

■ CASO PRÁCTICO: Análisis de Datos en Biogenesys (Proyecto Integrador)

1. Entendimiento del Problema / Negocio

- Objetivo: analizar el impacto del COVID-19 en los países latinoamericanos donde Biogenesys desea expandirse.
- Preguntas clave: ¿Qué países fueron menos afectados? ¿Dónde conviene expandirse primero?
- Stakeholders: equipo directivo de Biogenesys y analistas estratégicos.

2. Obtención de los Datos

- Fuente: dataset público de estadísticas de COVID-19.
- Formato: CSV.
- Herramientas usadas: Python (pandas), Jupyter Notebook.

3. Limpieza de los Datos

- Eliminación de columnas innecesarias como 'new_recovered' y 'cumulative_recovered'.
- Conversión de fechas al formato datetime.
- Normalización de nombres de columnas y valores nulos.

4. Análisis Exploratorio (EDA)

- Identificación de los países con mayor y menor cantidad de casos y muertes.
- Análisis de tendencia temporal por país.
- Visualización de distribución de casos acumulados y tasas por millón de habitantes.

5. Modelado de los Datos

- Creación de métricas personalizadas: tasa de mortalidad, tasa de contagio, crecimiento mensual.
- Filtrado por países de interés de expansión: LATAM.
- Agregaciones temporales por mes.

6. Visualización y Storytelling

- Gráficos de barras para comparar países.
- Gráficos de línea para evolución temporal.
- Dashboard final en Power BI mostrando KPIs por país.

7. Presentación del Análisis

- Exposición de insights: países con mejor control de la pandemia.
- Recomendación: priorizar países con baja mortalidad y alta recuperación.
- Presentación en demo con storytelling claro.

8. Documentación y Entregables

- Notebook con limpieza y análisis en Python.
- Informe PDF con conclusiones y recomendaciones.
- Tablero en Power BI embebido o como archivo .pbix.