

PROYECTO FINAL

Equipo 8:

MÓNICA ÁNGELES SÁNCHEZ
MARÍA DEL CARMEN HUITRÓN DOMINGO

1. Introducción

Desde su aparición en diciembre de 2019, el coronavirus SARS-CoV-2 se propagó rápidamente, desafiando a los sistemas de salud y alterando la vida cotidiana. Las medidas de contención, como el distanciamiento social y el confinamiento, se implementaron en muchos países para frenar la transmisión. A medida que la pandemia avanzaba, se evidenció la desigualdad en la respuesta sanitaria y económica entre naciones, así como la importancia de la cooperación internacional.

El análisis de los datos relacionados con la pandemia es crucial para entender su impacto y preparar respuestas más efectivas ante futuras crisis sanitarias. Este proyecto se enfoca en procesar un conjunto de datos sobre los casos de COVID-19 ocurridos a nivel mundial. El objetivo es analizar cuáles son los países más afectados, en términos de número de muertes, así como conocer los países con el mayor número de personas que lograron recuperarse de la enfermedad.

2. Objetivo general

Realizar un análisis exploratorio de un conjunto de datos sobre la pandemia de COVID-19 a nivel global con el fin de identificar si existe alguna relación entre los datos proporcionados.

3. Objetivos específicos

- Limpiar y organizar la tabla para garantizar que los datos estén listos para su análisis.
- Agrupar el número de muertes y de recuperaciones por región para analizar de mejor manera los datos.
- Identificar las regiones con mayor número de muertes y con mayor número de recuperaciones, para detectar en qué países es conveniente mejorar los sistemas de salud y cuáles son los países más preparados para afrontar problemas de salud.

4. Planteamiento del problema

La pandemia de COVID 19 resaltó las diferencias que existen en el sistema de salud de cada país y de la eficacia de las medidas que cada uno toma para lograr que su población se recupere. Este proyecto tiene como objetivo procesar un conjunto de datos de los números de muertes y recuperaciones de cada región para descubrir patrones y diferencias en los sistemas de salud de cada uno.

5. Posibles preguntas de investigación

- ¿Cuál es la media, la mediana y el rango de los casos confirmados, muertes y recuperaciones en el dataset?
- ¿Cuáles son los países con mayor número de muertes?
- ¿Cuáles son los países con mayor número de recuperaciones?
- ¿Existe alguna relación entre los datos?

6. Posible solución

La solución consiste en implementar estimados de locación para conocer mejor la distribución de los datos, utilizar tablas de frecuencias, histogramas, boxplots y violinplots para visualizar distribuciones y hacer una Regresión Lineal Simple para intentar hacer predicciones entre pares de variables en nuestro dataset.

7. Conclusión

En este proyecto, nos enfocamos en analizar la distribución de un conjunto de datos del número de casos, de muertes y recuperaciones que tuvo cada región al enfrentarse a la pandemia del COVID 19, utilizando correlaciones y aplicando distintos métodos como tablas de frecuencias, histogramas, boxplots y violinplots, entre otros.