# Contexto

La pandemia de COVID-19 marcó el año 2020 y tuvo un impacto en Colombia y en todo el mundo. La información sobre los casos positivos de COVID-19 en Colombia durante ese año es esencial para comprender el impacto de la pandemia en el país. Estos datos muestran la propagación del virus, la respuesta del gobierno y la capacidad del sistema de salud. En este resumen, exploraremos las tendencias, los desafíos y las lecciones aprendidas de los datos de casos positivos de COVID-19 en Colombia en 2020, ofreciendo una visión general de cómo la pandemia afectó al país y cómo se enfrentó a ella.

# Objetivos.

1. Identificar el rango de edad de las personas que tuvieron mayor afectación.
2. Identificar los departamentos con mayor afectación, con más casos de COVID confirmados.
3. Evaluar la eficiencia de las medidas de prevención, en relación con las personas recuperadas.

# Herramienta utilizada:

La herramienta que se utilizó para el análisis de los datos fue Power Bi.

# Metodología

La metodología que se utilizó fue CRISP DM, un proceso estándar de la industria para la minería de datos, un método probado para guiar sus operaciones de minería de datos. Se llevaron a cabo tres etapas. En la primera fase se vio la comprensión del negocio, que incluyó una descripción de la organización, sus objetivos generales y específicos, el proceso comercial y una evaluación de la situación actual de los recursos (Hardware, Software, fuentes de datos y personal involucrado), los objetivos u objetivos de minería de datos y los indicadores para análisis de criterios para medir el éxito. La comprensión de los datos en cuestión, sus características, la exploración de los datos, la verificación y la gestión de la calidad y los datos finales se llevaron a cabo en la segunda fase. La reparación de datos, las fuentes de datos, los modelos dimensionales, la descripción del esquema y la extracción, transformación y carga (ETL) se utilizaron en la etapa 3.

# Descripción de los datos.

Los datos utilizados se obtuvieron de una fuente de la gubernamental, así que datos son verificados y confiables.

# Procesamiento de los datos.

Los datos fueron transformados antes de empezar el análisis. Se organizó algunos nombres de columnas y en “pertenencia étnica se agregaron directamente los nombres de las etnias.

# Análisis de los datos.

Para el análisis se utilizaron diagramas de barras para una mejor visualización y comprensión de los datos.

# Resultados.

1. El informe mostró el rango de edad de las personas más afectadas por el COVID 19, que fue entre 20-40 años. En ese rango, se encuentra la mayor concentración de afectados.

2. Los departamentos con mayor afectación fueron, en primero lugar Bogotá con 316156; en segundo lugar, Antioquia con 163830 y en tercer lugar se encuentra Valle con 67785.

3. Al analizar los datos, se concluye que las medidas tomadas para la prevención y propagación de virus fueron eficientes. Porque el porcentaje de recuperados en comparación de personas fallecidas fueron superiores en una gran diferencia.

# Hallazgos claves.

Las zonas con mayor afectación fueron los departamentos más grandes.

# Conclusiones.

La pandemia de COVID-19 en 2020 tuvo un impacto significativo en Colombia, poniendo en evidencia una variedad de tendencias y desafíos que enfrentó el país. Los datos analizados con la herramienta Power Bi y la metodología CRISP DM brindan una visión detallada de cómo se propagó el virus y cómo el gobierno y el sistema de salud respondieron.

Se llegaron a varias conclusiones importantes:

El grupo de edad más afectado fue el de entre 20 y 40 años, lo que podría deberse a su mayor actividad laboral y social durante la pandemia.

Los departamentos con más casos confirmados fueron Bogotá, Antioquia y Valle, lo que se debe a una mayor densidad poblacional y actividad económica en estas áreas.

La eficacia de las medidas preventivas. Debido a que el porcentaje de personas recuperadas fue significativamente superior al de fallecidos, las medidas de prevención y control implementadas fueron efectivas. Esto demuestra que se ha llevado a cabo una gestión efectiva de la crisis sanitaria, pero es necesario evaluar y mejorar constantemente las estrategias de respuesta.

También se encontró que las áreas más afectadas se encontraban en los departamentos más grandes y urbanizados, lo que demuestra la importancia de utilizar enfoques distintos y recursos adecuados para diferentes regiones del país. Estos hallazgos destacan la importancia de la preparación, la respuesta rápida y la adaptación de estrategias basadas en datos precisos y verificables para futuras emergencias sanitarias.