

Guion del EDA: "Tendencias de YouTube en España a lo largo del tiempo"

0. Introducción del proyecto

- Objetivo: Estudiar las tendencias de YouTube en España.
- Interés: Analizar cómo viralizar un video y evaluar la calidad del contenido popular.

1. Obtención de los datos

- Fuente: Kaggle Dataset (Trending YouTube Videos 113 countries)
- Acciones:
 - Descargar el dataset completo.
 - Filtrar registros cuyo país sea "ES" (España).
 - Guardarlo como CSV separado.

2. Limpieza y preparación de los datos

- Eliminar columnas innecesarias.
- Eliminar registros duplicados.
- Procesar fechas:
 - Convertir publish_time y trending_date a formato fecha.
- Crear columnas auxiliares: año, mes, día de la semana, hora.
- Mapear category_id a nombres de categorías.
- Tratar datos faltantes si los hubiera.

3. Análisis exploratorio básico (EDA)

- Estadísticas:
 - Número total de videos tendencia.
 - Número de videos distintos.
 - Número de canales distintos.
 - Distribución por año.
 - Duración media.
 - Media de visitas, likes, comentarios.

4. Planteamiento de preguntas a resolver

- Parte 1: Cómo viralizar un video
 - Duración óptima para viralizarse.
 - Mejor día de la semana para publicar.
 - Categorías dominantes.
 - Análisis de títulos (patrones de éxito).
 - Relación entre vistas, likes y comentarios.
- Parte 2: Calidad del contenido
 - Categorías con mayor engagement.
 - Engagement vs duración.
 - Canales con contenido superficial vs educativo.
 - Características del "contenido basura".

5. Visualizaciones

- Histograma de duración de videos.
- Barras apiladas de categorías por año.
- Heatmap día de la semana vs videos tendencia.
- Scatterplots likes vs views, comments vs views.
- Boxplot de duración vs engagement.
- Wordcloud de palabras en títulos.
- Timeline de tendencias por categoría.

6. Conclusiones

- Respuestas a las preguntas planteadas.
- Reflexión final sobre viralidad y calidad de contenido.

7. Presentación final

- Introducción breve.
- Metodología (datos y limpieza).
- Análisis y visualizaciones.
- Preguntas resueltas.
- Conclusiones.
- Opcional: limitaciones y mejoras futuras + presentación interactive (canva o similares)