

Análisis a los datos epidemiológicos del COVID 19 en los Estados Unidos

Proyecto Final Diplomado Ciencia de Datos

Eduardo López Mahaén

Introducción al proyecto

En el 2020 la enfermedad COVID 19 se empezó a extender alrededor del mundo, el covid es una enfermedad infecciosa provocada por el SARS-COV-2, está enfermedad es respiratoria y puede no tener efectos graves en el cuerpo humano, sin embargo en ciertas condiciones como en personas de la tercera edad, o gente que posee alguna enfermedad degenerativa puede desembocar en el agravamiento de está enfermedad, pudiendo provocar el fallecimiento.

Según la OMS, la creación de una vacuna y el acceso a esta ha sido fundamental para poner fin a la pandemia de COVID 19, se ha hecho una inversión significativa en la investigación y comportamiento del virus, el desarrollo de la enfermedad y la creación y mejoramiento de diversas vacunas.

Se considera que la pandemia de COVID 19 terminó el mes de mayo del 2023 y aunque todavía a la fecha hay contagios, ya se considera como una enfermedad endémica y gracias al control que se ha tenido, han disminuido los contagios y los decesos.

Pero, ¿Realmente la vacunación fue lo que provocó que terminara la pandemia?

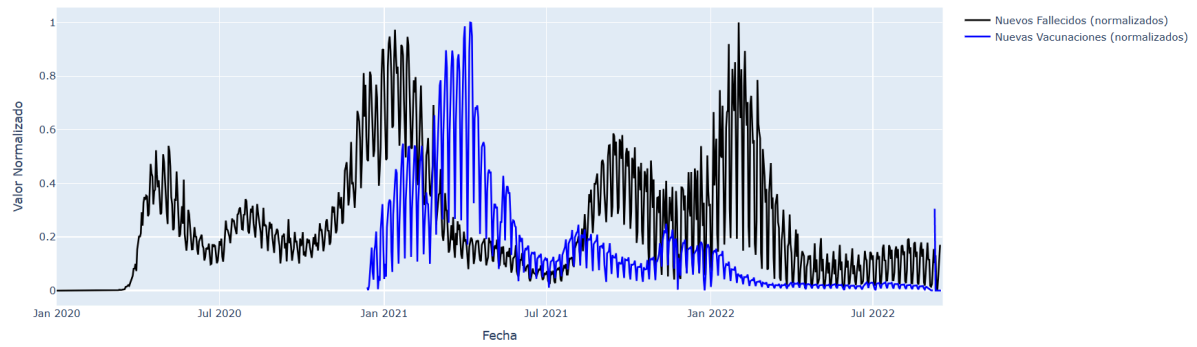
¿Hay alguna correlación entre la disminución de los casos de COVID con la aplicación de la vacuna?

En el presente trabajo se busca realizar un pequeño análisis a los datos Epidemiológicos de los casos de contagio, fallecimientos, pruebas y vacunación de COVID 19 en los Estados Unidos, además de analizar si hay alguna correlación entre la disminución de casos de COVID con la aplicación de la vacuna.

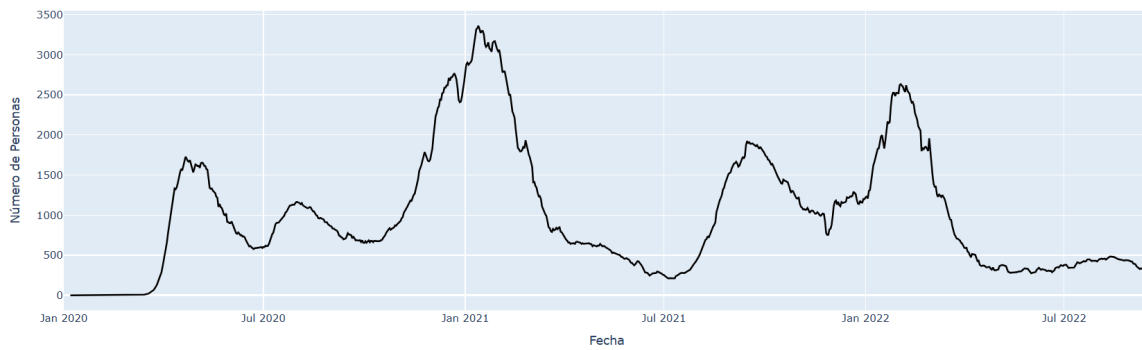
El conjunto de datos que se usará para este trabajo es el repositorio de datos abiertos del COVID 19 de Google Health, el cual fue un proyecto abierto para recopilar datos de más de 20,000 diferentes ubicaciones, extrayendo datos de diferentes datasets como la OMS, New York Times y Eurostat. En estos se encuentra información económica, políticas de salud, geografía, climatología, censos poblacionales y la evolución del actuar del gobierno hacía la enfermedad, datos epidemiológicos, el cómo se distribuyó la vacunación en el país, datos de acceso a la vacuna, entre otros.

Análisis exploratorio

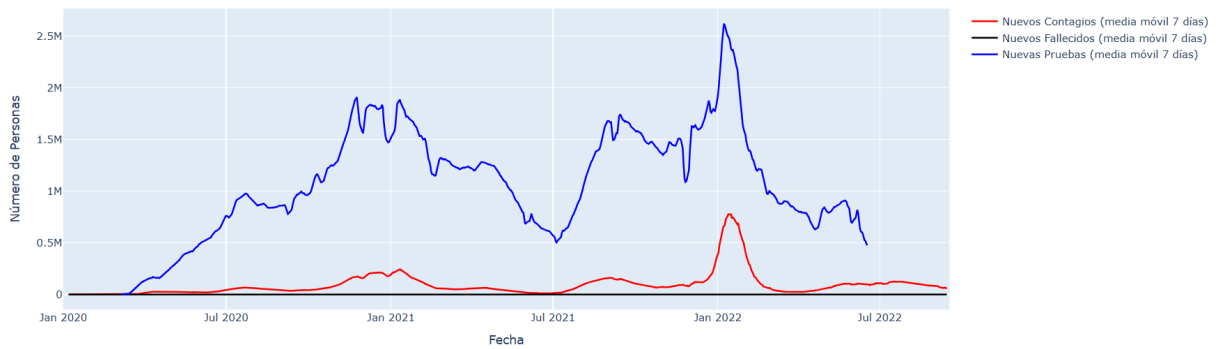
Comparación entre Nuevos Fallecidos y Nuevas Vacunaciones (normalizados)



Comparación de Datos Nuevos



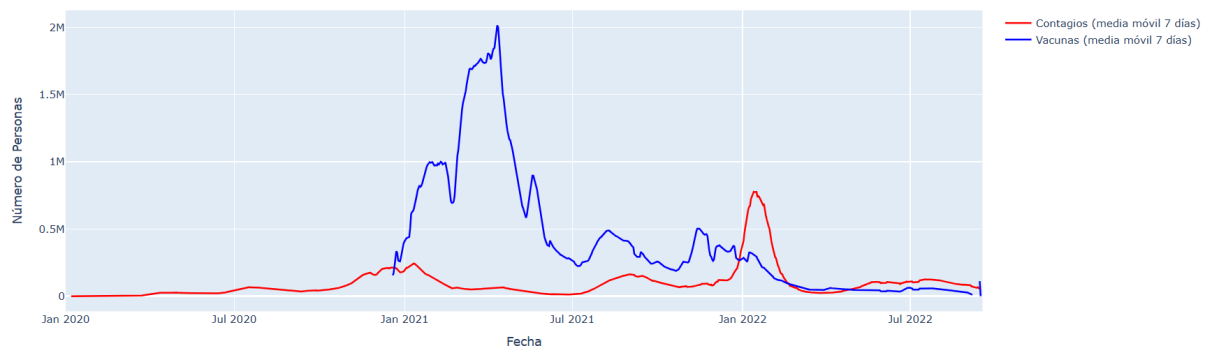
Comparación de Datos Nuevos



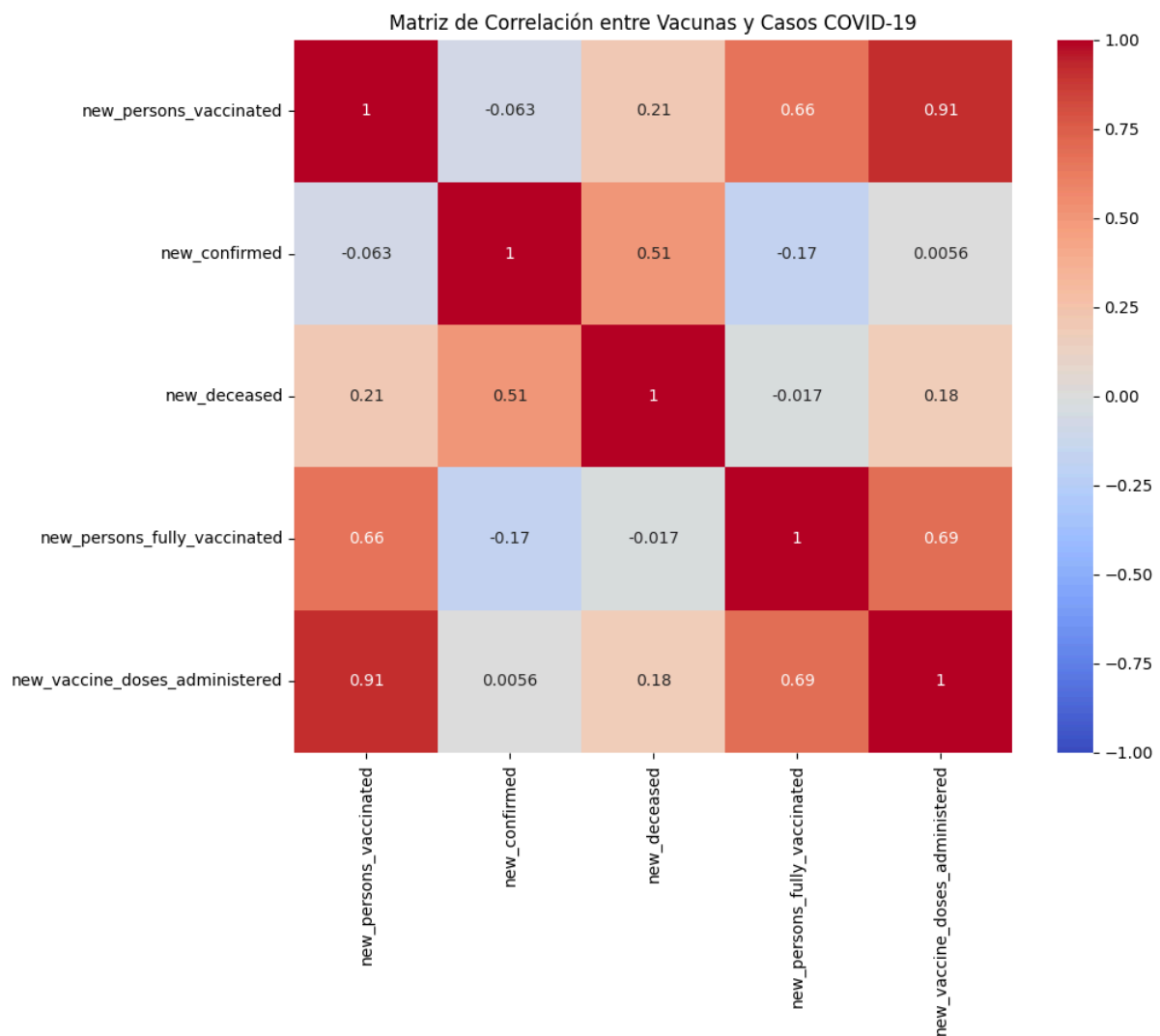
Comparación entre Contagios Acumulados y Vacunas Acumuladas



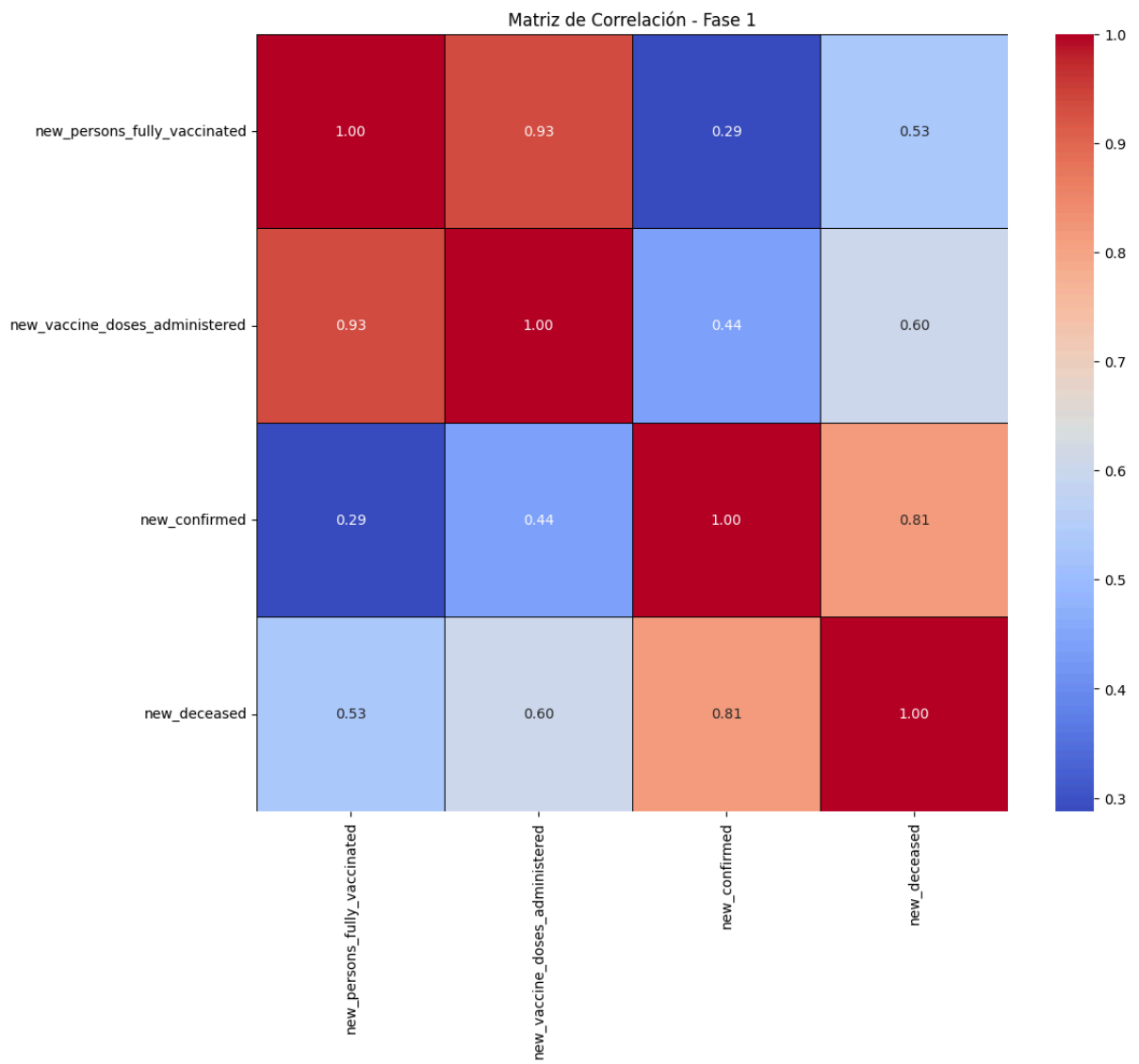
Comparación entre Contagios y Nuevas Vacunas

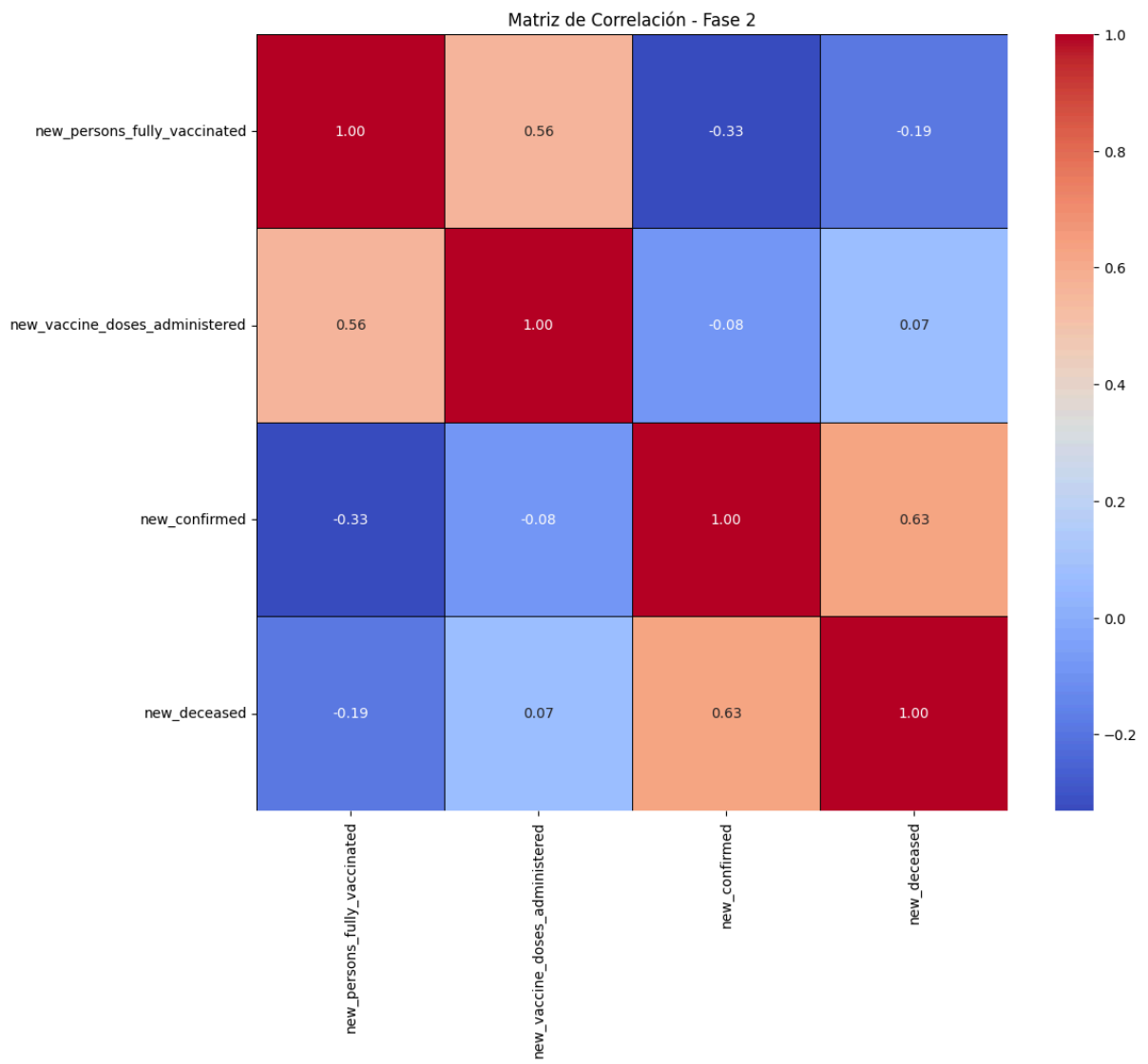


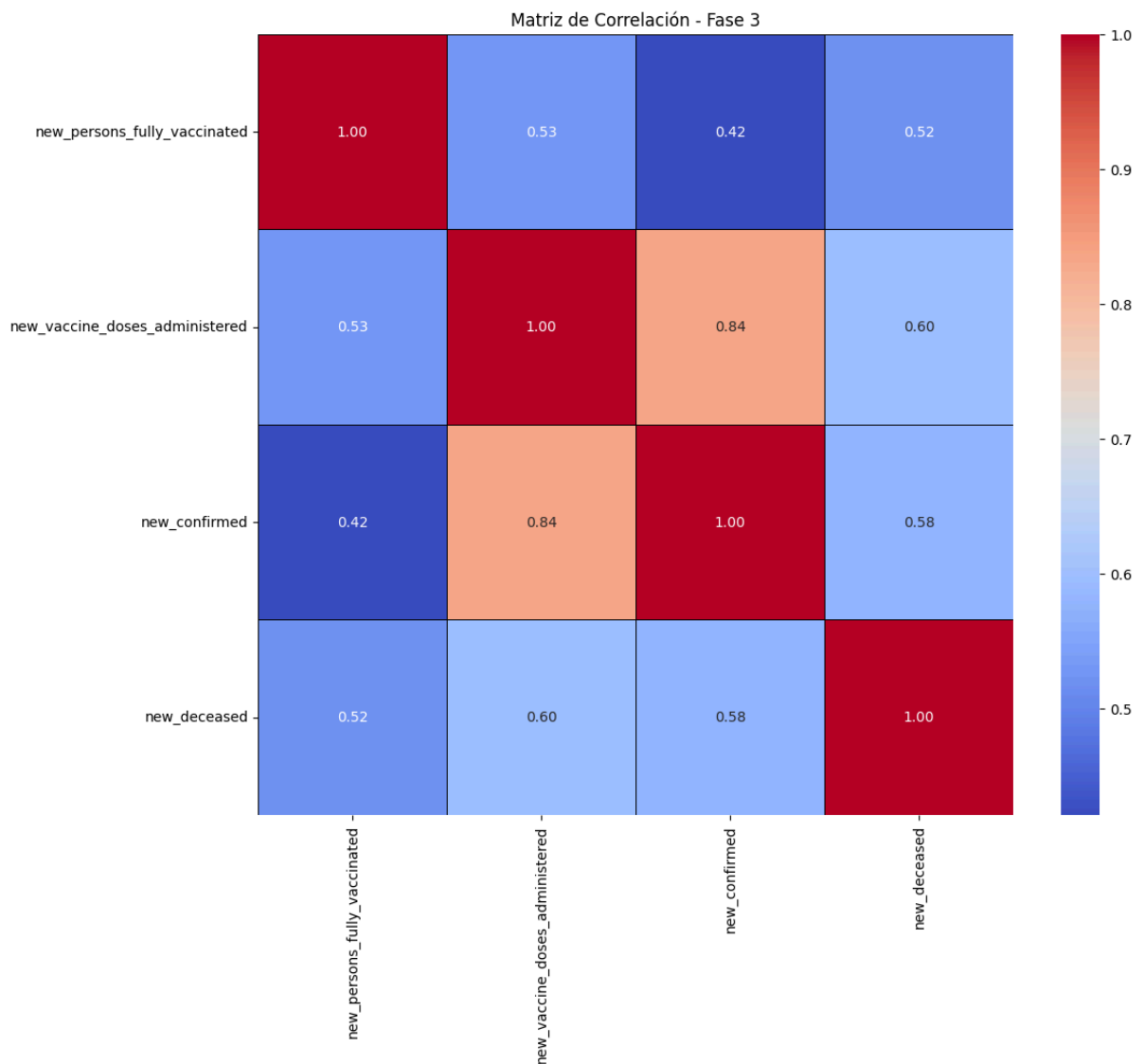
Resultados



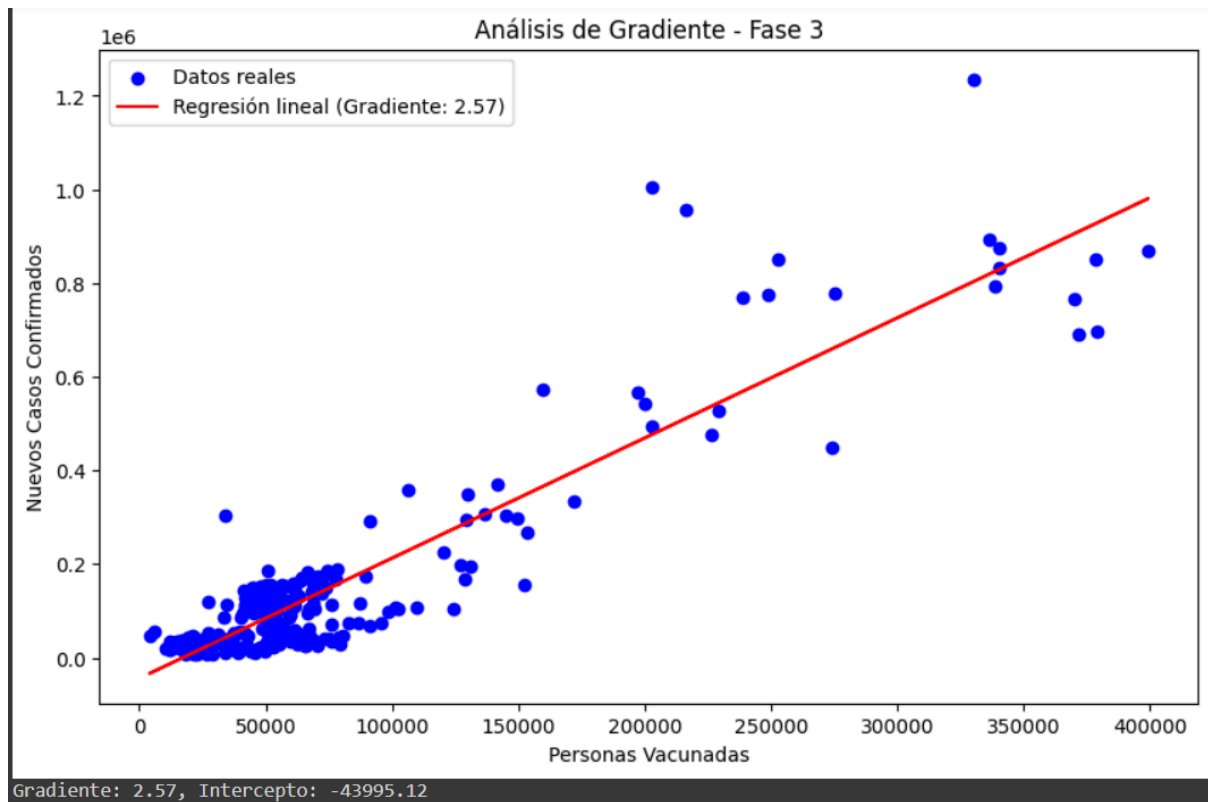
Se puede notar que la correlación entre personas vacunadas y casos confirmados es muy baja, incluso negativa, a esto se entendió que posiblemente haya habido diferentes ciclos de contagios y nuevas cepas, así que se generaron otras 3 matrices, una por año.







Se puede ver que la relación en la fase 1 y en la fase dos son muy bajas, en la fase 1 siendo el primer año porque no se había desarrollado una vacuna. Pero en la fase 3 podemos ver que hay una correlación de 0.83, a lo cual meteremos los datos de esta fase para ver que tan certera es esta información.



Un gradiente de 2.57 sugiere que por cada persona vacunada, los nuevos casos confirmados aumentan en 2.57, lo cual es inesperado. Se esperaría que sucediera lo contrario y que los casos empezaran a disminuir, cosa que no pasa.

Conclusiones

Al realizar las pruebas del modelo se pudo ver que no hay correlación entre las personas vacunadas y nuevos casos, esto depende de muchos factores como el cuidado que mantuvieron las personas con los nuevos repuntes, el cómo se redujeron políticas de home office, que los alumnos regresaran a sus colegios antes de que la cepa de omicrón empezara a propagarse.

Referencias

https://www.who.int/es/health-topics/coronavirus#tab=tab_1

<https://www.who.int/es/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines>

<https://health.google.com/covid-19/open-data/>