

ESTRATEGIA DE BIOGENESYS

Nombre del autor: Nicole Viviant

Email: nicoleviviant@hotmail.com

Cohorte: DA-PT05

Fecha de entrega: 22/02/2025

Institución: NicolYZe

Introducción

El propósito del proyecto se basa en el análisis estratégico de datos con el fin de entender y ayudar a la empresa Biogenesys a expandirse en latinoamérica.

Hemos realizado un análisis exploratorio de datos junto a técnicas de limpieza de los mismos, con los cuales, realizando diferentes gráficos y visualizaciones, nos encaminaron a entender la incidencia del Covid-19.

Con el fin de entender este análisis con mayor profundidad, hemos realizado un dashboard interactivo mediante el cual podremos elegir que datos explorar.

Desarrollo del proyecto

Recibida nuestra base de datos, realizamos una limpieza de filas, columnas y datos nulos reduciendo así la cantidad de datos y espacio a trabajar.

Hemos exportado nuestra propia base de datos con la limpieza ya realizada a un nuevo archivo con el cual trabajamos para conseguir los insights. Este archivo contiene la información solo de los países seleccionados, como Colombia, Perú, Argentina, Chile, Brasil y México., como también los datos posteriores al 01-01-21.

Los datos nulos o faltantes se han rellenado con la media de cada columna.

DATA ANALYTICS

EDA e insights

Las dosis de vacunas administradas lograron reducir la tasa de muertes en todos los países. Después de noviembre del 2021, países como Brasil, Chile y México, tuvieron un pico de muertes que luego se fue reduciendo a medida que pasaban los meses. Todos los países, menos Argentina, coinciden en el aumento de aplicaciones de vacunas a partir de mayo del 2021. Argentina comienza a aumentar las dosis administradas a partir de marzo del 2022 luego de tener un alto pico de casos confirmados a principios del 2022, los cuales, luego se fueron reduciendo.

Análisis del dashboard

El dashboard cuenta con una portada, donde a partir de diferentes botones, podremos elegir ver las visualizaciones por país, por temperatura, por mes, también ver variaciones y matrices. En cada hoja de visualización cuenta con un botón para volver a la portada principal y que así sea más sencilla la exploración del dashboard. No cuenta con filtros ya que estos no son compatibles con scripts de Python.

Conclusiones y Recomendaciones

Teniendo en cuenta todo el análisis trabajado durante estas semanas, la conclusión de la ubicación más óptima para la expansión de laboratorios farmacéuticos, sería sin ninguna duda Brasil. Ya que múltiples visualizaciones lo ponen como el país con más muertes, más casos, con la temperatura como gran influencia de la propagación del virus.

Durante enero del 2022, Brasil tuvo, entre los demás países seleccionados, el mayor pico de casos confirmados, como también el pico con más recuperaciones, aunque el periodo con más muertes en el país fue en abril del 2021. Esto puede tener relación con las dosis de vacunas administradas, ya que las muertes se redujeron a partir de que las personas empezaron a vacunarse.

Brasil cuenta con una expectativa de vida muy baja comparada a los demás, seguido por México, como también un bajo PIB per cápita.

Reflexión personal

Aun teniendo una base en python, lo aprendido en este modulo fue todo un reto para mi, porque no tenia conocimientos de lo dado, pero si considero que son métodos muy importantes e interesantes que dan al análisis de datos más versatilidad y manejo, En este proyecto utilicé técnicas y maneras de avanzar en este las cuales fui aprendiendo a lo largo de todo el curso y trabajos pasados. Considero que en el camino tuve y tendré errores, pero confío y creo en mi y en la base de conocimiento que me brinda y me ha brindado Henry durante este tiempo.

EXTRA CREDIT

Realice una hoja exclusiva para visualizaciones con geopandas, utilizando una fuente externa incluida en la carpeta. También realice predicciones con machine learning, las cuales todavía no consigo entender y pulir bien, pero lo intente.

