
GESTIÓN DE DATOS PROYECTO

El coronavirus 2019 (COVID-19) es una enfermedad infecciosa causada por el coronavirus 2 del síndrome respiratorio agudo severo (SARS-CoV-2). Ésta se identificó por primera vez en 2019 en Wuhan, China, y desde entonces se ha extendido a todo el mundo, lo que resultó en la pandemia de coronavirus entre 2019-20. Los síntomas comunes incluyen fiebre, tos y dificultad para respirar. El dolor muscular, la producción de esputo y el dolor de garganta son menos comunes. La tasa de muertes por número de casos diagnosticados es, en promedio, del 3,4%, oscilando entre el 0,2% en los menores de 20 años y aproximadamente el 15% en los mayores de 80 años. Esta pandemia ha revolucionado la vida de miles de personas alrededor del mundo, experimentando diferentes comportamientos según distintos factores que hacen que no tenga una misma evolución en todo el mundo, si no que se puede apreciar claramente que hay países que se han visto mucho más afectados que otros.

Con el fin de mantener un registro y recopilar datos importantes para saber más de este virus y así tomar mejores decisiones a futuro es que se creó un repositorio en el que se ingresan todos los reportes diarios de infectados, fallecidos, recuperados, etc. Este repositorio va creciendo conforme pasa el tiempo, agregando un nuevo archivo CSV diariamente. Cada uno de estos archivos contiene información mayormente referente a USA pero no es indiferente a lo que sucede en Latinoamérica, por lo que los reportes diarios de Chile también se encuentran contabilizados. El repositorio de JHU CSSE se puede encontrar en el siguiente [link](#), en él, se puede encontrar información asociada a lo que contiene cada uno de los archivos, las columnas que los componen y muchos otros datos adicionales que pudiesen ser de interés.

Tomando en consideración todo lo anterior se pide a usted, que por medio de la utilización de este repositorio de respuesta a los siguientes requisitos que tienen relación a la difusión del coronavirus, su análisis, generación de visualizaciones, predicciones y comparaciones, tomando como base los archivos CSV contenidos en este repositorio de reportes diarios. A medida que avance en el listado de requisitos es probable que necesite usar más de un archivo CSV a la vez (esto no será indicado textualmente, usted debe decidirlo).

1. Mostrar todos los datos recopilados de un CSV en particular por pantalla, el conjunto de datos seleccionado debe contener un intervalo de días (no usar solo un CSV) y ser usado en los ítems posteriores, es decir, ese rango debe continuar siendo analizado.
2. Información sobre el conjunto de datos (número de columnas, tipos de datos, entre otros).
3. Información sobre los datos faltantes /inconsistentes dentro del conjunto de datos.
4. Número de casos confirmados, fallecidos, recuperados y activos por país/región.
5. Número de casos confirmados, fallecidos y recuperados por país/región y provincia/estado.
6. Número de casos confirmados, fallecidos, recuperados en China según provincia/estado.
7. Número de casos confirmados, fallecidos, recuperados en USA según provincia/estado.
8. Número de fallecidos por país.
9. Listado de países que no cuentan con recuperados.
10. Número de fallecidos por coronavirus por país.
11. Número de recuperados de coronavirus por país.
12. Top 10 de los países con más casos de coronavirus confirmados indicando su última fecha de actualización.
13. Gráfico de líneas (con diferentes colores) que muestre la cantidad de casos confirmados, fallecidos, recuperados y activos por país/región donde las muertes hayan superado los 50 casos.
14. Gráfico de barras que muestre la cantidad de fallecidos confirmados por coronavirus según provincia/estado en USA.
15. Gráfico de barras que muestre la cantidad de casos activos según provincia/estado en USA.

16. Gráfico de barras que compare la cantidad de casos confirmados, fallecidos, recuperados y activos según provincia/estado en USA.
17. Gráfico que muestre la cantidad de casos confirmados en todo el mundo a través del tiempo según mes.
18. Número total de casos confirmados en todo el mundo.
19. Número total de fallecidos en todo el mundo.
20. Listado de país/región que experimentaron rebrotes (un día sin casos confirmados y luego la aparición de nuevos casos).
21. Se deben generar al menos 5 salidas adicionales que resuman la evolución del COVID a lo largo de 2 años como mínimo, estas salidas deben ser creadas por usted y relacionar al menos 5 variables cada una además de la fecha.
22. Generar un reporte final usando Python, la idea es generar un pdf que explique en resumen lo descrito en el ítem 21 (estilo dashboard).

Consideraciones:

- La resolución del proyecto debe ser subida a Ev@ en un archivo .zip con su apellido (apellido.zip).
- Es necesario que los CSV usados en el desarrollo sean incorporadas al archivo adjunto solo en el caso que el tiempo de consulta a estos repositorios tome mucho tiempo.
- Códigos similares/duplicados serán evaluados con nota mínima.
- Modalidad individual.
- Plazo de entrega máximo 12 de Julio en horario de clases.