

PRUEBA SOLEMNE N°2

TALLER DE PROGRAMACIÓN II

1. Objetivo

Objetivo Aplicar técnicas de limpieza, transformación y visualización de datos en Python, utilizando exclusivamente el dataset proporcionado (**Forma_B.csv**), y demostrar resultados en un Jupyter Notebook y en una presentación de clase.

2. Entregables

1. **Jupyter Lab** (.ipynb) con todo el análisis y código documentado.
 2. **Presentación** (máx. 8 diapositivas en .pptx o .pdf) que resuma los resultados clave. 10 minutos máximo de presentación por grupo.
 3. **Dashboard en Streamlit (.py):**
Incluyendo visualizaciones interactivas basadas en los gráficos generados en el análisis, y al menos un elemento interactivo (botón, slider o selector).
-

3. Entorno recomendado

- Python 3.x
 - Librerías: pandas, matplotlib
-

4. Flujo de trabajo y pasos

4.1 Carga y exploración inicial

- Importar el CSV en un DataFrame.
- Mostrar df.info() y df.describe() para conocer estructura y estadísticas básicas.

4.2 Limpieza de datos

- Rellenar valores faltantes en columnas relacionadas con causas de muerte con 0.
- Rellenar valores faltantes en 'code' con 'Desconocido'.

4.3 Normalización y renombrado de columnas

- Convertir todos los nombres a minúsculas, sin espacios (_).
- Renombrar **todas** las columnas al español, ejemplo:
 - ✓ 'country' → 'pais'
 - ✓ 'code' → 'codigo'
 - ✓ 'year' → 'año'
 - ✓ 'hiv_aids' → 'vih_sida'
 - ✓ 'respiratory_infections' → 'infecciones_respiratorias'

4.4 Conversión de tipos y estadísticas

- Asegurar que las columnas numéricas están en formato float o int.
- Calcular media, mediana y desviación estándar de 'total_muertes' usando métodos de Pandas.

4.5 Cálculo de muertes totales

- Crear columna 'total_muertes' como la suma de todas las causas de muerte por fila.

4.6 Filtrado de países destacados

- Determinar el percentil 75 de 'total_muertes'.
- Seleccionar países con 'total_muertes' \geq percentil 75.

4.7 Análisis por año

- Agrupar por año para calcular el promedio de muertes anuales ('mortalidad_anual').
- Identificar las 5 causas más comunes y los 5 países con más muertes acumuladas.

4.8 Visualizaciones

- KDE de 'total_muertes'.
- Gráfico de torta con la distribución de las 5 causas de muertes más comunes.
- Barras de los 5 países con más muertes acumuladas.
- Línea de tendencia de 'total_muertes' por año.
- Comparación de las 2 causas principales a lo largo del tiempo.

4.9 Implementación de un Mini-Dashboard en Streamlit

Como parte del trabajo, se debe desarrollar un dashboard sencillo utilizando Streamlit, que permita visualizar de forma interactiva algunos de los gráficos generados en el análisis.

Requerimientos mínimos del dashboard

- Debe incluir al menos dos visualizaciones previamente generadas en la sección 4.8.
- Debe incorporar al menos un elemento interactivo, como un botón, slider o menú desplegable.
- Debe permitir modificar o filtrar parcialmente la información visualizada a partir del elemento interactivo.
- El dashboard debe presentarse en un archivo independiente llamado dashboard.py.

5. Formato y documentación

- Cada sección del Notebook debe incluir título y explicación breve en Markdown.
- El código debe estar limpio, con comentarios que aclaren los pasos.

En la presentación, incluir:

- Objetivo y metodología de trabajo.
 - Hallazgos principales (con gráficos clave)
 - Conclusiones y recomendaciones
-

6. Criterios de evaluación (100 pts)

- Limpieza y renombrado de columnas: 5 pts
- Cálculo y estadística de total_muertes: 10 pts
- Creación y filtrado: 15 pts
- Agrupación y análisis anual: 10 pts
- Calidad y claridad de visualizaciones: 8 pts
- Documentación y estructura del Notebook: 12 pts
- Calidad de la presentación oral y diapositivas: 20 pts
- Nuevo: Mini-dashboard en Streamlit (funcionalidad e interactividad): 20 pts

7. Fechas límite

- **Entrega del Notebook:** 23 de diciembre 2025 hasta las 21:00 horas.
- **Entrega de la Presentación:** 23 de diciembre 2025 hasta las 21:00 horas.
- **Exposición en clase:** 23 de diciembre 2025 a las 19:00 horas.