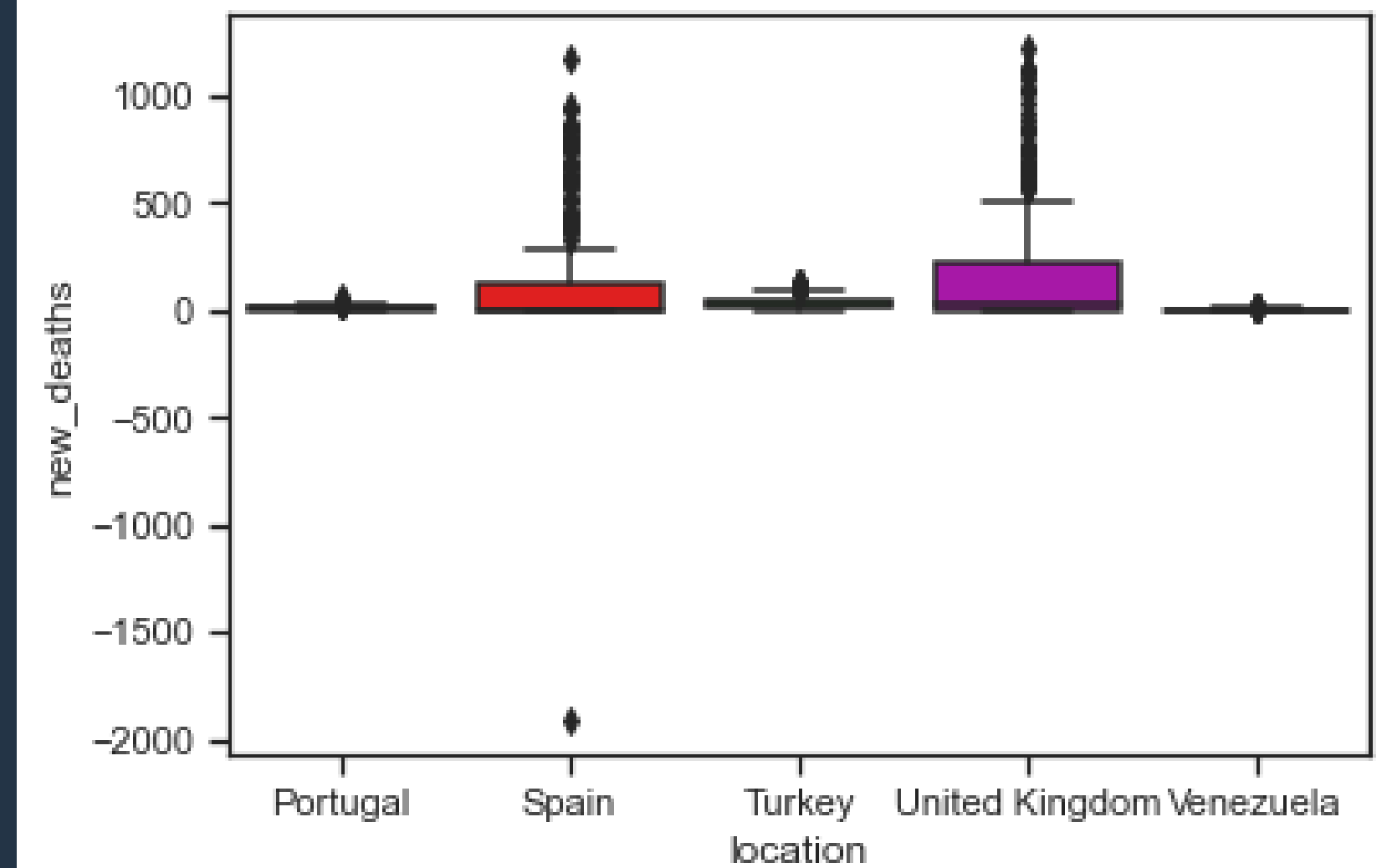
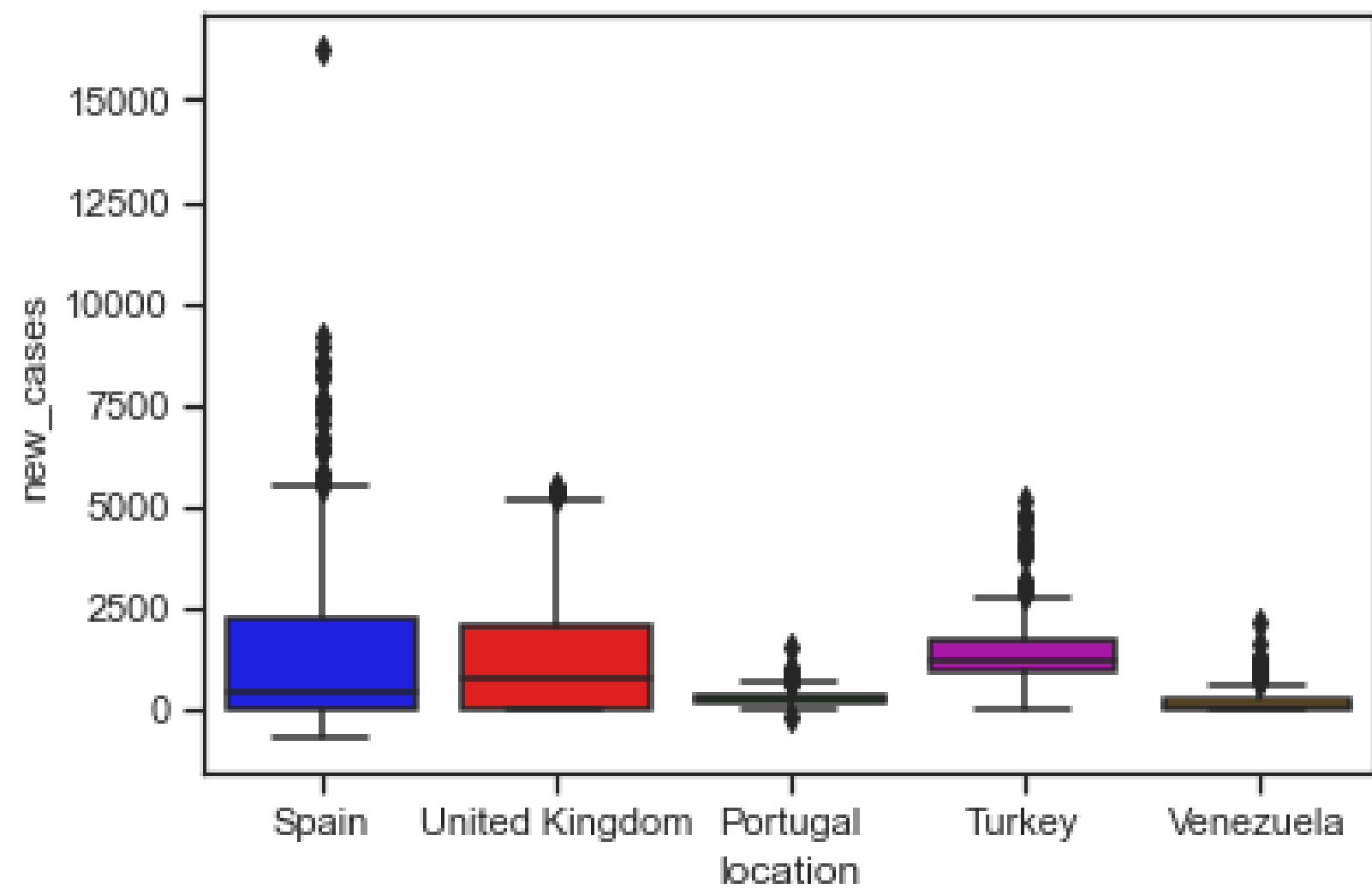


GRUPO D

Outliers

Casos y Muertes diarias

Texto con aspectos a destacar



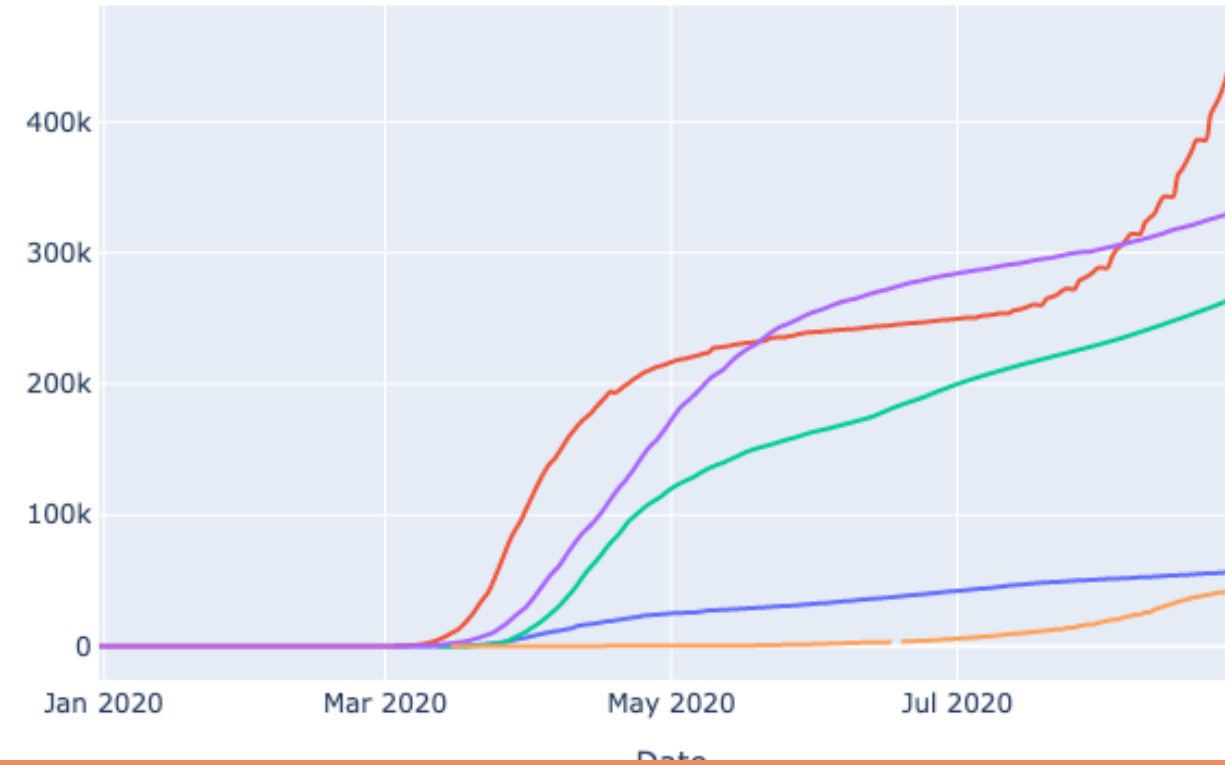
GRUPO D

CASOS - TEST

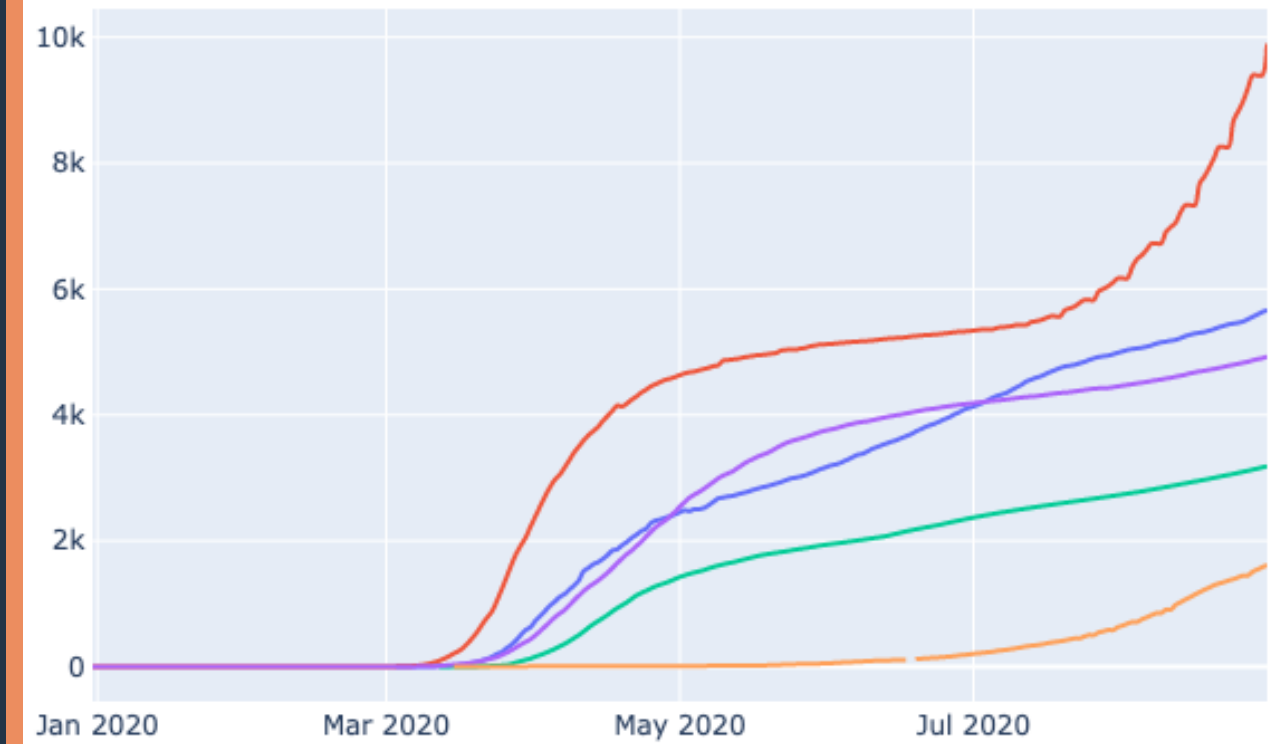
Portugal
España
Reino Unido
Turquía
Venezuela



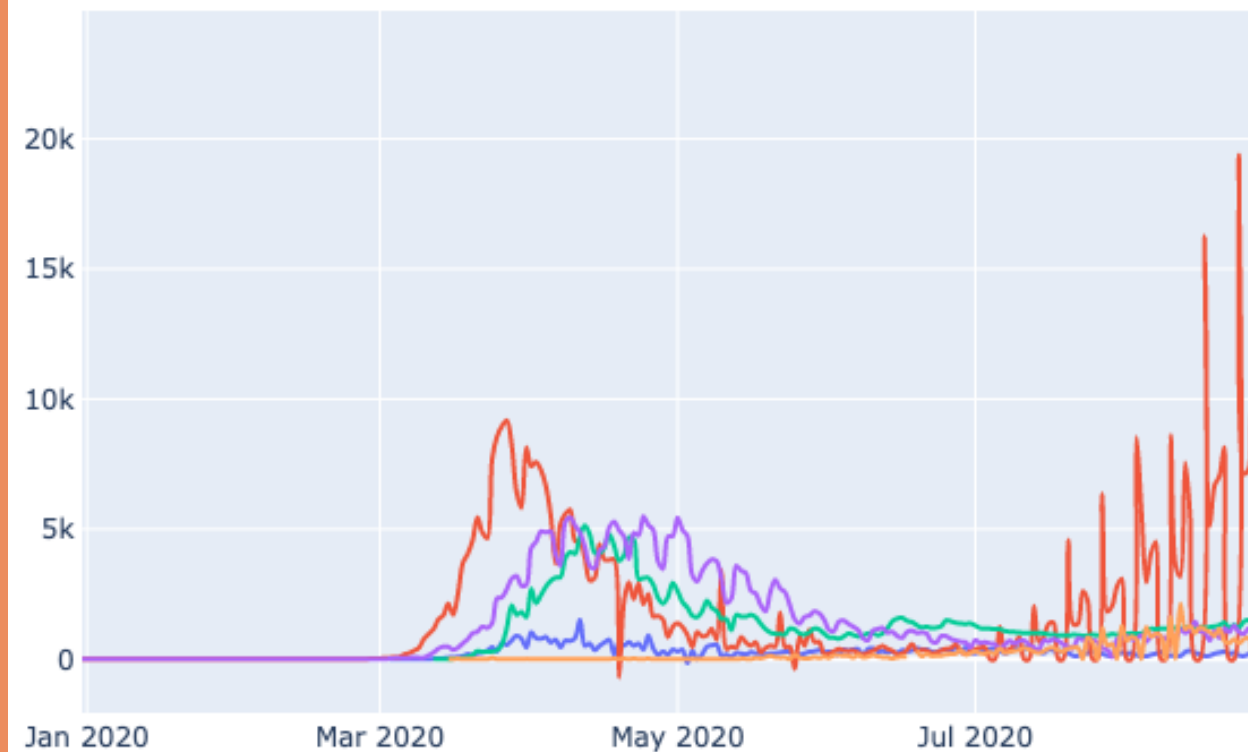
Total Cases by Date and Country - Group D



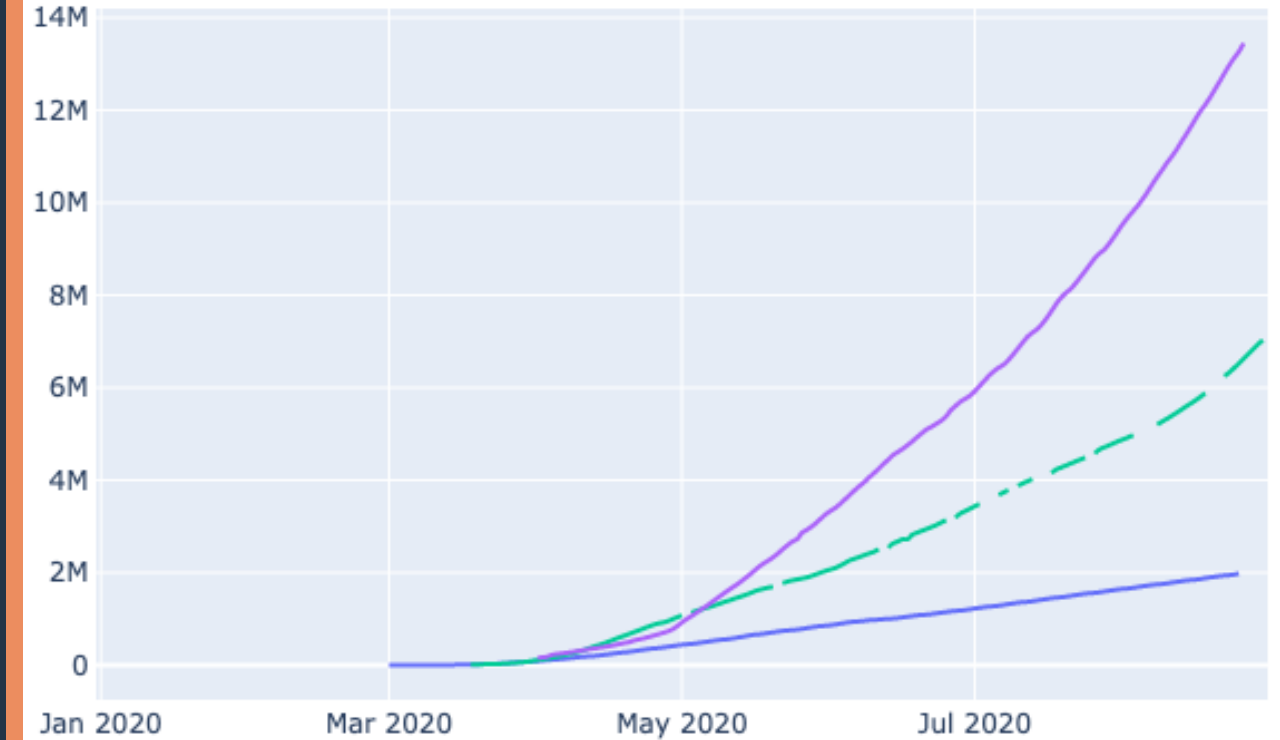
Total Cases Per Million by Date and Country - Group D



New Cases by Date and Country - Group D



Total Tests by Date and Country - Group D



GRUPO D MUERTES

Portugal



España



Reino Unido



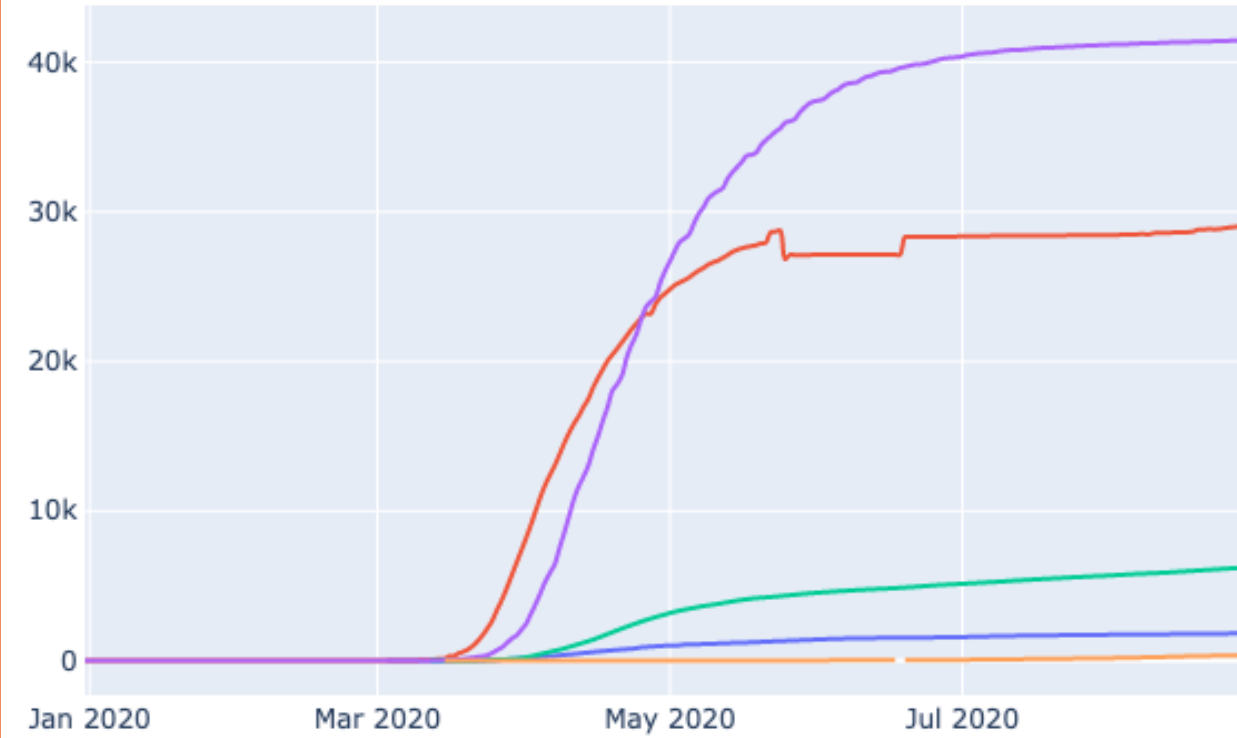
Turquía



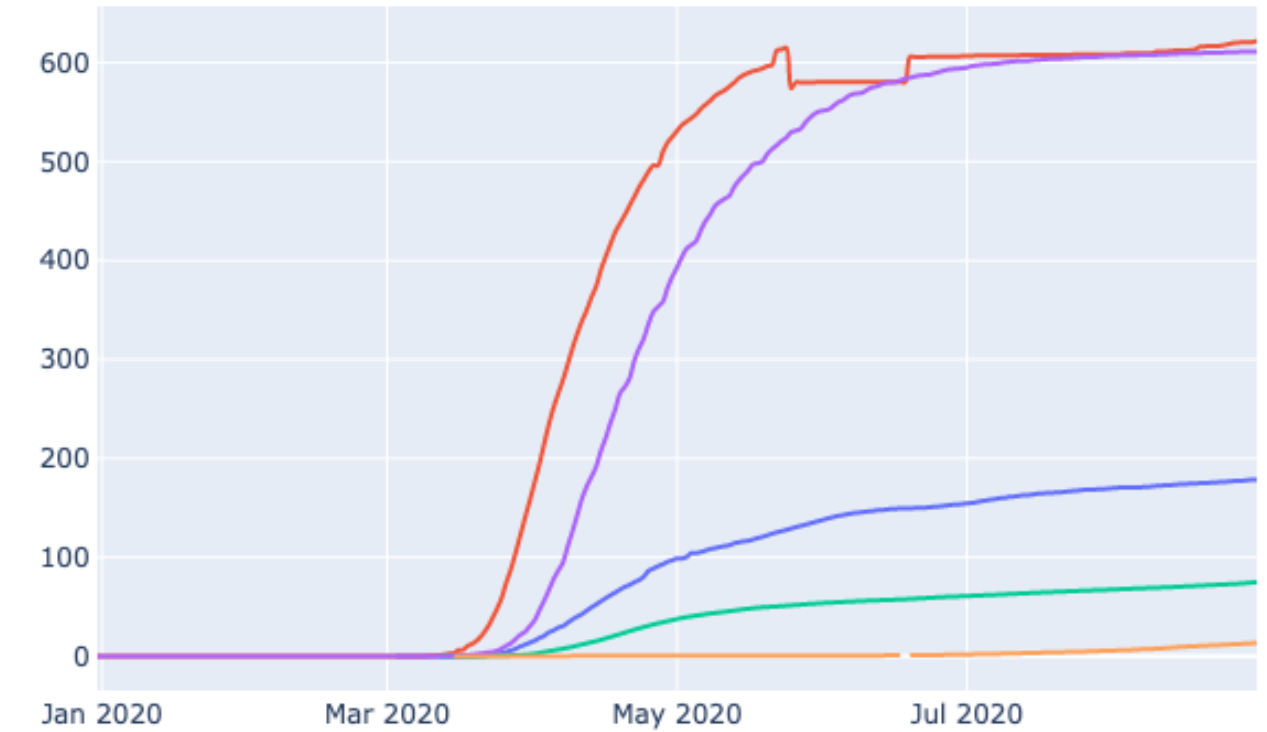
Venezuela



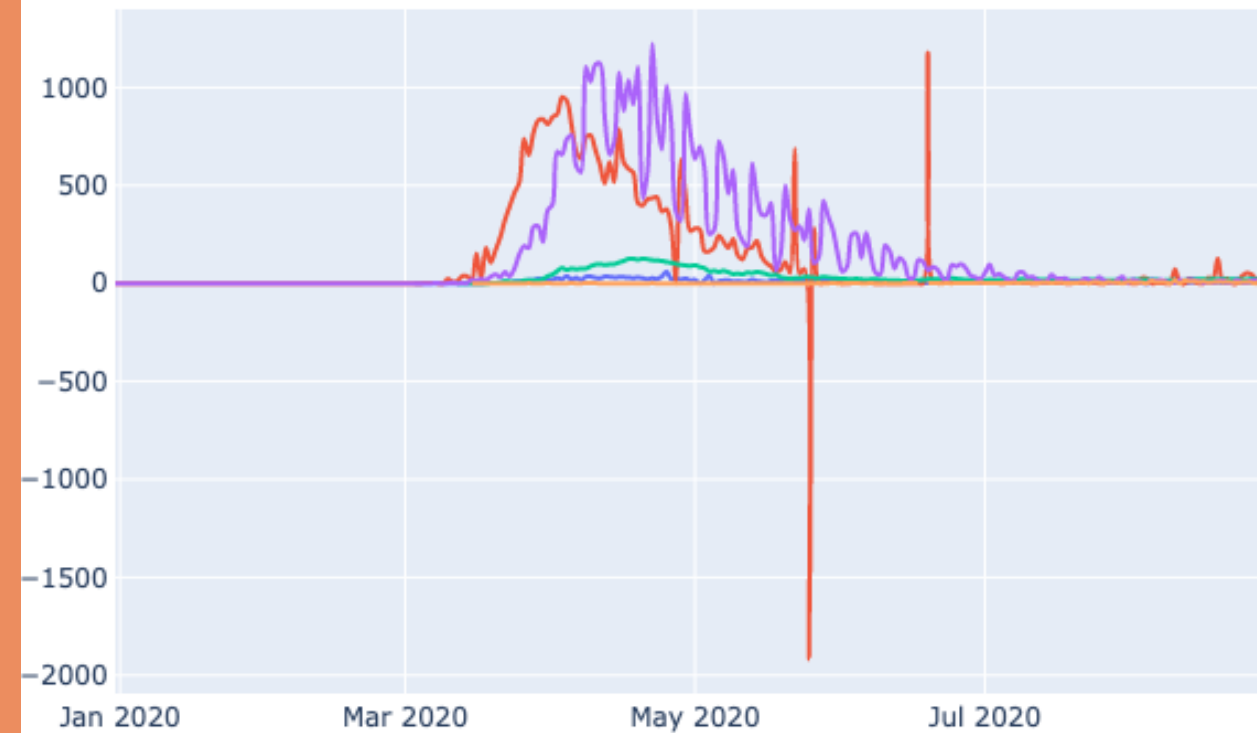
Total Deaths by Date and Country - Group D



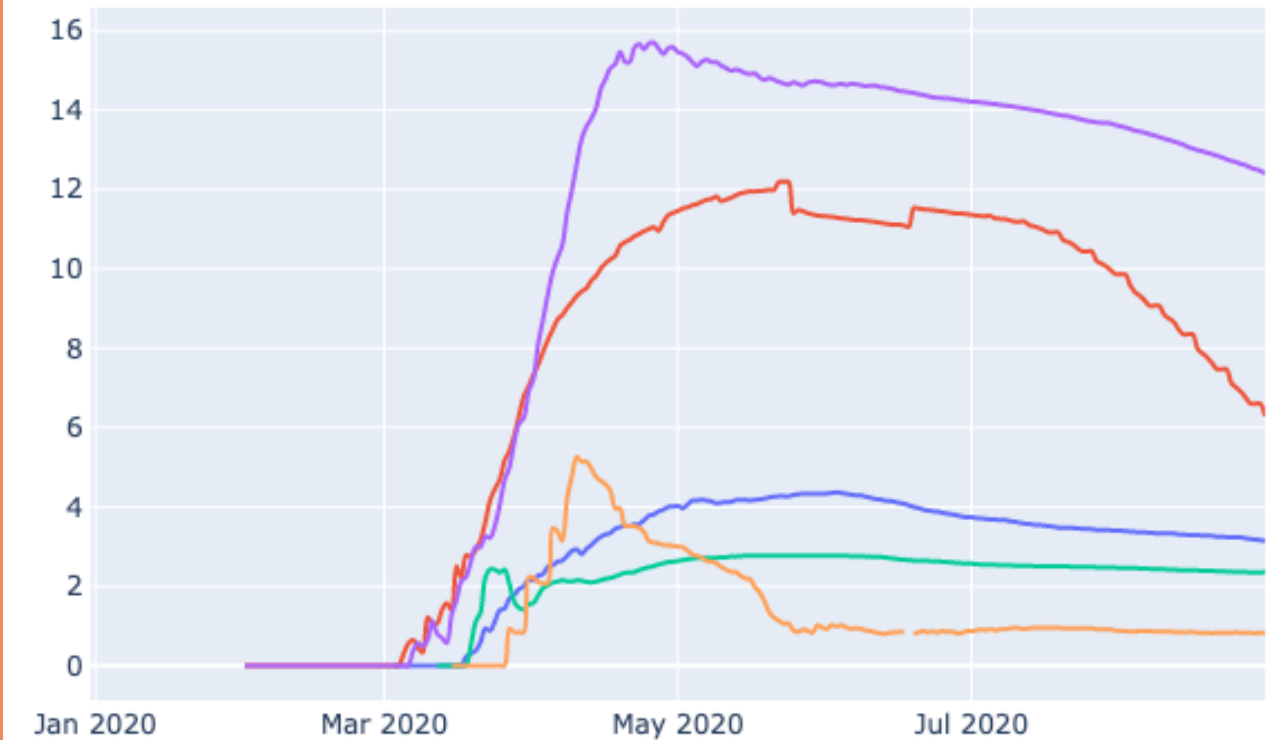
Total Deaths Per Million by Date and Country - Group D



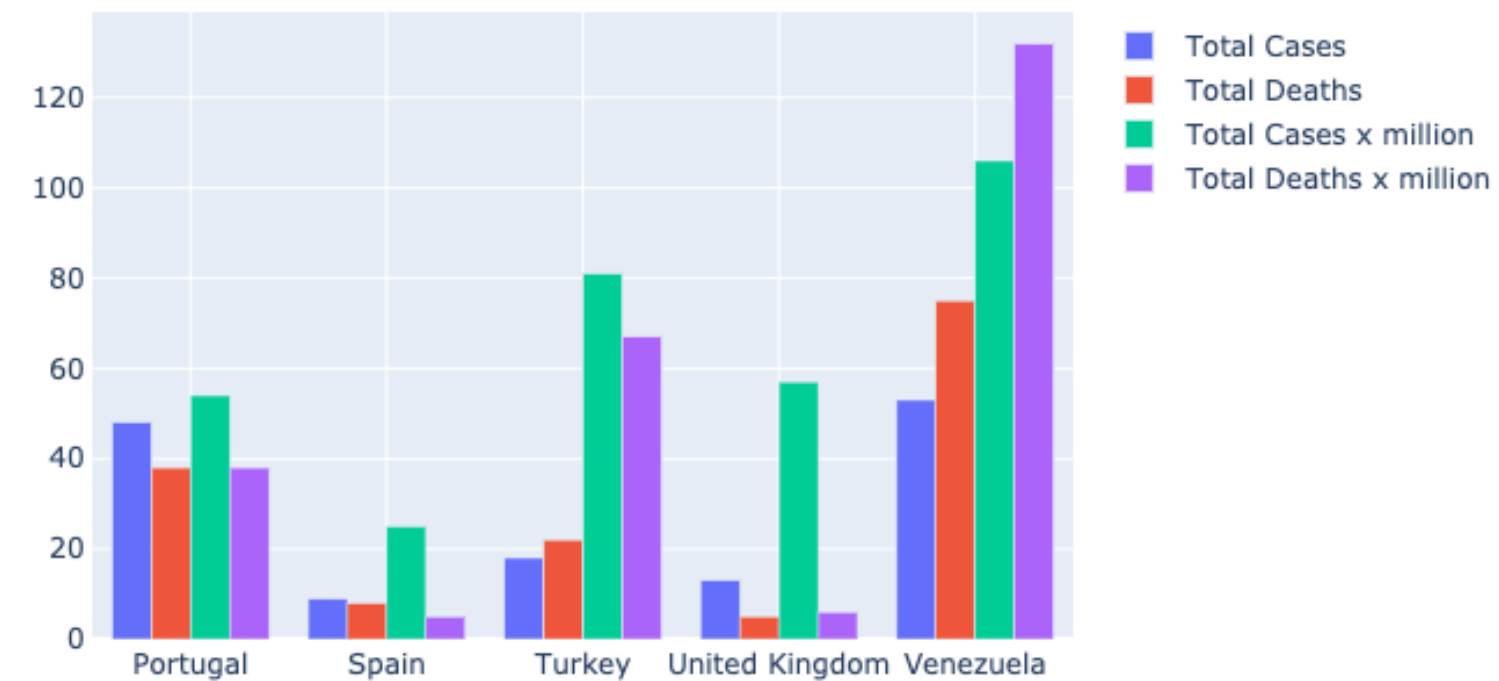
New Deaths by Date and Country - Group D



Mortality Rate by Date and Country - Group D



Group D Position World Ranking - Covid19

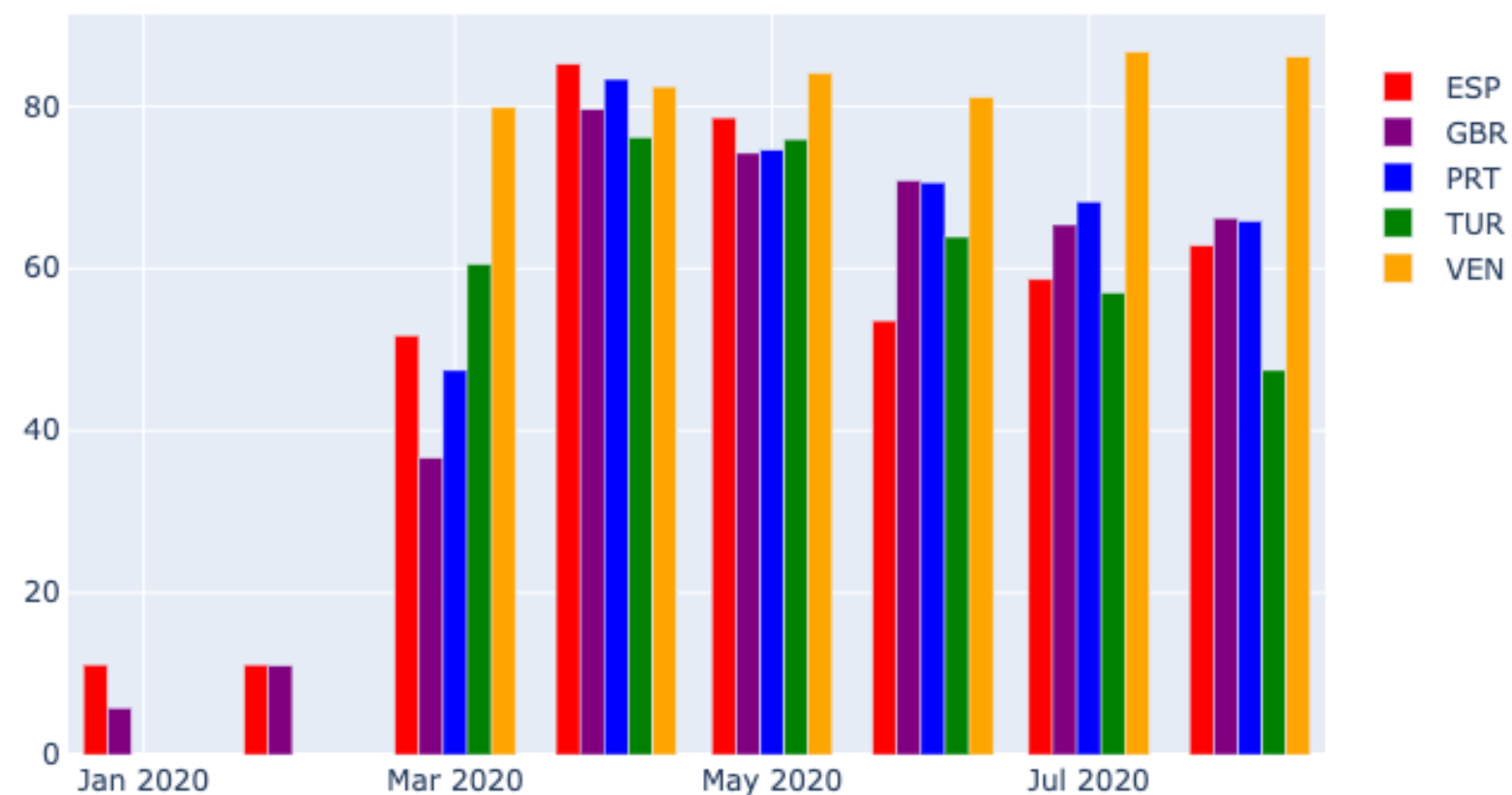


GRUPO D

POSICION RANKING

- Ranking a 21 de Agosto de 2020
- España y Gran Bretaña se mantiene en posiciones altas del ranking.
- Los países varían sus posiciones no solo por si mismo sino también influidos por el comportamiento de los demás.

Stringency Index per country (Monthly mean) - Group D



GRUPO D

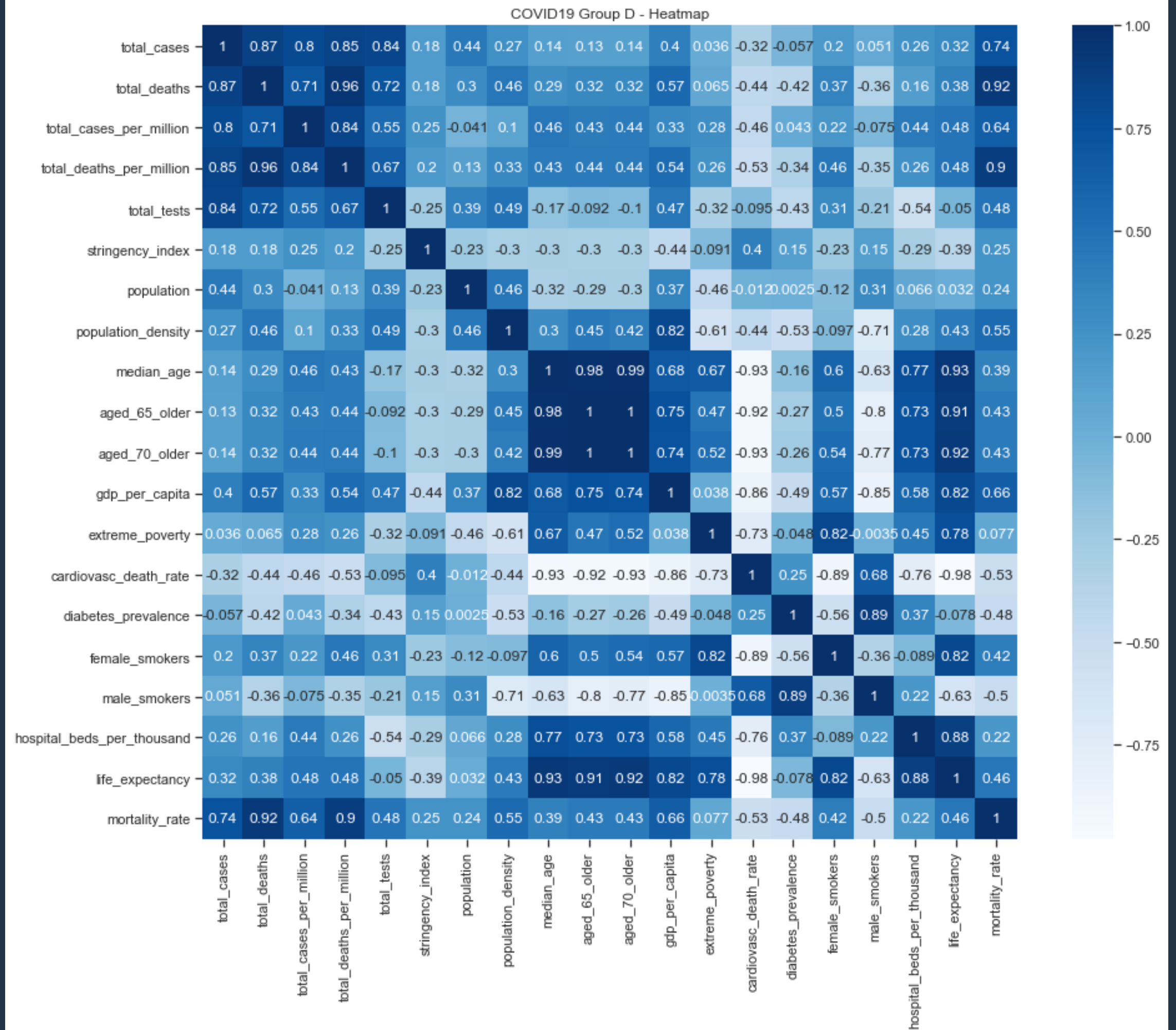
STRINGENCY INDEX

- Mide como de estrictas han sido las medidas tomadas por cada país para combatir la propagación del virus.
- Se representa la media por mes.

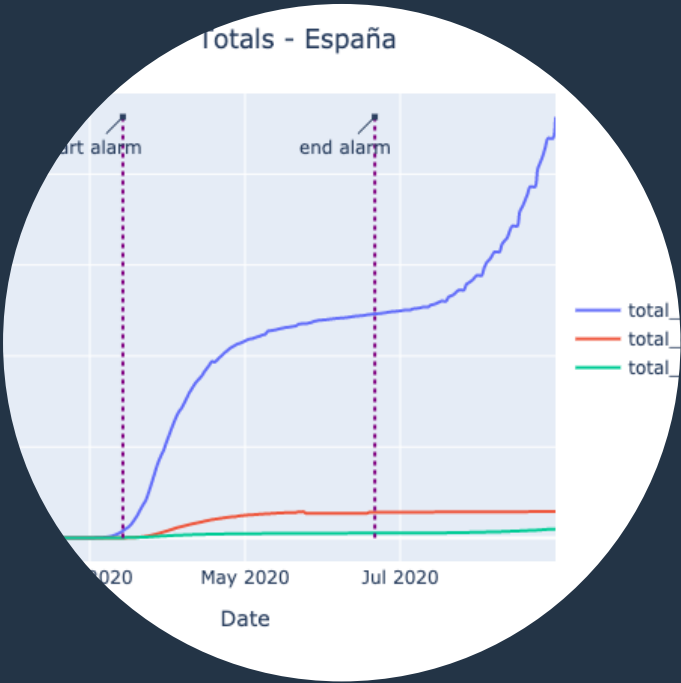
GRUPO D

CORRELACIÓN

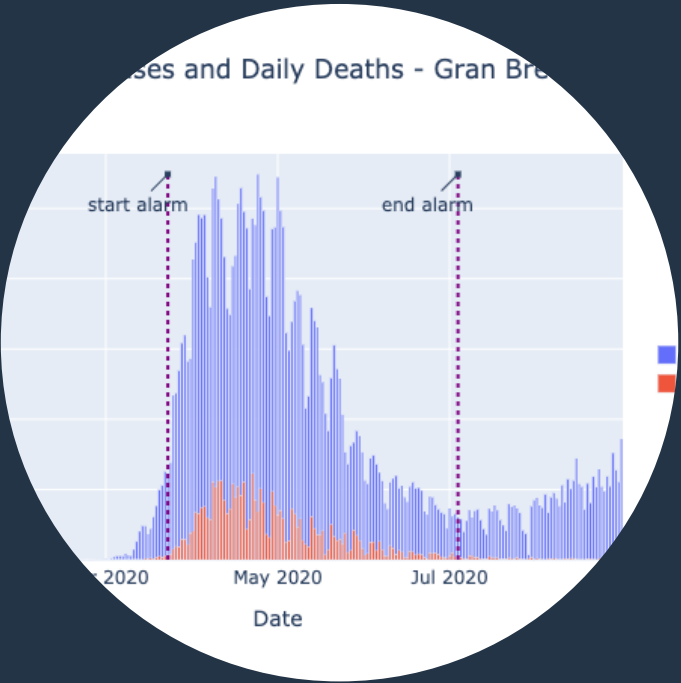
- Correlaciones altas entre variables que miden directamente la incidencia del virus (casos, muertes y test)
- Correlación alta entre test efectuados e índice de mortalidad. Mayor n° de casos y muertos junto con alto índice de mortalidad provocan mayor n° de test
- Sorprendente no correlación entre factores que a priori se puede pensar que favorecerían el aumento de los casos y los muertos (diabetes, edad,, pobreza, stringency index)
- Correlación negativa alta entre stringency index y test realizados. Sin medidas estrictas se opta por realizar mas test para luchar contra la propagación.



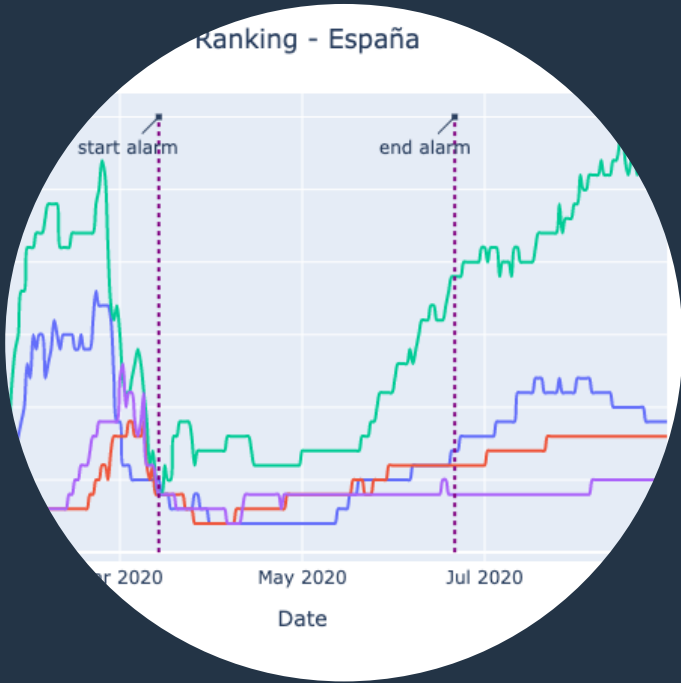
EDA por país



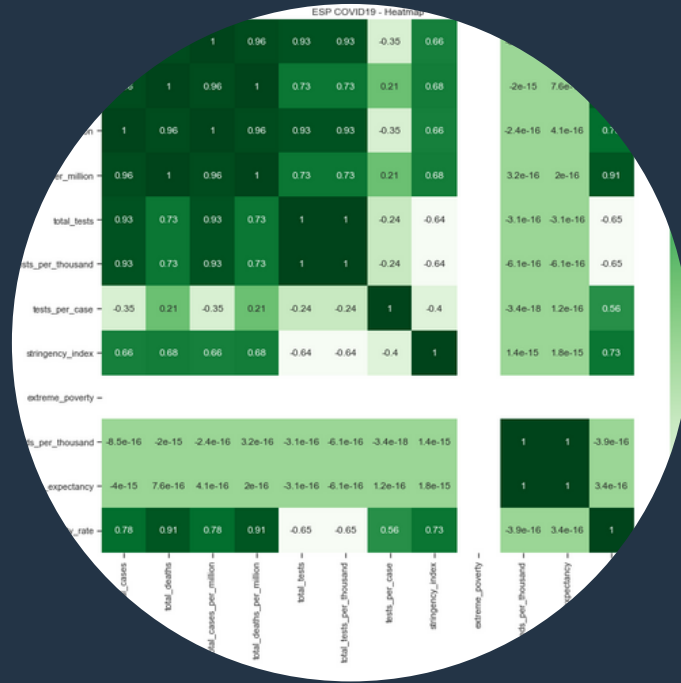
Totales



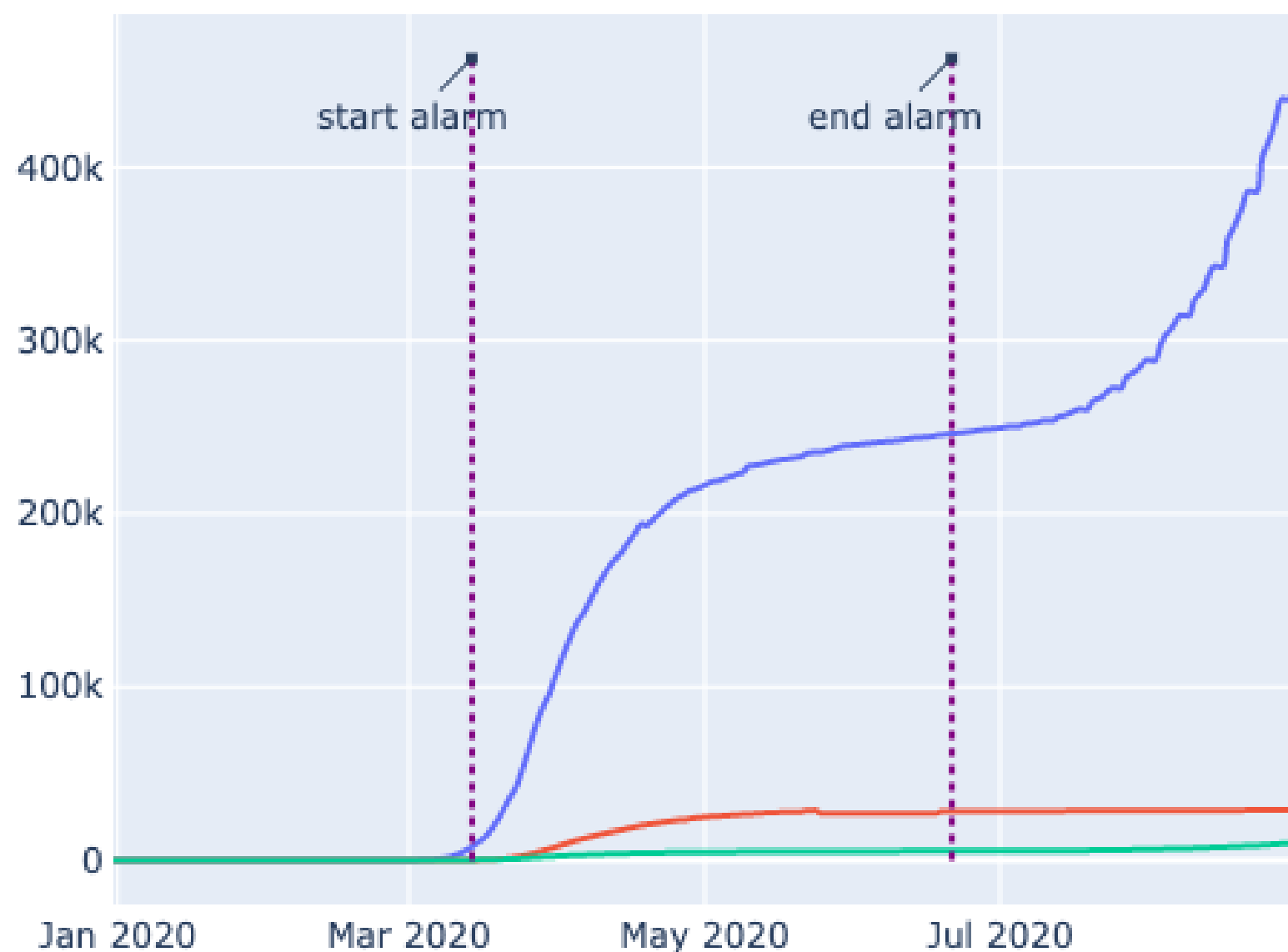
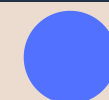
Diarias



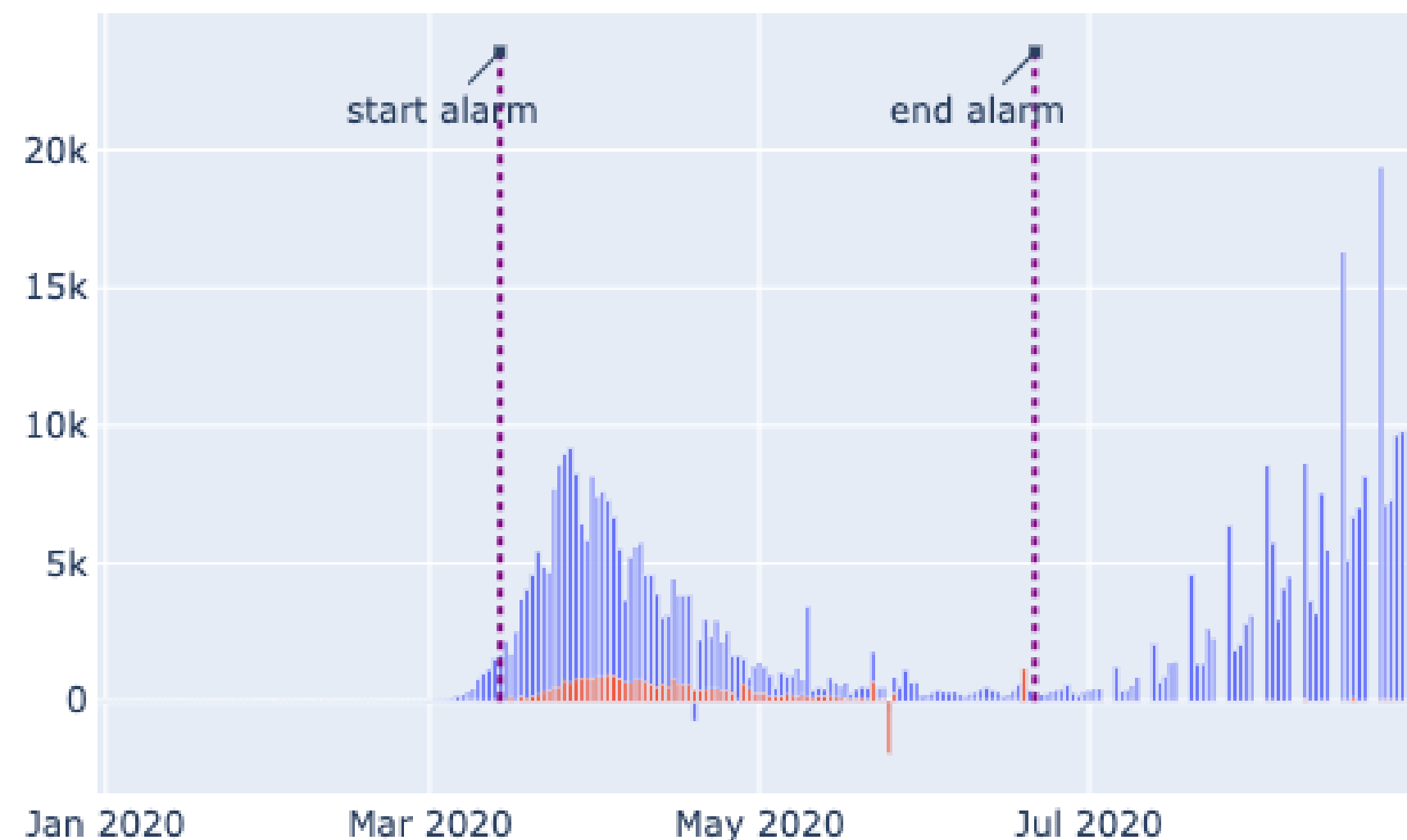
Ranking



Correlación



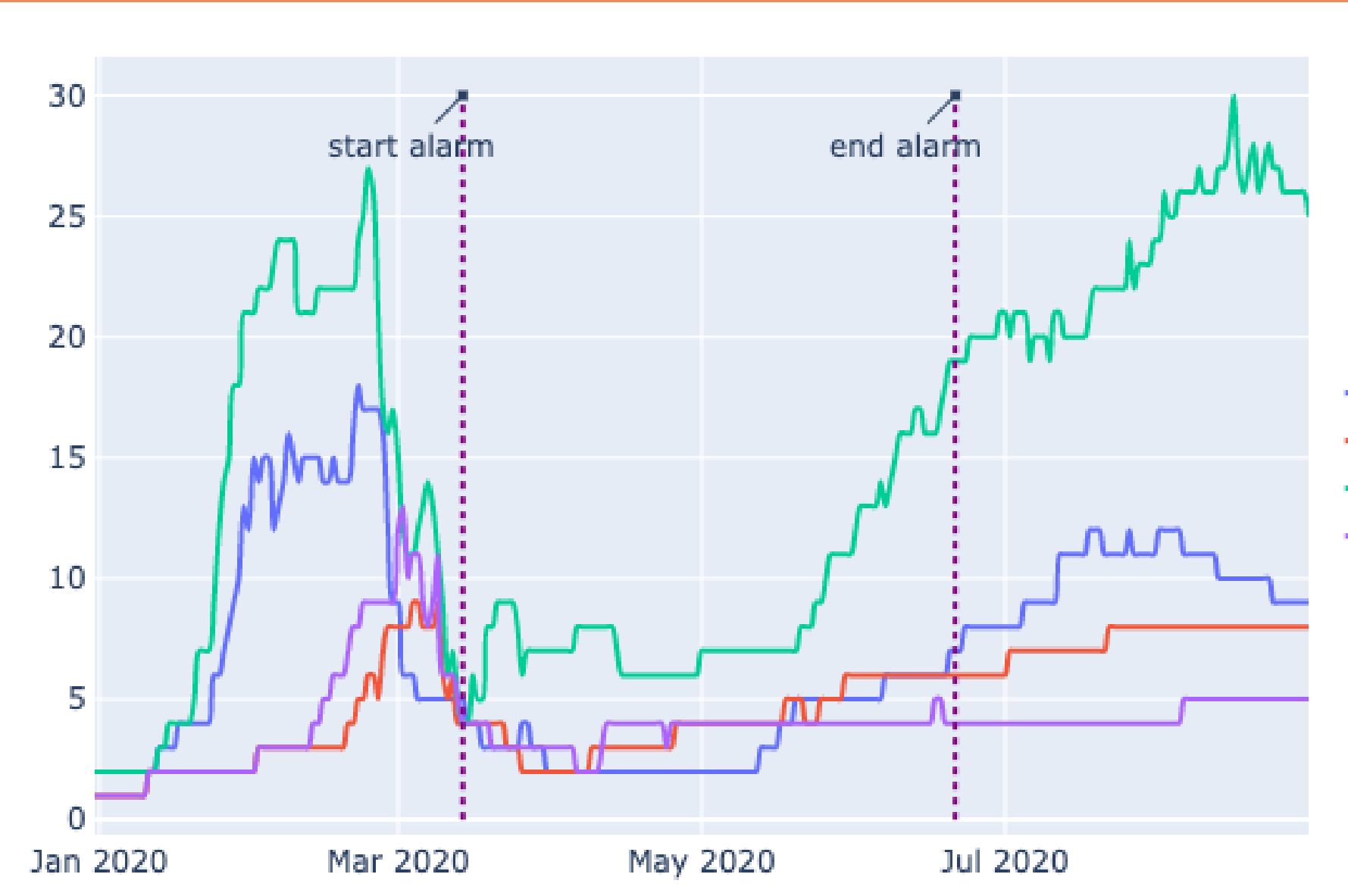
CIFRAS TOTALES



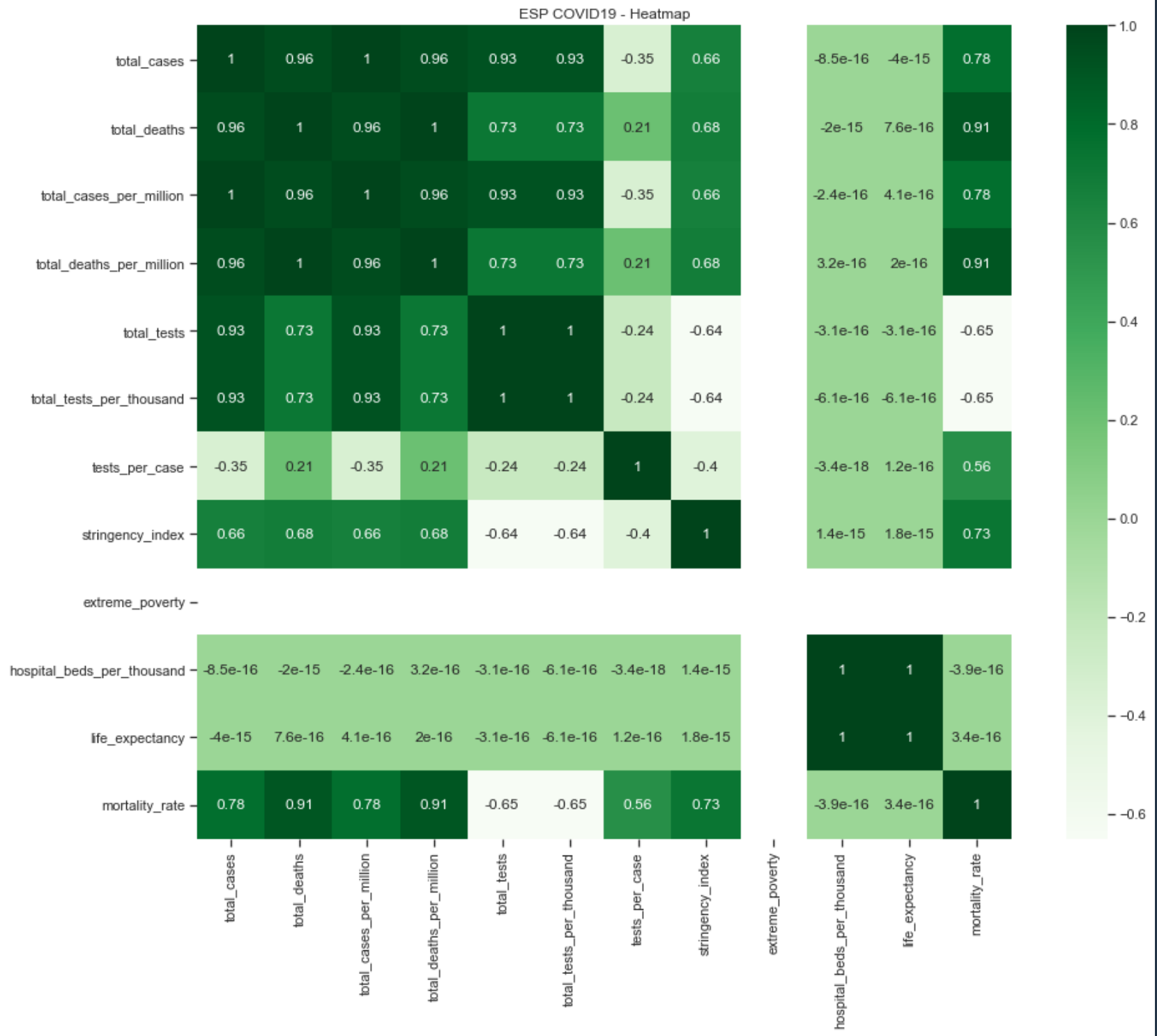
CIFRAS DIARIAS

Ajuste de datos. Casos el 19/4, el 10/5 y el 22/5.
Muertes el 22/5, 25/5 y el 19/6.

ESPAÑA



RANKING



CORRELACIÓN

Correlaciones positivas altas entre todos los tipos de variables de casos y muertos. También el índice de mortalidad con ellos.

PORTUGAL

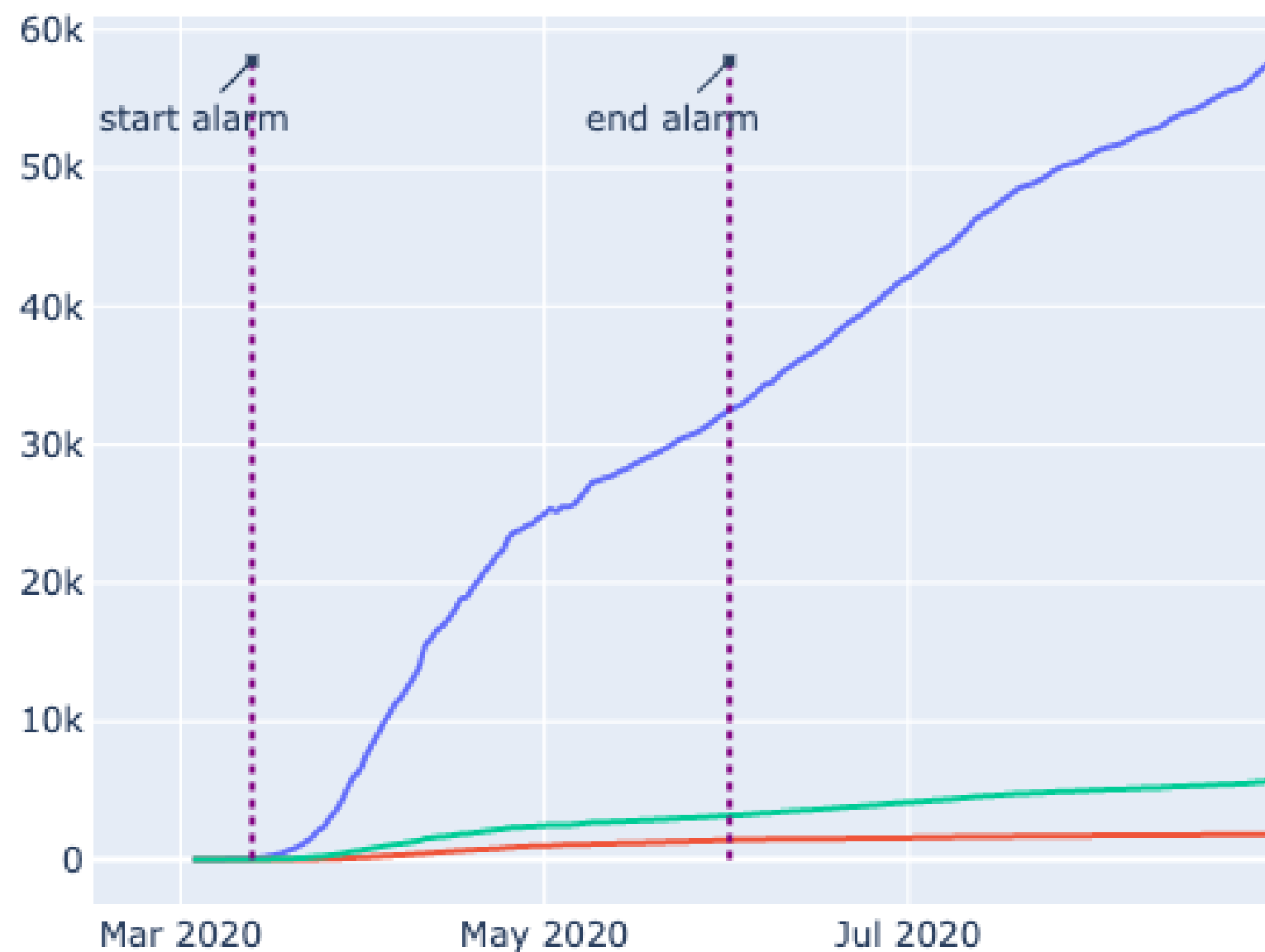
Muertes



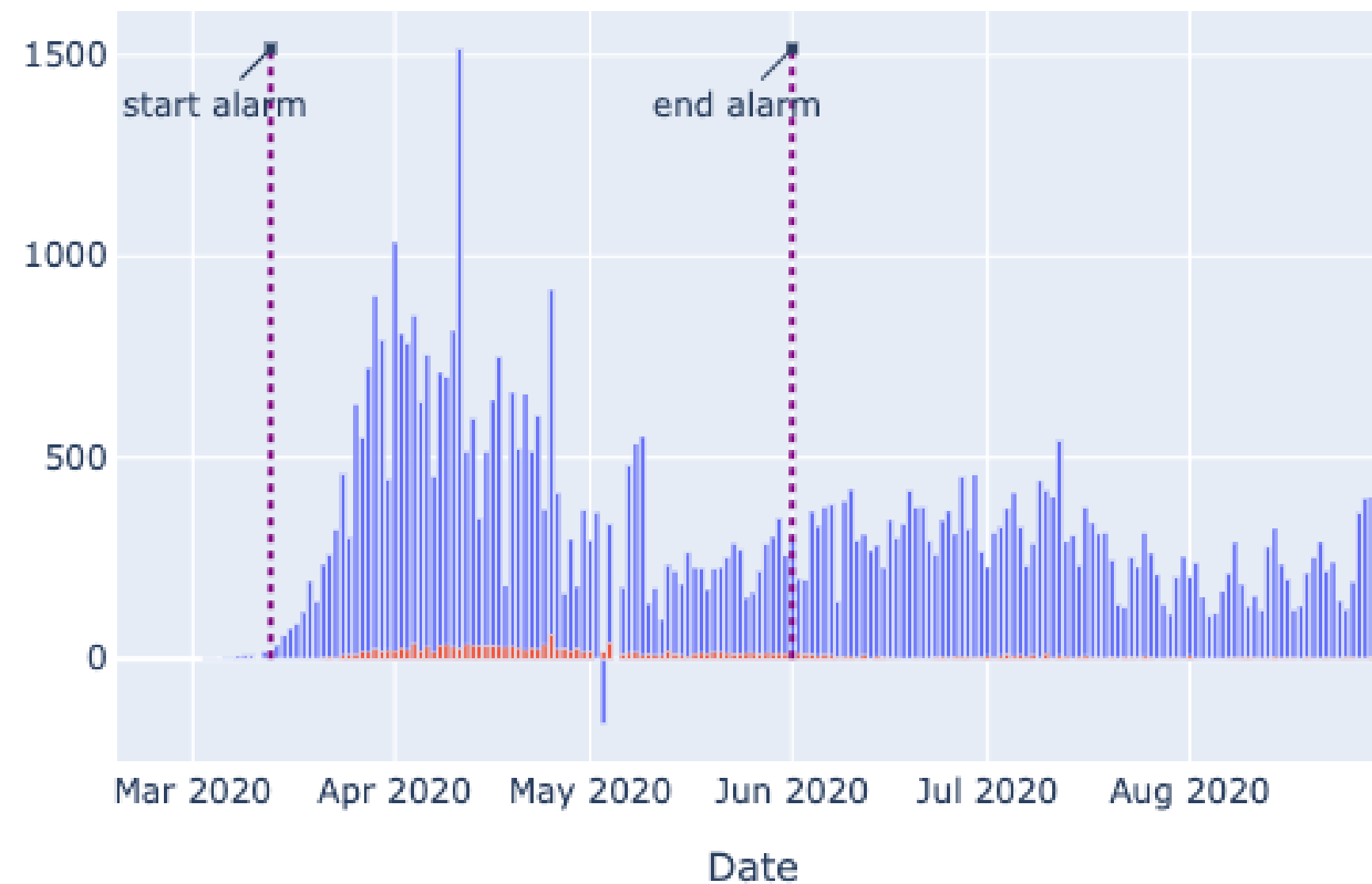
Casos



Casos por millón



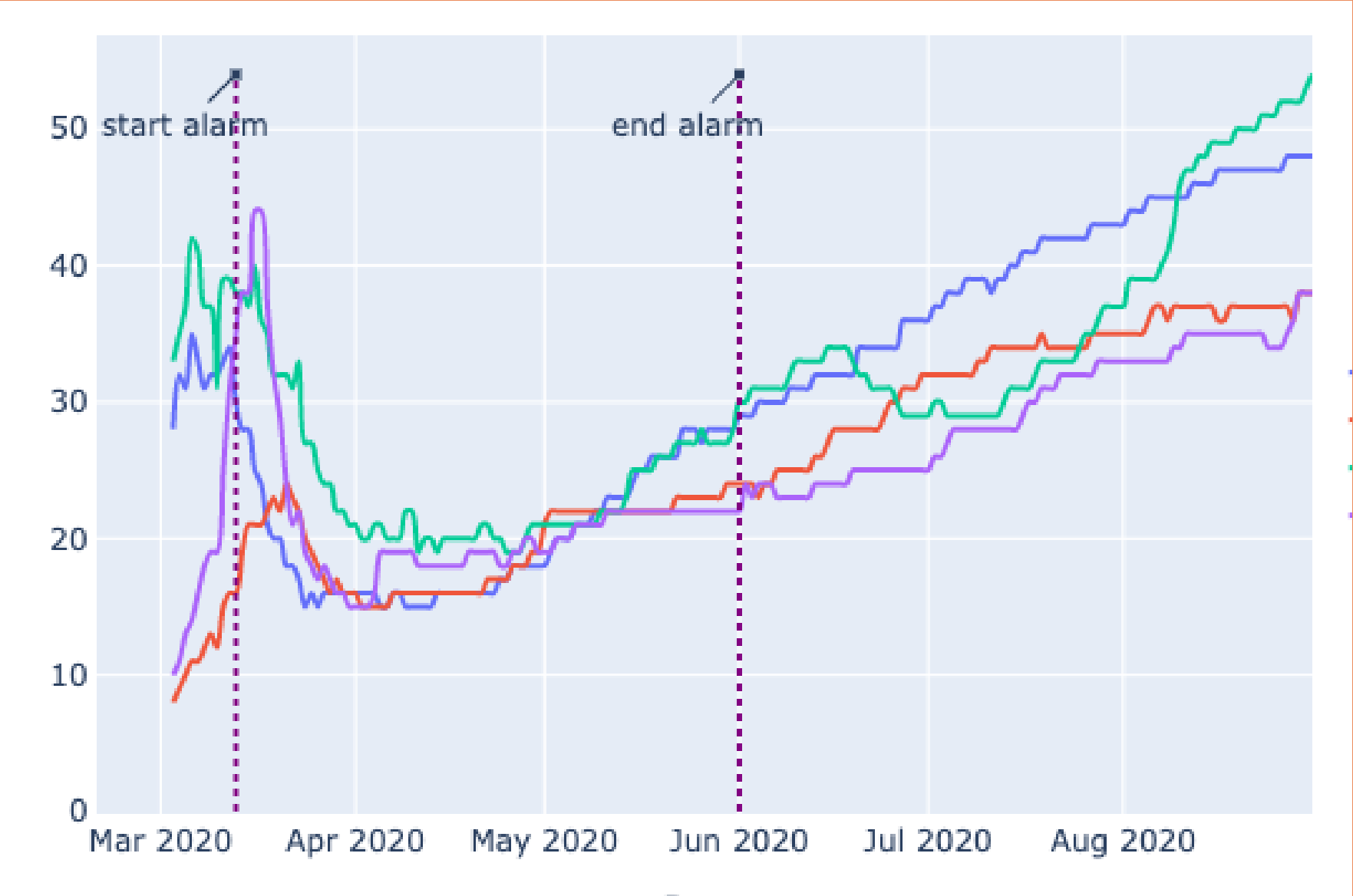
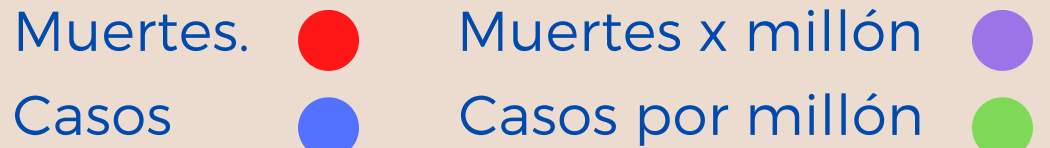
CIFRAS TOTALES



CIFRAS DIARIAS

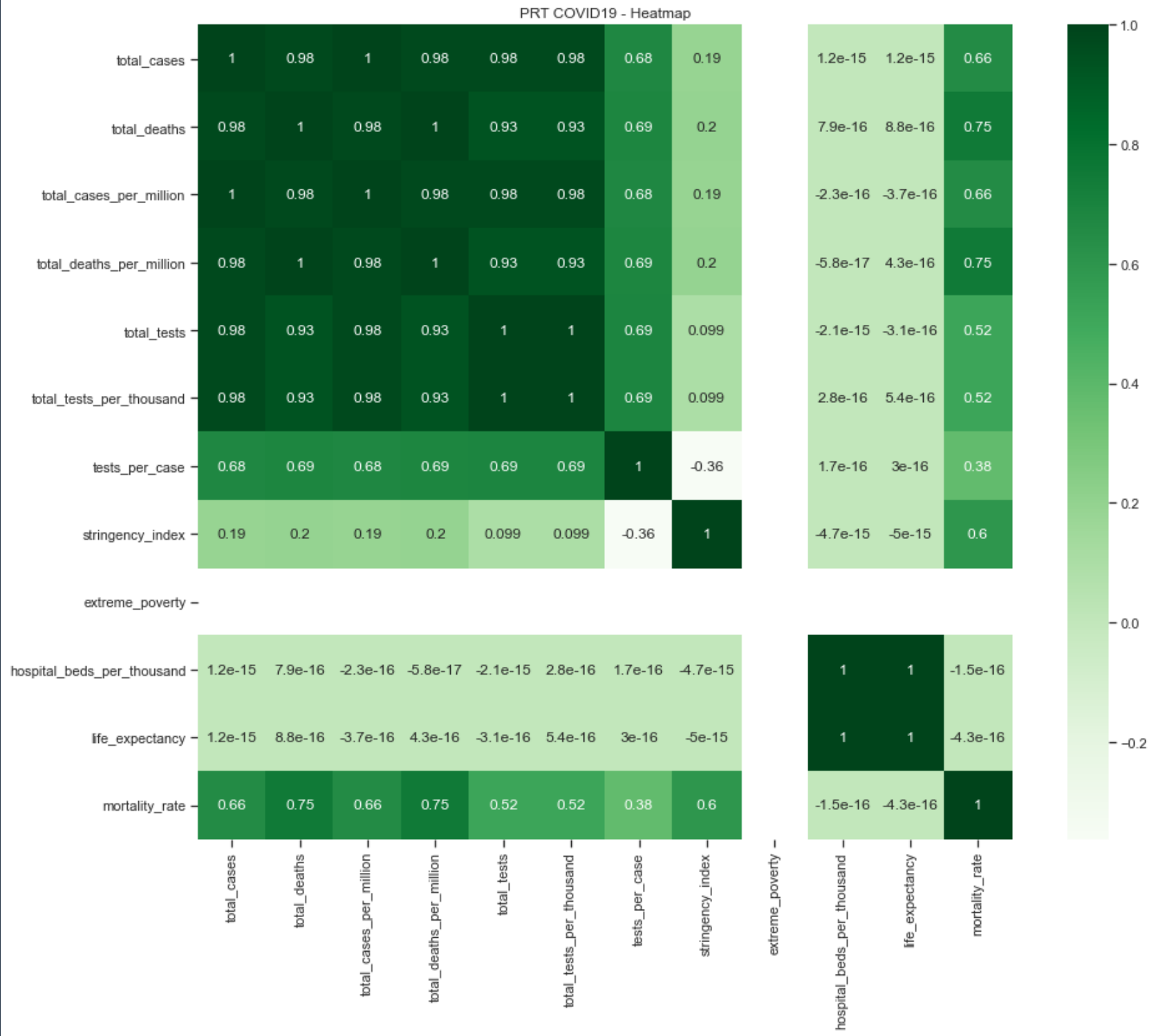
Ajuste de datos. Casos el 19/4 y el 10/5.
Muertes el 25/5 y el 19/6.

PORTUGAL



RANKING

Presentations are communication tools that can be used as demonstrations.



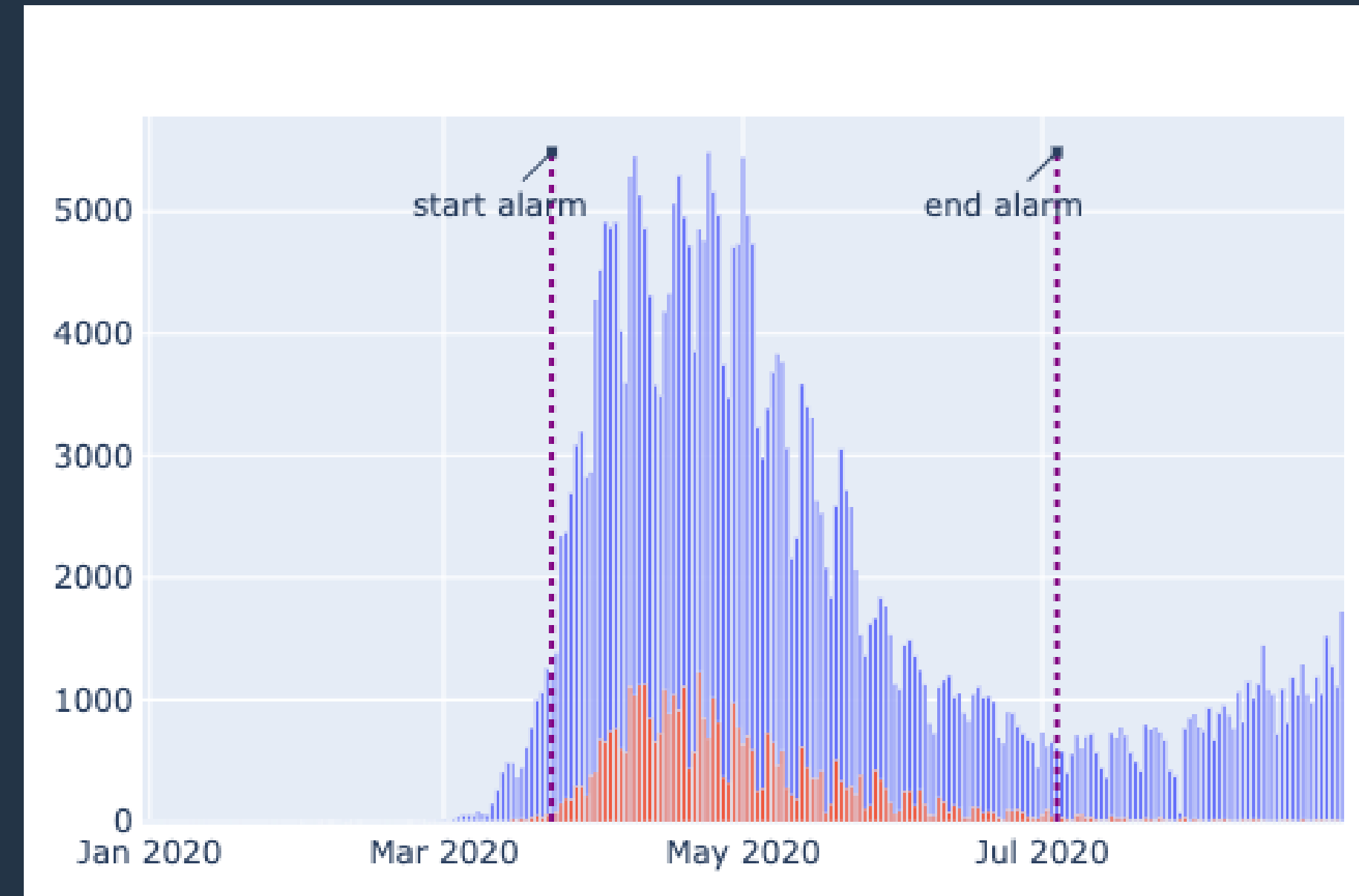
CORRELACIÓN

Correlaciones positivas muy altas entre todos los tipos de variables de casos y muertos. Tambien el índice de mortalidad con ellos



CIFRAS TOTALES

Presentations are communication tools that can be used as demonstrations.

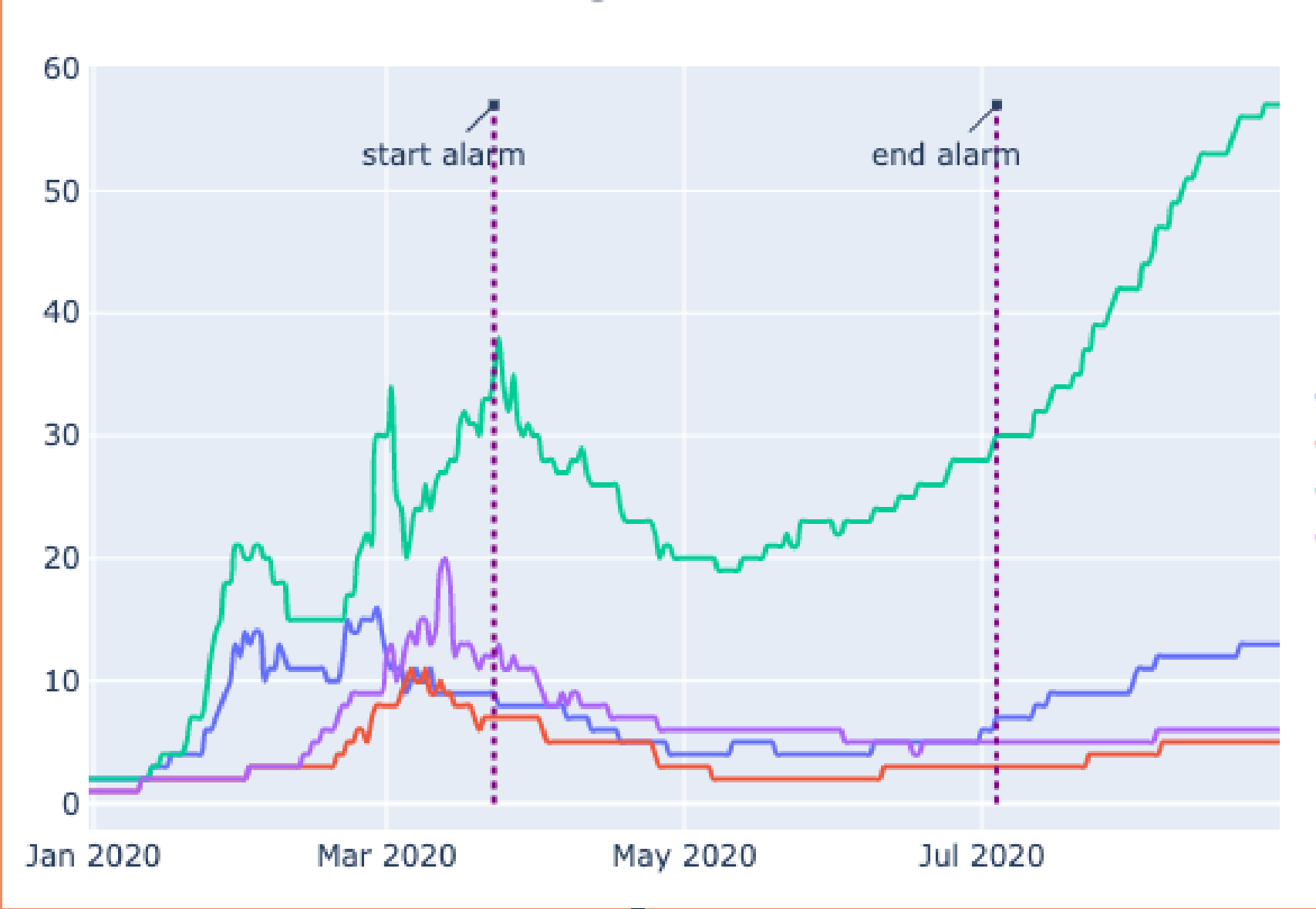


CIFRAS DIARIAS

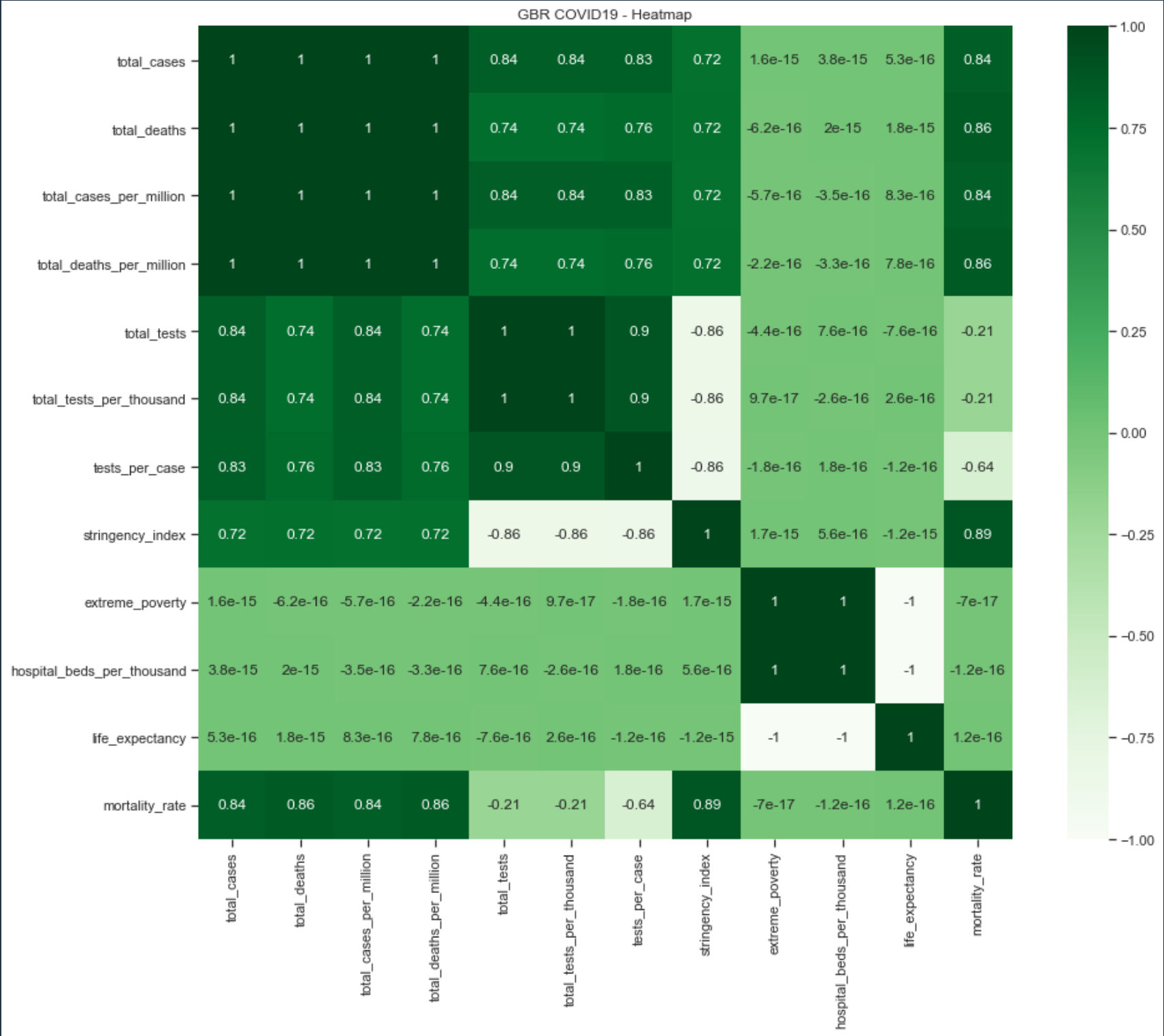
Casos y muertes diarias

REINO UNIDO

Muertes. ● Muertes x millón ●
Casos ● Casos por millón ●



RANKING



CORRELACIÓN

Índice -1 entre camas de hospital y esperanza de vida. El resto de correlaciones encontrados en otros países se repiten en GB

VENEZUELA

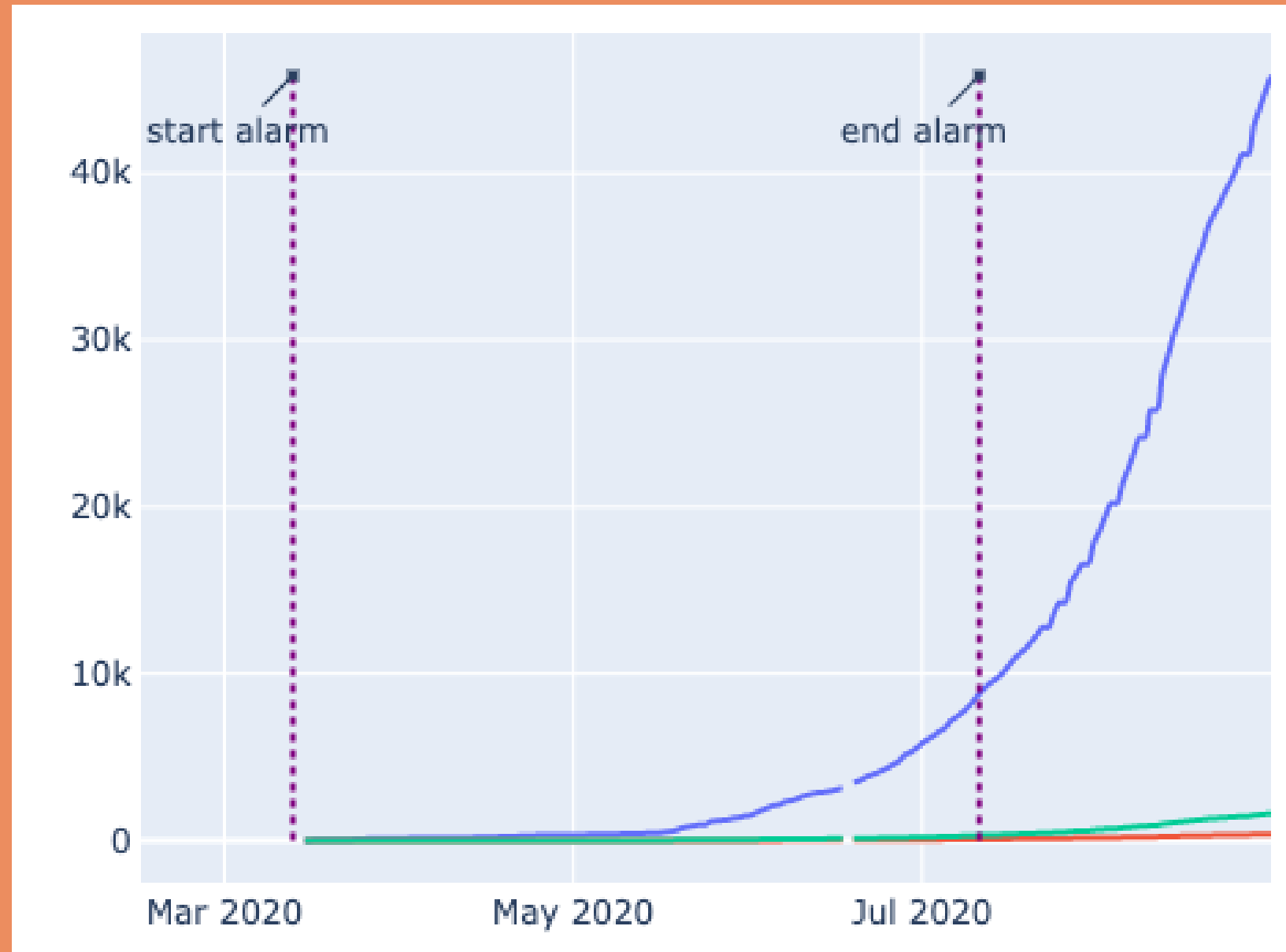
Muertes



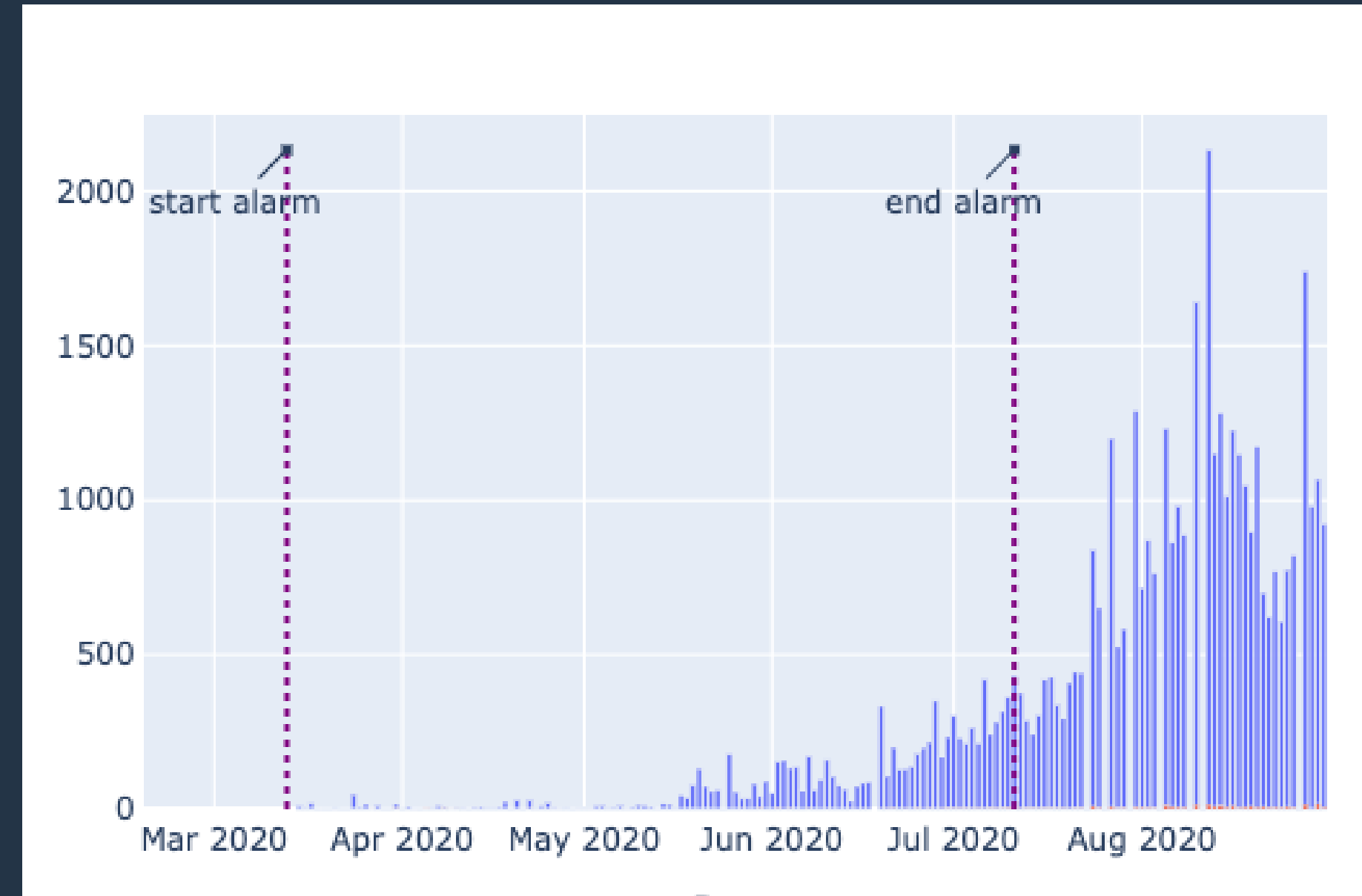
Casos



Casos por millón



CIFRAS TOTALES

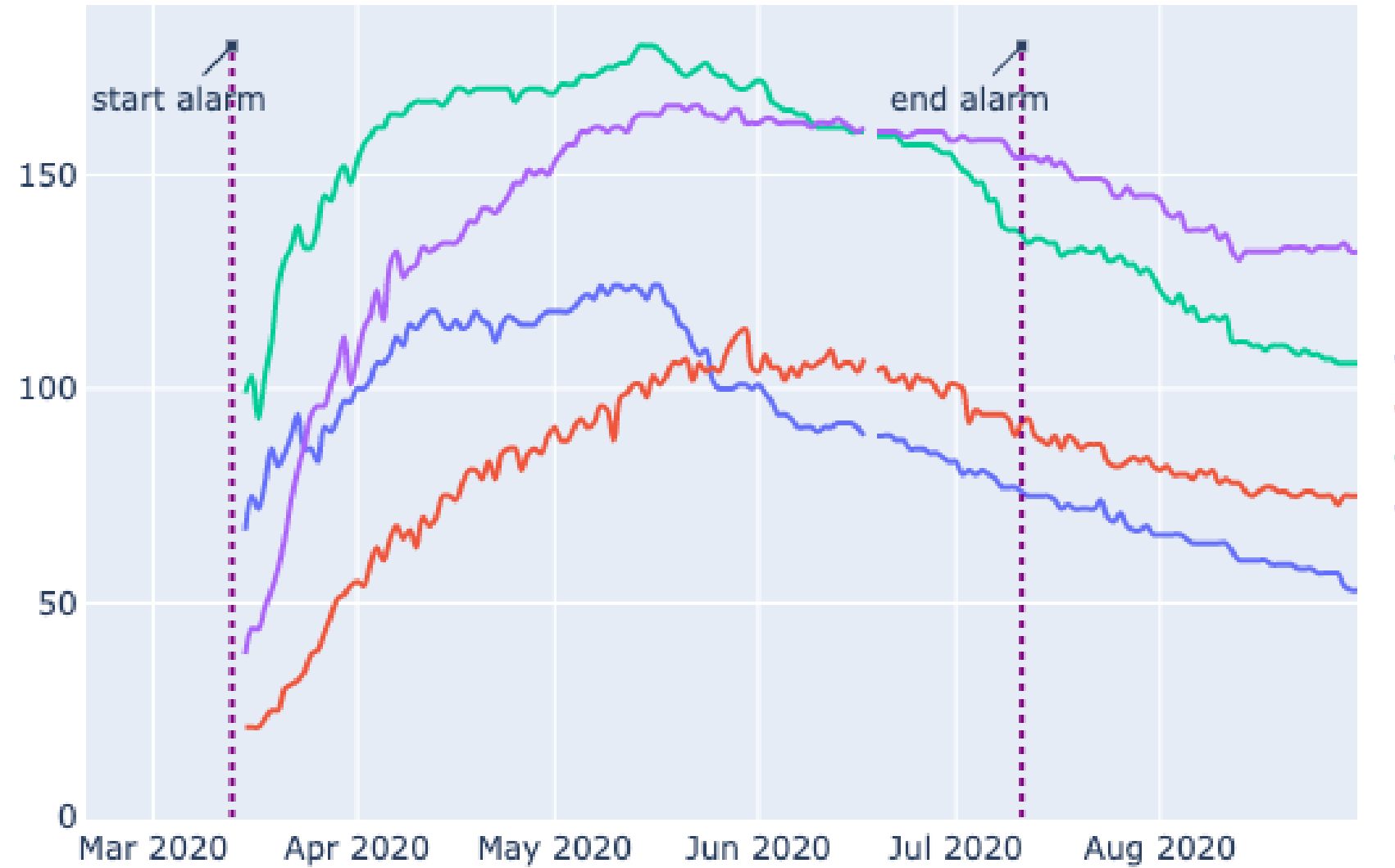


CIFRAS DIARIAS

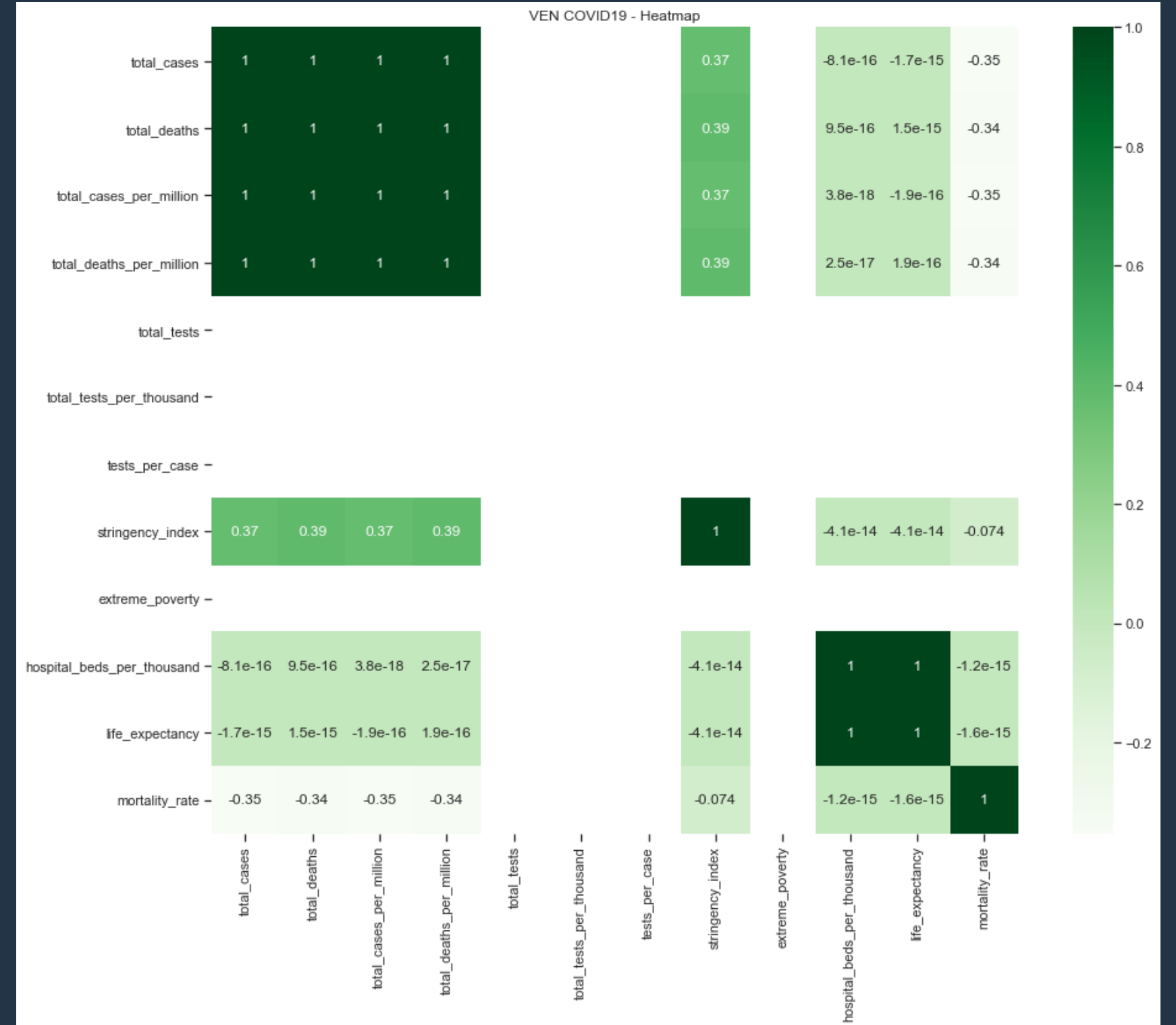
Casos y muertes diarias

VENEZUELA

Muertes.	●	Muertes x millón	●
Casos	●	Casos por millón	●



RANKING



CORRELACIÖN

Al no proporcionar datos de alguna de las columnas estas aparecen vacías

TURQUÍA

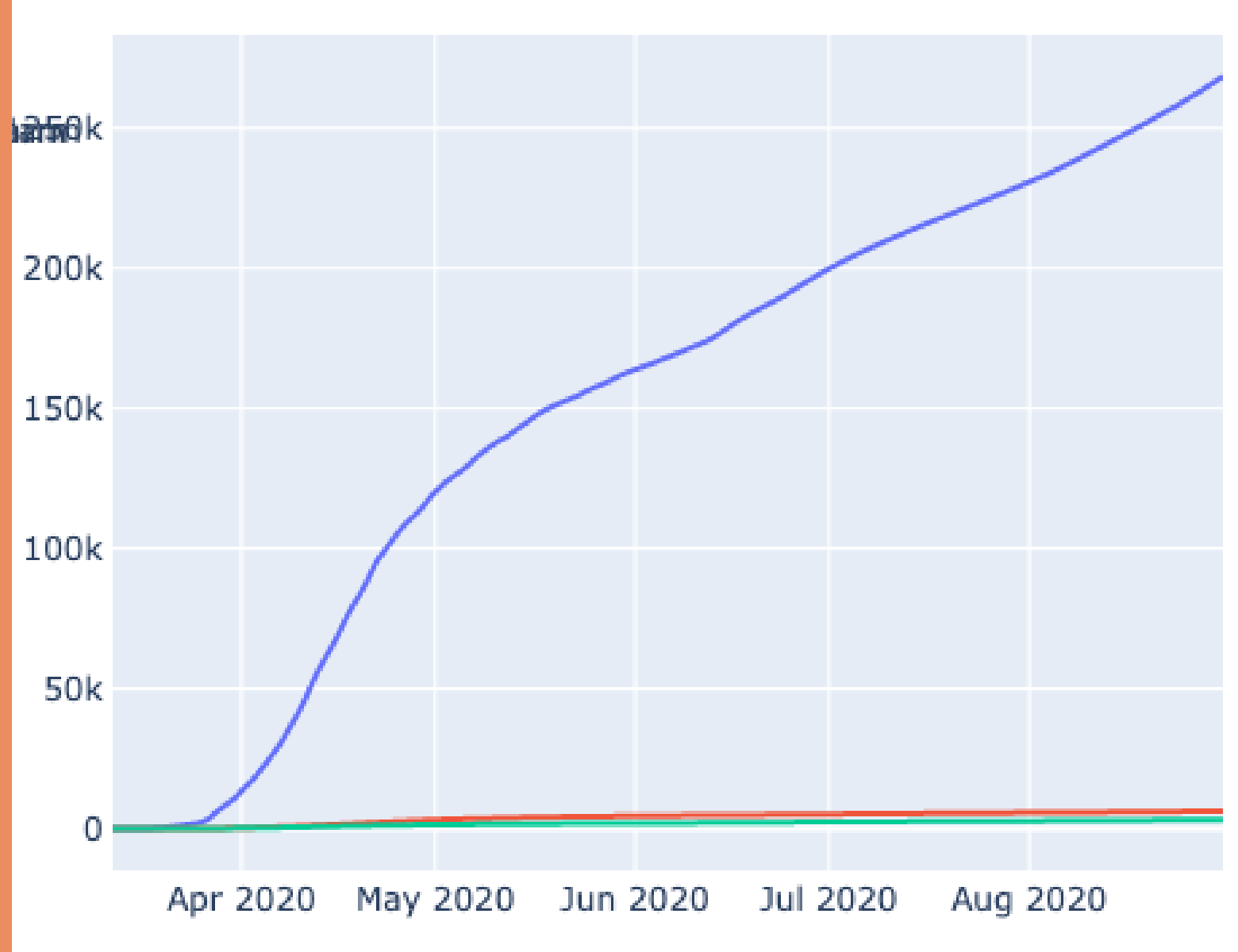
Muertes



Casos

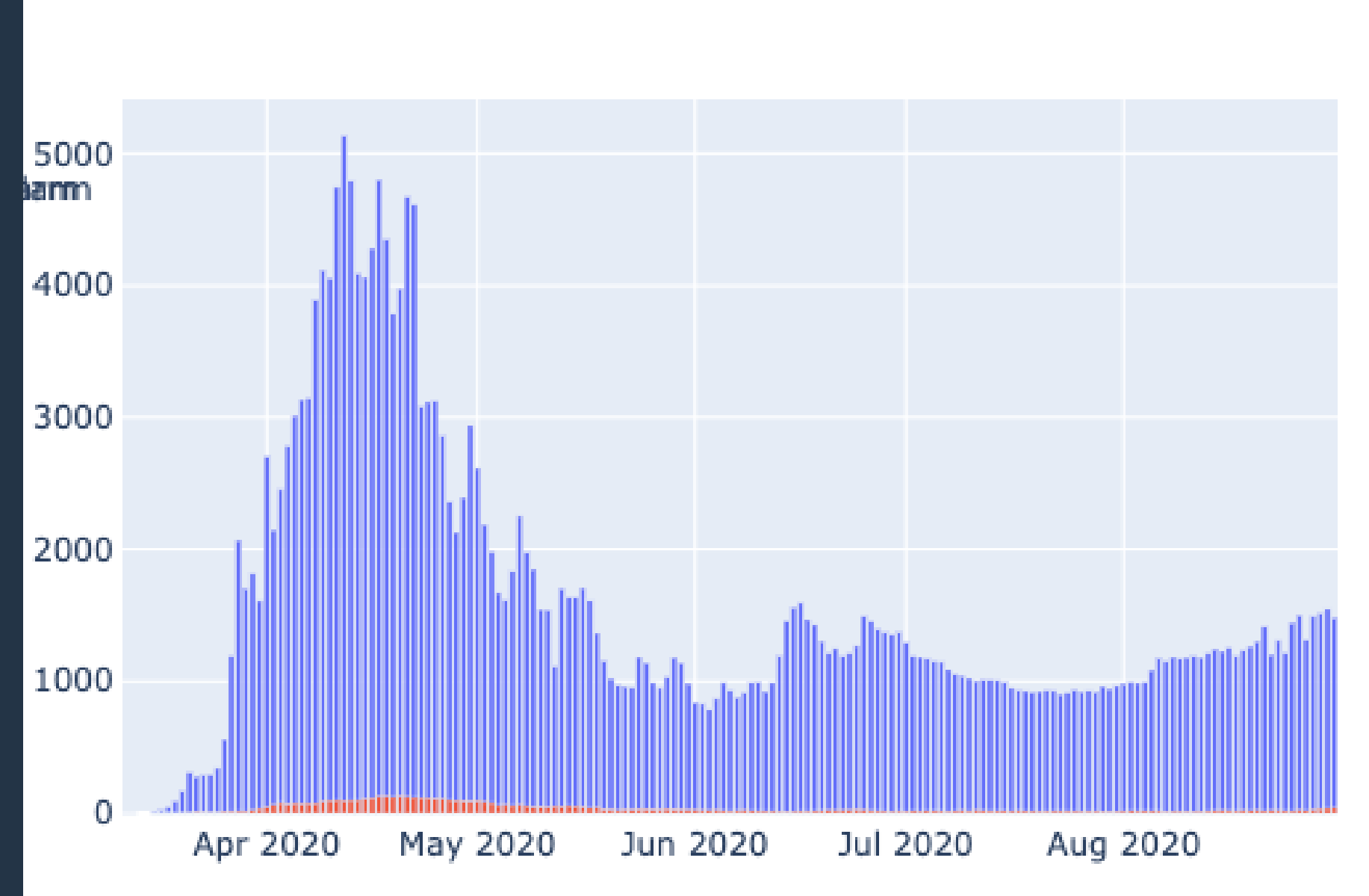


Casos por millón



CIFRAS TOTALES

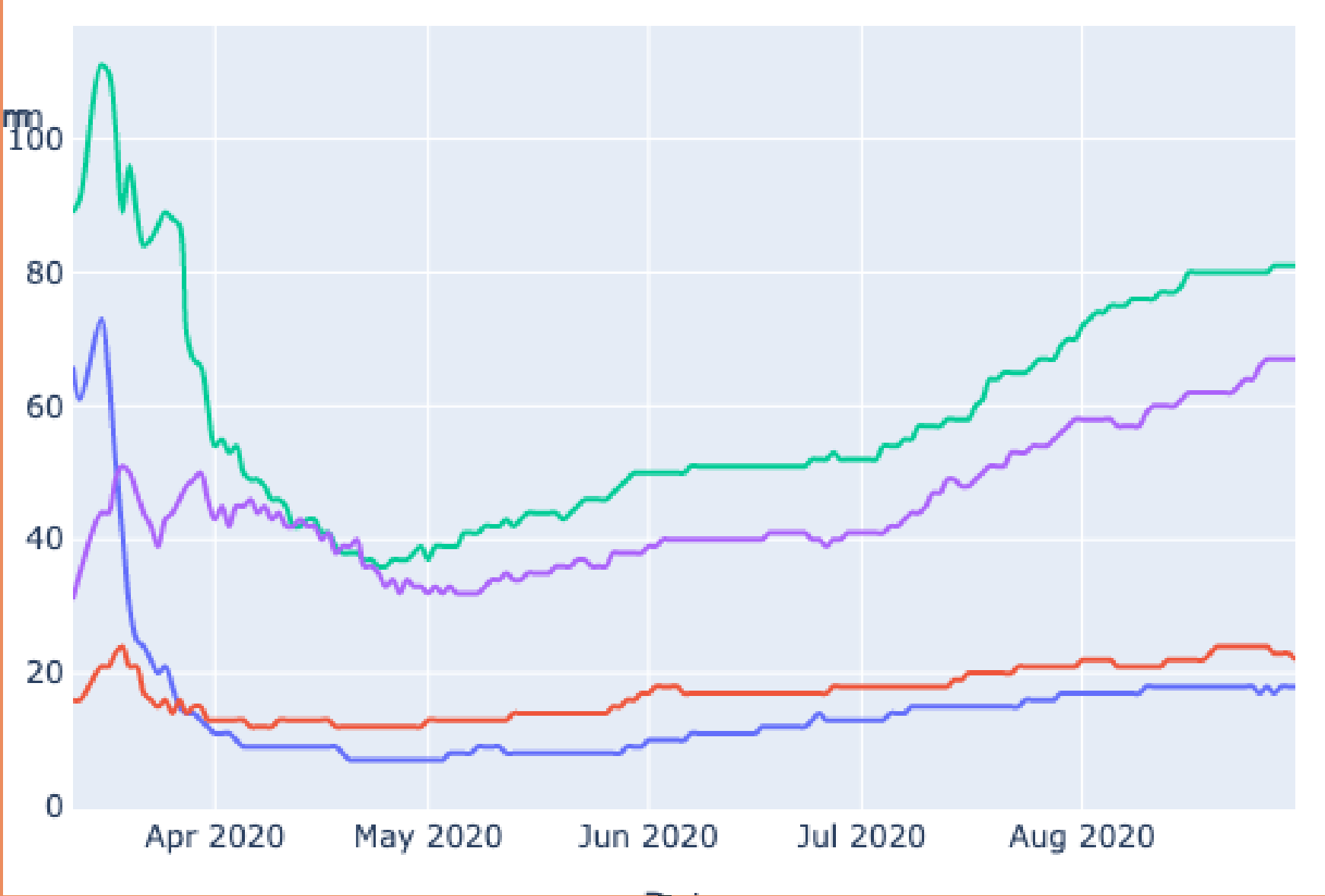
Turquia no decretó el estado de alarma.



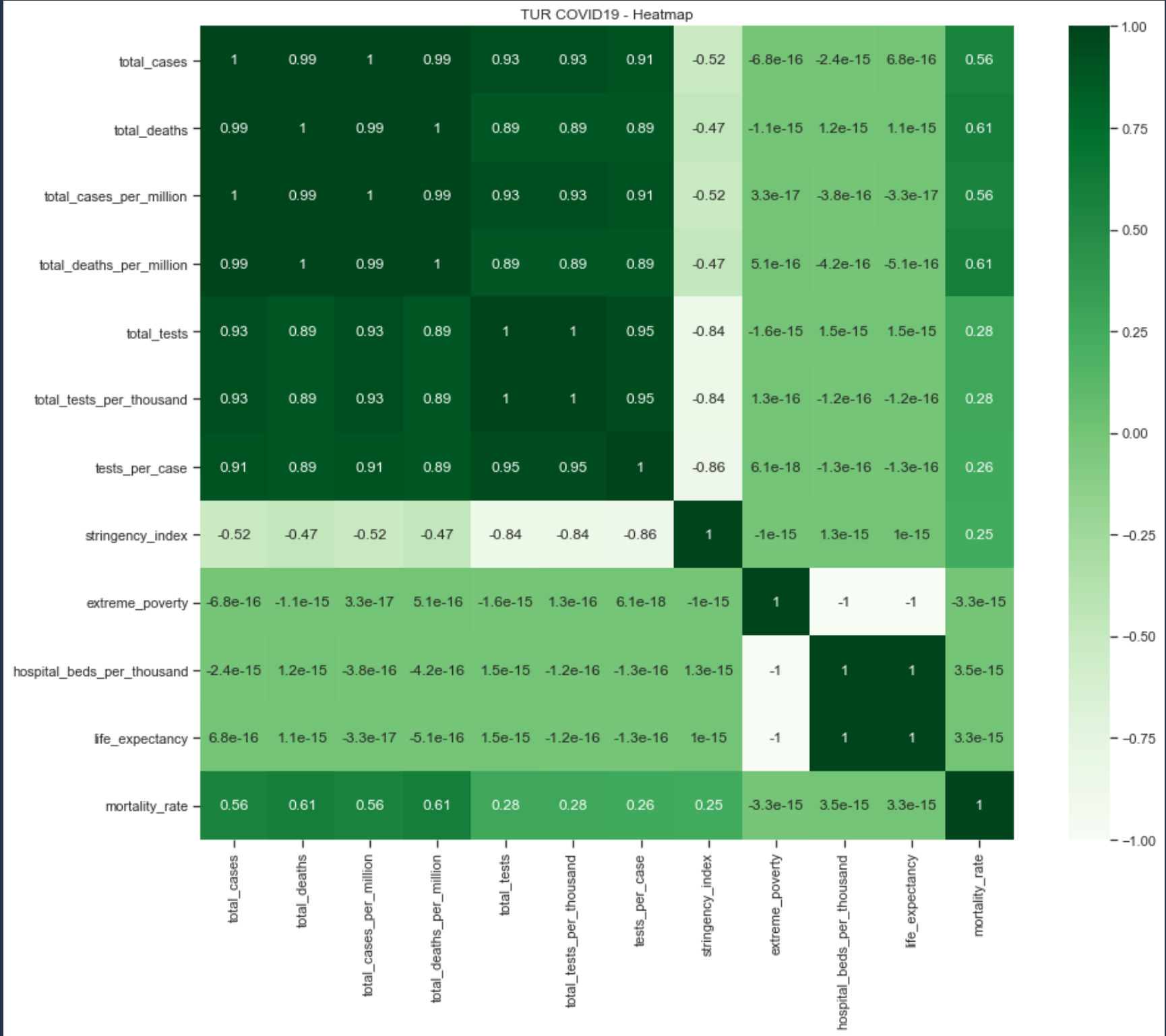
CIFRAS DIARIAS

Casos y muertes diarias

TURQUÍA



RANKING

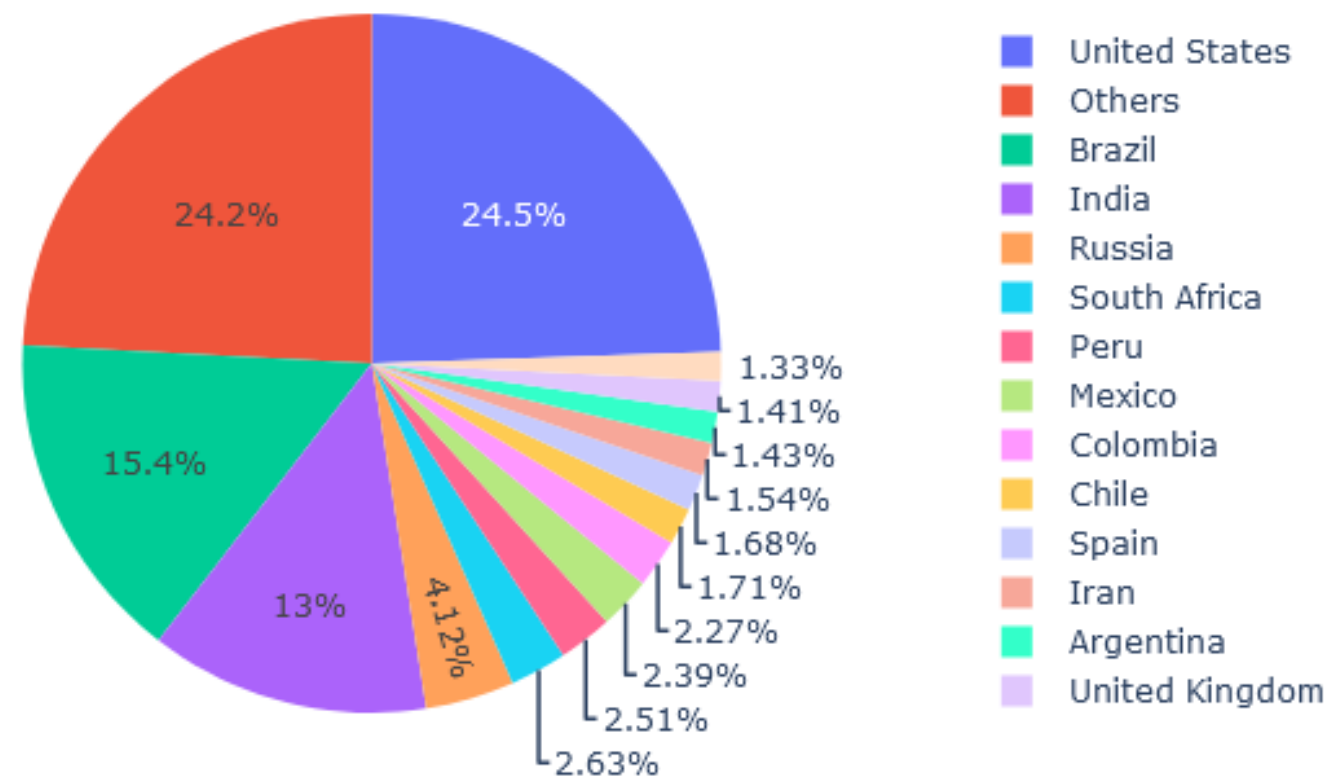


CORRELACIÓN

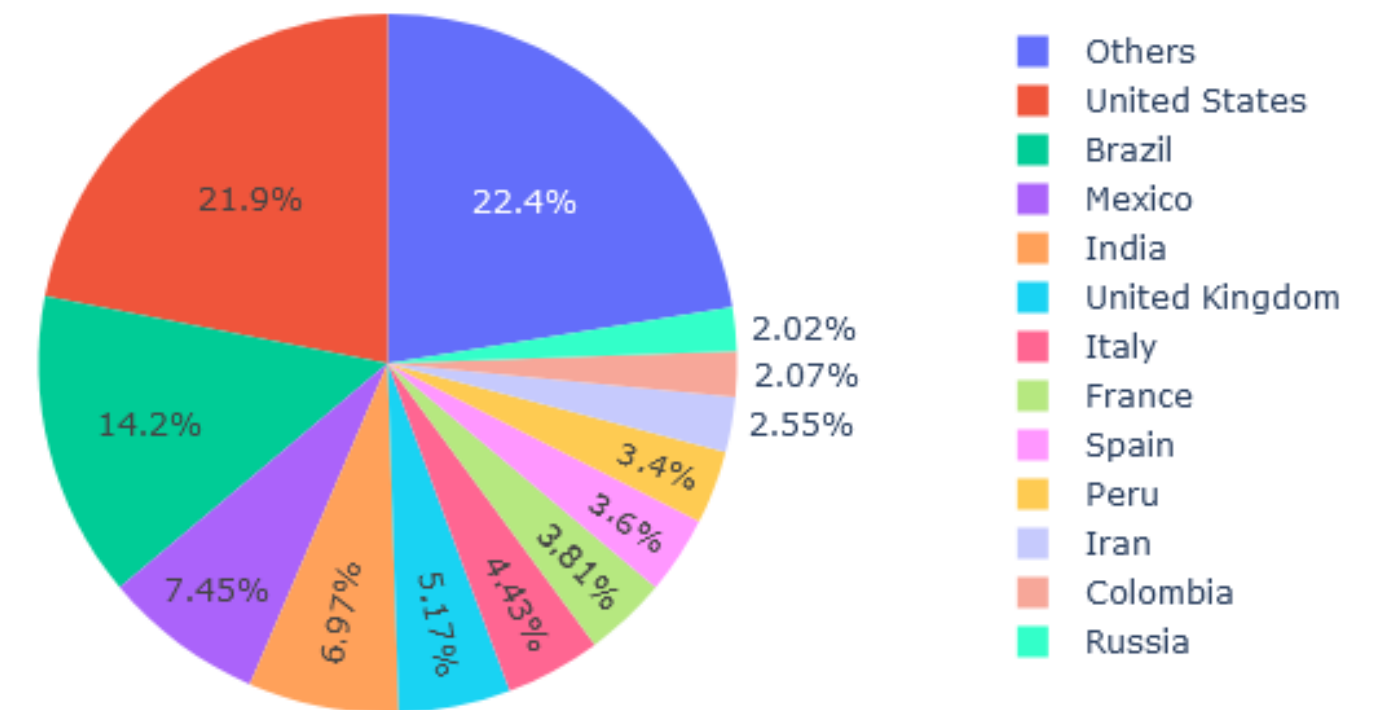
Correlaciones positivas muy altas entre todos los tipos de variables de casos y muertos. Tambien el índice de mortalidad con ellos

Covid 19 - EDA Global

WORLD Total Cases - Covid19



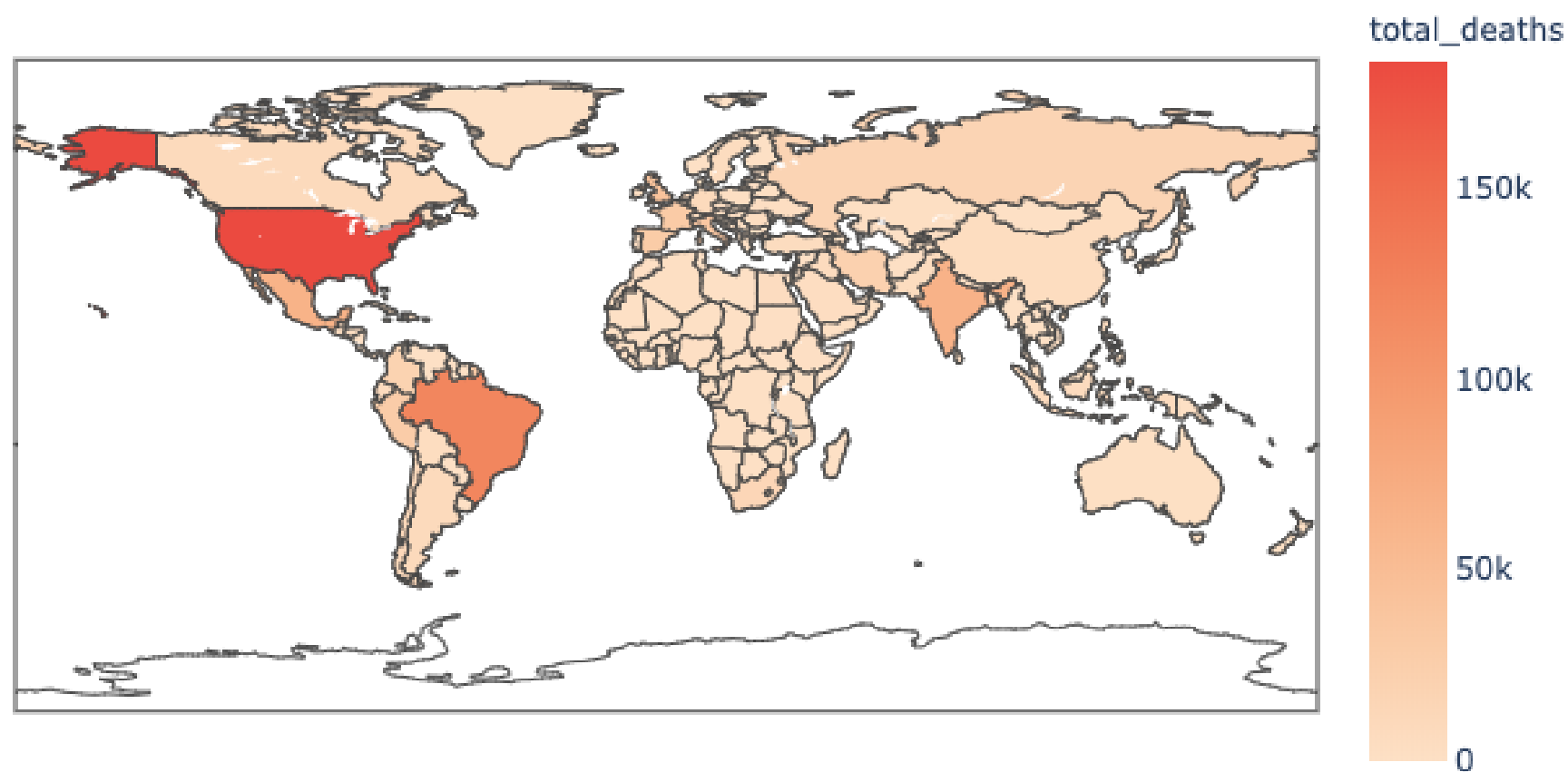
WORLD Total Deaths - Covid19



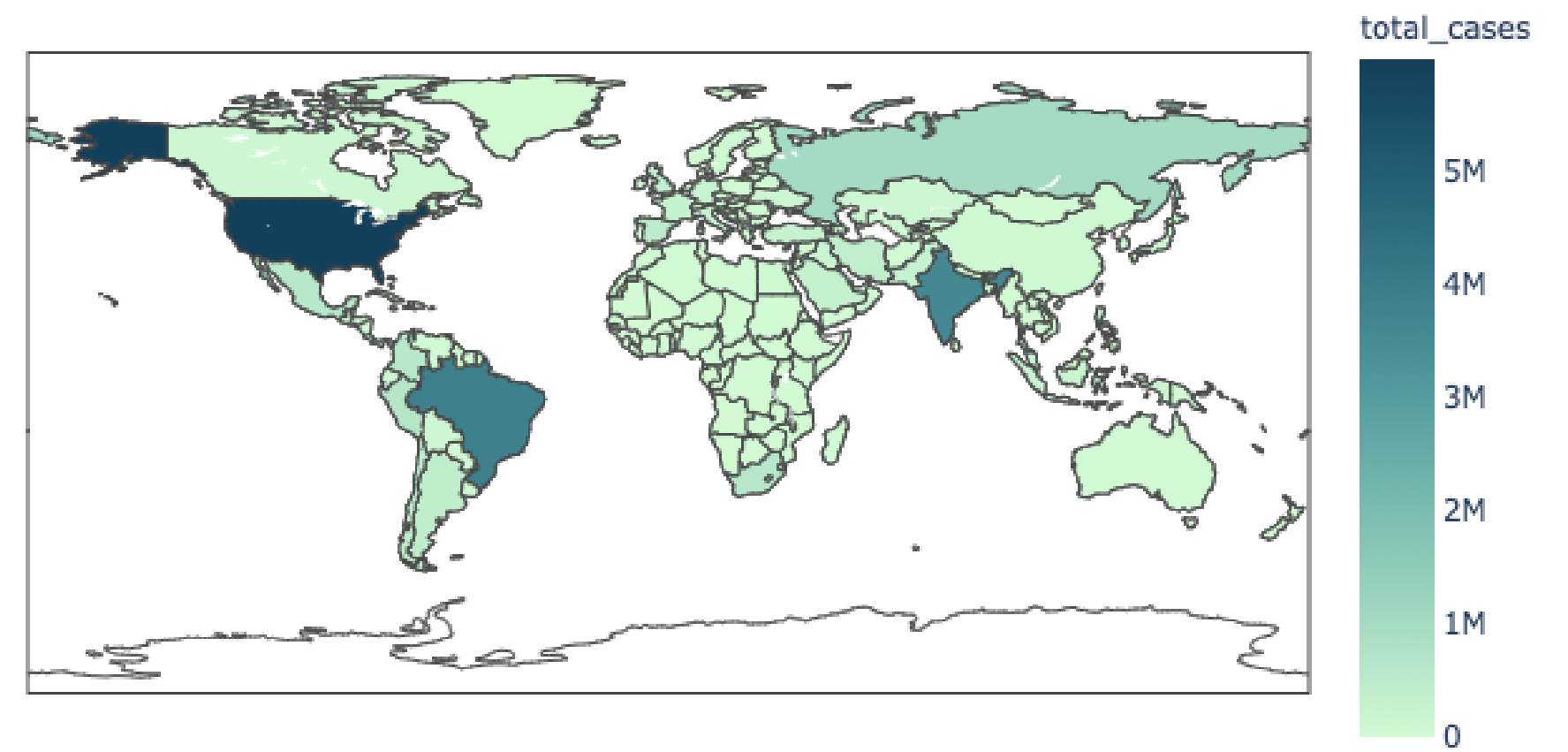
Casos y muertos totales en los países más afectados. En España y Gran Bretaña se puede inferir la diferente incidencia en cuanto a los fallecimientos.

Covid 19 - EDA Global

Total Deaths per Country - World Map



Total Cases per Country - World Map

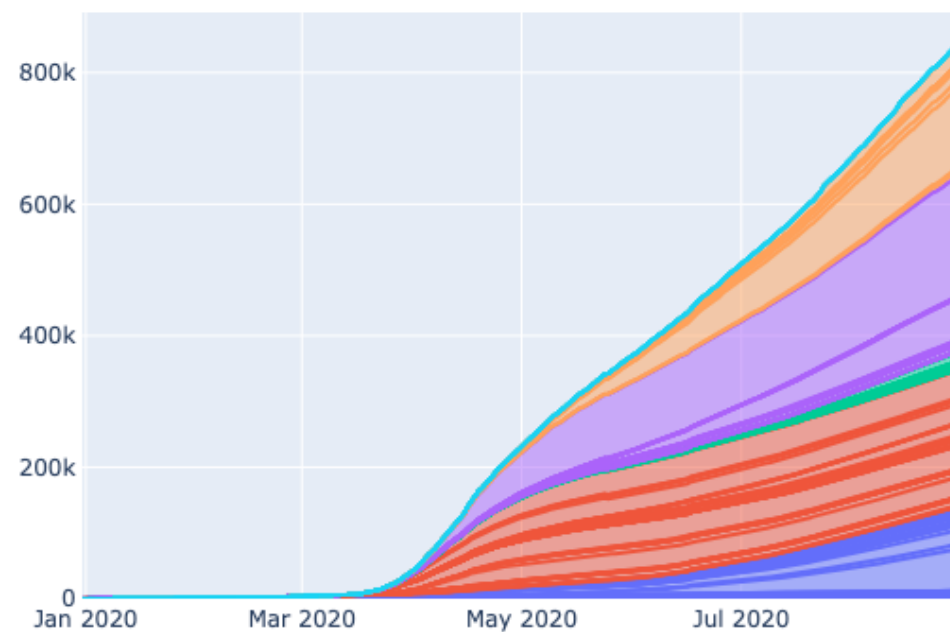


Presentations are communication tools that can be demonstrations, lectures, speeches, reports, and more.

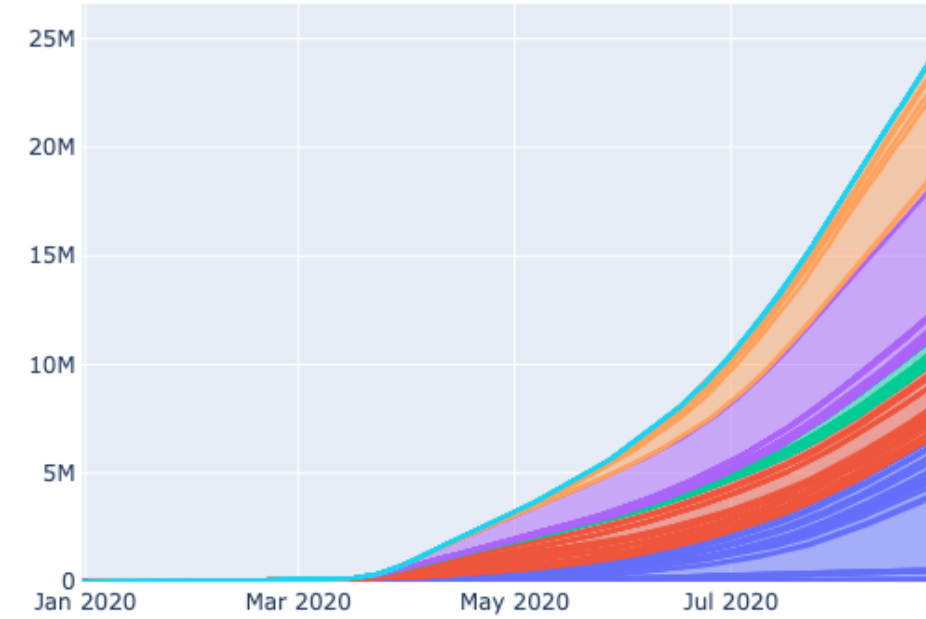
Covid 19 - Continentes

● Europa ● Asia ● Africa ● América Norte ● América Sur ● Oceanía

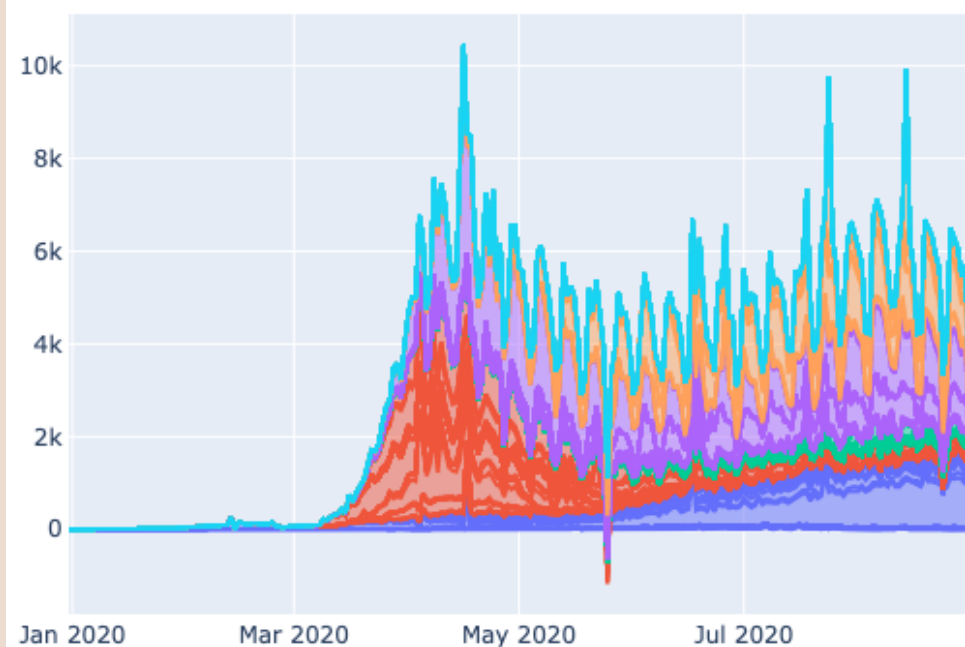
Total Deaths per continent - Area Graph



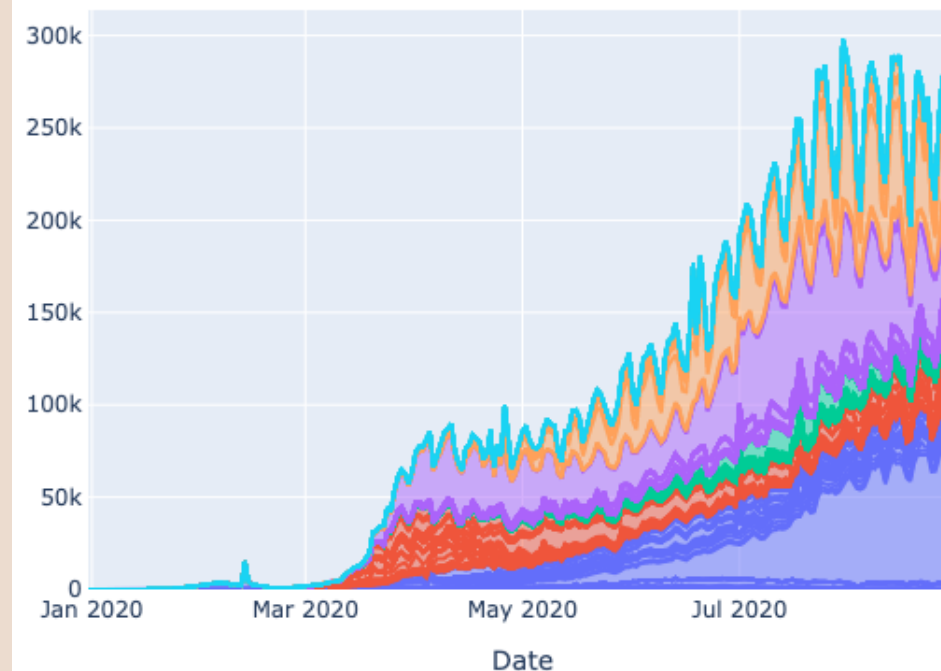
Total Cases per continent - Area Graph



New Deaths per continent - Area Graph



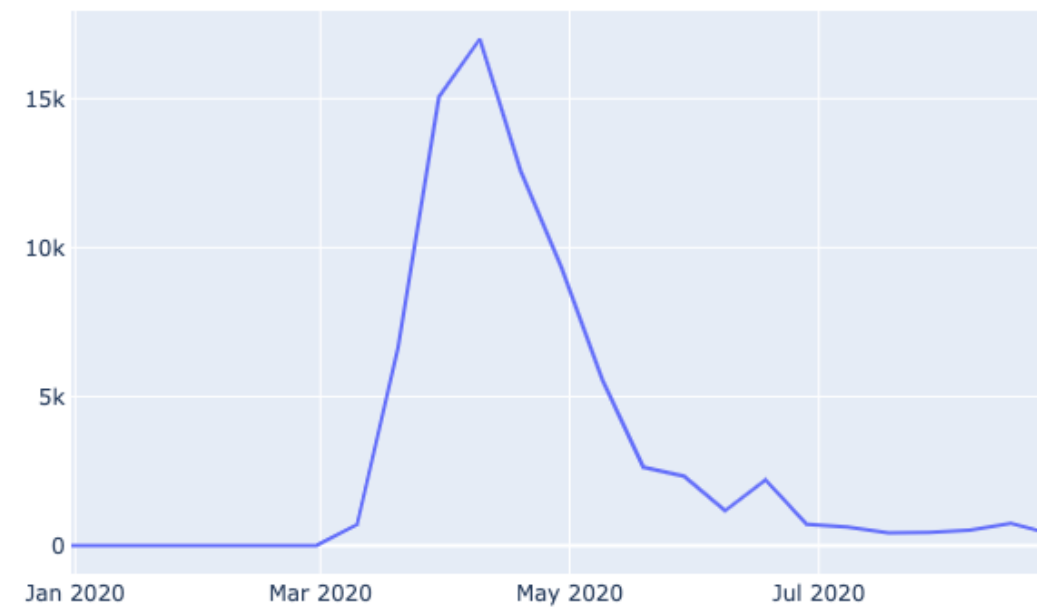
New Cases per continent - Area Graph



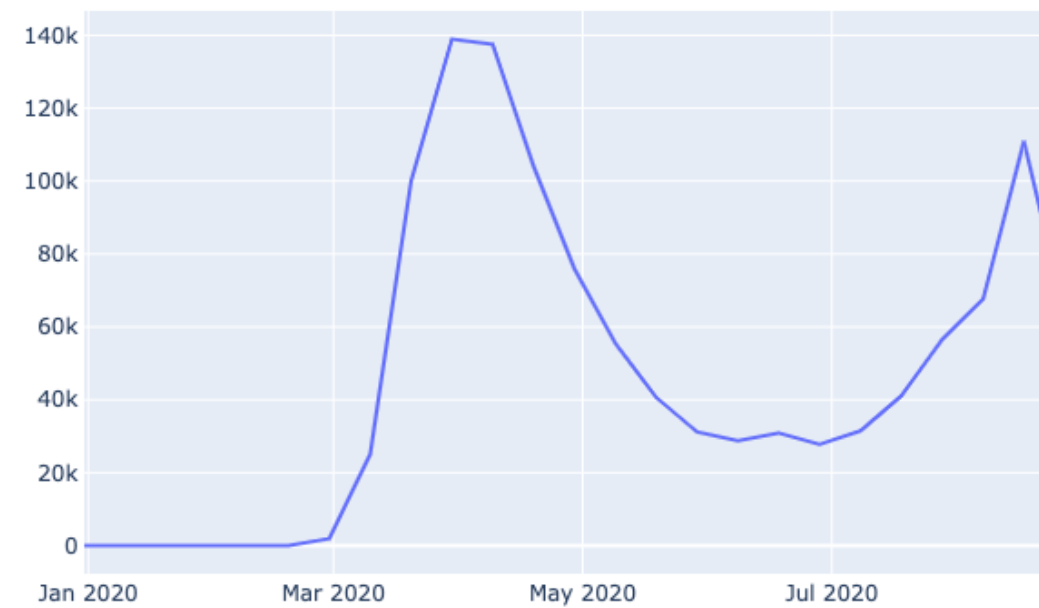
- Aqui se puede apreciar que todos los continentes han sido afectados por el virus y como la pandemia se ha ido moviendo de este a oeste empezando en Asia y Oceanía, pasando a Europa y después a América. La incidencia en África no ha sido alta

Covid 19 - Progresión (10d)

Group D - 10 days Progression (Daily Deaths)

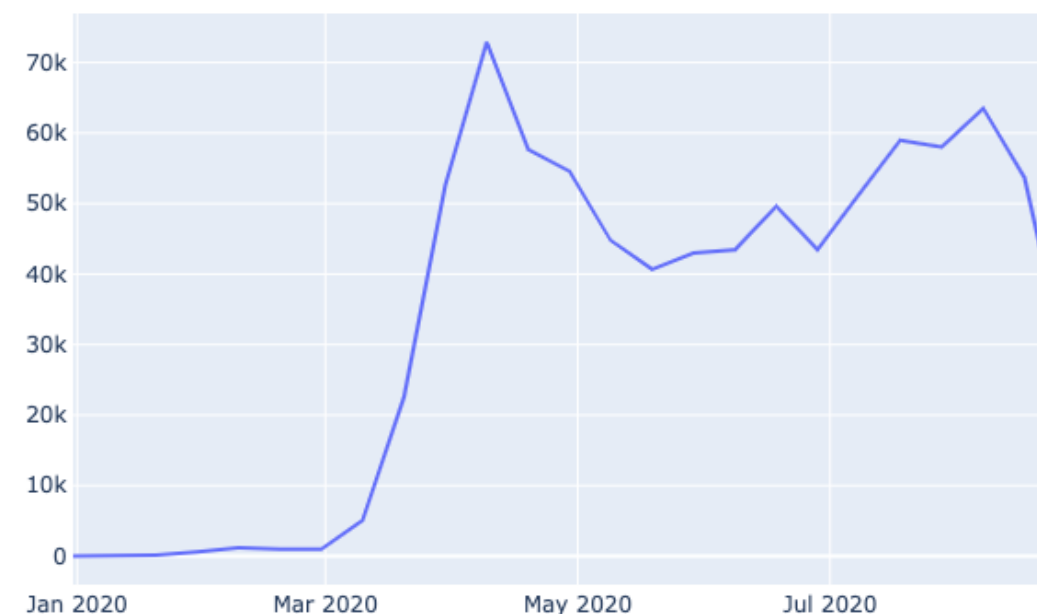


Group D - 10 days Progression (Daily infected cases)

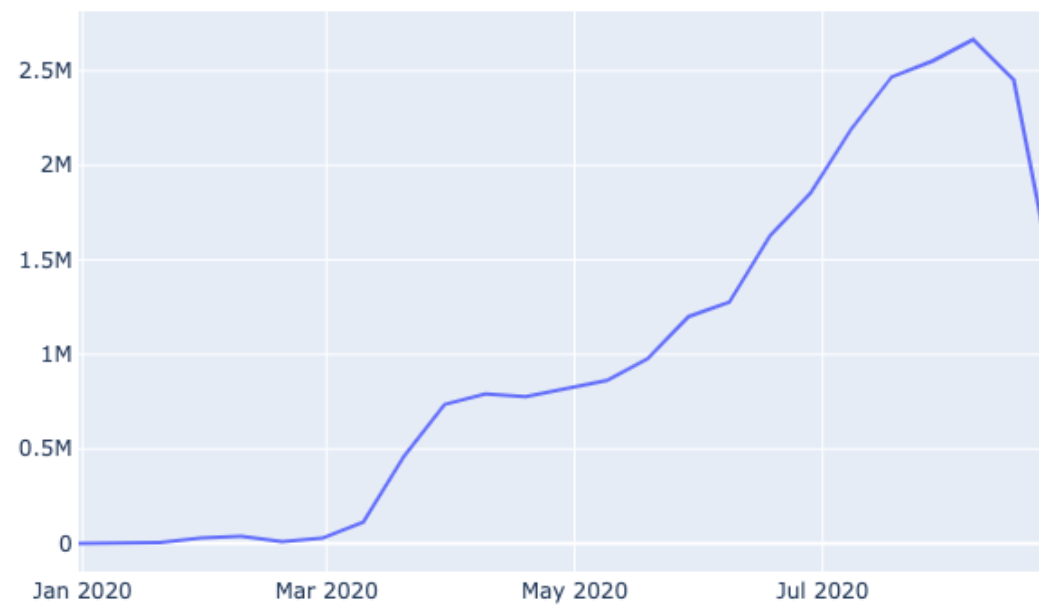


GRUPO D

World - 10 days Progression (Daily deaths)



World - 10 days Progression (Daily infected Cases)



MUNDO

Conclusiones



- Los casos se incrementan rápidamente en Marzo y 10-15 días después se observa lo mismo en el caso de las muertes.
- Algunos países del grupo D imponen estados de alerta que reducen la pendiente de las curvas de casos y fallecimientos.
- Los más afectados en cifras absolutas y relativas en muertos y en casos son ESP y GB.
- Índice de mortalidad varía entre países.
- Correcciones en los de datos, producen outliers
- Dudas sobre los datos de algunos países
- Dos-tres semanas tras el final del estado de alerta, la curva de nuevos casos vuelve a crecer parecida a Marzo. No ocurre igual con los muertos quizás por mejor tratamiento y hospitales no saturados.
- En último mes fuerte incremento en el número de test realizados.

*"Torture the data,
and it will confess to anything"*

RONALD COASE

Nobel Prize in Economics

¡Gracias!

“