

Control 4 Visualización de la Información

En este control, realizarás una comparación de dos décadas específicas y los datos por género, generando gráficos de streamgraph y lluvias de palabras para explorar y visualizar estas comparaciones. En este notebook exploraremos algunas visualizaciones donde el texto es protagonista a través del dataset `guaguas` (`1920-2020.csv`).

Archivo proporcionado:

Parte 1: Comparación de Décadas y Género (70 puntos)

1. Seleccionar Décadas:

- Escoge dos décadas específicas para comparar (por ejemplo, 1950-1960 y 1980-1990).

2. Filtrar por Género:

- Filtra los datos para obtener los nombres más populares por género (masculino y femenino) **para las décadas seleccionadas**.

3. Generar Visualizaciones:

- Utiliza el código proporcionado para crear gráficos de streamgraph para cada combinación de década y género, y aplica el suavizado gaussiano. Elimina las etiquetas del eje.
- Alternativamente, puedes crear una nube de palabras para cada década por género.

4. Análisis de Resultados:

- Interpreta los gráficos generados y describe cualquier patrón o tendencia observable.
- Compara los resultados entre las dos décadas seleccionadas y entre los géneros.

Requisitos Adicionales:

- Asegúrate de que tus gráficos sean claros y estéticamente agradables.
- Incluye títulos y leyendas apropiadas para facilitar la interpretación de los gráficos.
- Presenta tus hallazgos en un Jupyter Notebook, incluyendo el código utilizado y las visualizaciones generadas.

Puntos a Evaluar:

- Selección y justificación de las décadas (10 puntos).
- Filtrado correcto de los datos por género (10 puntos).
- Calidad y claridad de los gráficos generados (30 puntos).
- Análisis e interpretación de los resultados (20 puntos).

Parte 2: Lluvia de Palabras (30 puntos)

1. Selección de Décadas y Género:

- Utiliza las mismas décadas y géneros seleccionados en la Parte 1.

2. Filtrar los Datos:

- Filtra los datos para obtener los nombres más populares por género (masculino y femenino) para las décadas seleccionadas.

3. Generar la Lluvia de Palabras:

- Crea una visualización de lluvia de palabras (word cloud) para cada combinación de década y género.
- Utiliza colores y tamaños de fuente para representar la frecuencia de los nombres. Los nombres más frecuentes deben ser más grandes y destacados.

4. Análisis de Resultados:

- Interpreta las lluvias de palabras generadas y describe cualquier patrón o tendencia observable.
- Compara los resultados entre las dos décadas seleccionadas y entre los géneros.

Requisitos Adicionales:

- Asegúrate de que las lluvias de palabras sean claras y estéticamente agradables.
- Incluye títulos y leyendas apropiadas para facilitar la interpretación de las visualizaciones.
- Presenta tus hallazgos en un Jupyter Notebook, incluyendo el código utilizado y las visualizaciones generadas.

Puntos a Evaluar:

- Filtrado correcto de los datos por género (5 puntos).
- Calidad y claridad de las lluvias de palabras generadas (15 puntos).

- Análisis e interpretación de los resultados (10 puntos).

Entrega:

- Sube el Jupyter Notebook con el código, las visualizaciones y el análisis a Canvas antes de la fecha límite.
- Asegúrate de que el Notebook se ejecute sin errores y que todo el contenido sea comprensible.