**Proyecto Integrador: Análisis de Datos**

Emmanuel Kaled Frausto Isunza A01786354

José Eduardo Balderas Rojas A01754542

Pensamiento Computacional para Ingeniería

Grupo 414

Jueves 17 de octubre de 2024

**Introducción**

La pandemia de COVID-19, iniciada en diciembre de 2019 y declarada como un caso epidemial desde marzo de 2020 ha generado una crisis de salud pública sin precedentes en la historia reciente. México como los demás países, padeció graves consecuencias en múltiples niveles: desde los daños a la salud de la población hasta los efectos en la economía y la vida cotidiana. El primer caso de COVID-19 en México fue detectado oficialmente en febrero de 2020 y el virus se extendió rápidamente por todo el país, lo que llevó a medidas de contención como confinamientos, limitación de la movilidad y uso de equipo de protección personal. A lo largo de la pandemia uno de los rastreadores clave a diario fue fundamental para vigilancia y la toma de decisiones. Los datos ayudaron a las autoridades y a la sociedad en general a comprender mejor cómo la enfermedad se comportó con el tiempo y qué tan eficaces fueron los esfuerzos para controlarla, así como a observar señales de nuevos patrones y posibles períodos de mayor riesgo. En este proyecto, se desarrolló un código en Python que analiza el archivo de casos diarios de COVID-19 en México, permitiendo observar la evolución de los casos en cada estado y a nivel nacional además de generar estadísticas y gráficas que ayudan a comprender mejor esta problemática.

**Desarrollo**

El código en Python tiene como propósito procesar y analizar datos reales de casos diarios de COVID-19 en México, registrados en un archivo CSV (*covid.csv*). Este archivo contiene una columna para cada estado y una columna nacional que resume los casos totales del país por fecha. El programa le despliega un menú al usuario para que elija que acción quiere que se realice, la primera opción es desplegar los datos generales, en donde se ven las estadísticas de todos los Estados del País y además las estadísticas generales, la segunda opción es que el usuario elija un Estado en especifico para ver sus estadísticas y también una gráfica de los casos diarios de ese Estado y la última opción es para simplemente terminar el programa.

La primera función, *cargar\_datos\_csv*, abre el archivo *covid.csv* y extrae la columna de interés (por ejemplo, datos nacionales o de un estado en particular) junto con las fechas correspondientes. Estos datos se almacenan en listas separadas, *fechas y datos*, para facilitar el procesamiento posterior. Una vez cargados los datos, el programa realiza cálculos estadísticos clave como: La media de casos diarios, la varianza y desviación estándar y el máximo y mínimo de casos. La función *graficar\_datos* utiliza la librería matplotlib para crear gráficos de líneas que visualizan los datos de casos diarios a lo largo del tiempo cuando se elige la opción 2 y se quiere ver los datos de un Estado en específico, este gráfico incluye etiquetas en los ejes para la fecha y el número de casos además de una leyenda y una cuadrícula que facilita la lectura de los datos.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

*Figura 1. Estadísticas generales*

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

*Figura 2. Estadísticas Generales*

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

*Figura 3. Estadísticas de un Estado en específico.*

**Conclusión**

Este proyecto de análisis de datos de COVID-19 en México nos ha permitido aprender a obtener una visión detallada y práctica del comportamiento de la pandemia en el país. La implementación del código en Python nos ha facilitado observar cómo los casos se comportaron en los distintos estados y a nivel nacional aportando un enfoque analítico en el procesamiento de la información. El diseño de un programa propio nos permitió adaptar el análisis a las necesidades específicas generando gráficos y estadísticas de acuerdo con el estado. Esta flexibilidad es una ventaja significativa en comparación con el uso de herramientas de terceros, ya que nos permite ajustar el análisis conforme surgen nuevas preguntas o necesidades, el tener acceso a datos y visualizaciones en tiempo real con herramientas propias es crucial para una comprensión clara de situaciones complejas (en este caso una pandemia). Este enfoque permite identificar tendencias y posibles puntos de inflexión en los casos, apoyando la toma de decisiones en momentos críticos.

**Referencias**

García, C. & Villasis, M. (13 de septiembre de 2021). *COVID-19: marca un antes y después en México, ¿punto y aparte?, ¿punto y seguido?* [*https://www.scielo.org.mx/scielo.php?script=sci\_arttext&pid=S0035-00522020000200043*](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0035-00522020000200043)

Mayo Clinic (27 de marzo de 2024). *Enfermedad del Coronavirus 2019 (COVID-19)*. <https://www.mayoclinic.org/es/diseases-conditions/coronavirus/symptoms-causes/syc-20479963>

MedlinePlus (s.f.). *Enfermedad por coronavirus 2019 (COVID-19).* <https://medlineplus.gov/spanish/ency/article/007768.htm>

ORACLE (2020). *¿Qué es la ciencia de datos?* <https://www.oracle.com/mx/what-is-data-science/>