

Python intermedio – Proyecto final

Parte 1

A partir de una marca o proyecto definir el siguiente esquema partiendo de la siguiente pregunta: ¿Qué problema de negocio tenés hoy en día? Responder por lo menos 3 de las siguientes preguntas.

Definir problema	Con qué datos	Modelo	Storytelling	Impacto
Se debe analizar la cantidad de videos subidos a YouTube de los 100 canales top de la plataforma para analizar si su numero de videos subidos influye en el incremento de número de suscriptores.	Dataset: YouTube Channels Dataset Obtenido de: https://www.kaggle.com/datasets/taimoor888/top-100-youtube-channels-in-2024?resource=download		Un proyecto de coaching quiere incorporar un nuevo curso de aprendizaje y uso de YouTube para generar ingresos pasivos a través de subir "faceless" videos de tendencia generados con IA y se quiere conocer cuales canales presentan un incremento de vistas por video, en base a si tienen un mayor número de suscriptores o menor.	La creación de un modelo de negocios en base a un nuevo curso que se pueda ofrecer. El impacto se verá reflejado en el incremento de ventas del curso aumentando la popularidad de la plataforma de coaching.

Parte 2

Aprovechando el buscador de google dataset, presenta un máximo de 3 sets de datos que ayuden a comprender y explotar tu proyecto. También podés elegir datasets propios siempre y cuando no tengan problemas de privacidad.

En este caso he elegido un data set de la página Kegel ya que maneja datasets más sencillos y de dominio público. En mi caso elegí un dataset de YouTube.

<https://www.kaggle.com/datasets/taimoor888/top-100-youtube-channels-in-2024?resource=download>

Parte 3

Teniendo en cuenta la marca o proyecto seleccionado y los datasets elegidos, responde en no más de 2 páginas las siguientes preguntas, que te ayudarán a comprender tus datos:

- **¿Qué problema comercial está tratando de resolver la propuesta?**

Identificar si el número de suscriptores influye en el número de vistas por video con el objetivo de determinar cómo se estructurará el curso y si el enfoque principal será hacia tener más suscriptores o vistas por video.

- **La propuesta se centra en mostrar cómo a través del dominio de python y diferentes librerías para análisis de datos, se puede obtener skill necesarios para poder crear dashboards o informes personalizados en el rol de data analyst**

- **¿Quiénes son los usuarios finales? ¿Son técnicos o no técnicos? ¿Qué valor esperan obtener de este producto?**

El usuario final es el departamento de diseño e implementación de cursos y el equipo de marketing. El equipo del departamento de diseño e implementación de cursos necesita tener esta información para determinar cómo se estructurará el curso y si el enfoque principal será hacia tener más suscriptores o vistas por video. Mientras que el equipo de marketing necesita saber cuál será la estructura de los cursos para poder hacer las campañas publicitarias correspondientes al objetivo del curso.

- **Los usuarios finales son personas con conocimientos técnicos de python que quieran aprender acerca de analítica de datos y como a través de diferentes librerías se puede llegar a aprender skills necesarias en un rol de data analyst o BI en una empresa**

- **¿Qué son los datos? ¿Los datos son estáticos o dinámicos?**

Los datos son registros estructurados que describen los atributos de diferentes canales de YouTube (número de suscriptores, vistas, uploads, etc.) para analizar patrones, como la relación entre suscriptores y vistas.

El dataset contiene información sobre canales de YouTube, como:

- **Ranking:** Posición según popularidad.
- **Username:** Nombre del canal.
- **Subscribers:** Número de suscriptores.
- **Uploads:** Cantidad de videos subidos.

- **Views:** Total de vistas del canal.
- **Country:** País de origen del canal.

Los datos de este archivo son estáticos, ya que representan una captura en el tiempo. Sin embargo, los datos reales de YouTube son dinámicos y cambian constantemente.

- **¿Cuál es la granularidad y la calidad de los datos que se recopilan?**

La granularidad es a nivel de canal, no a nivel de video individual. Esto limita el análisis a tendencias generales del canal. La calidad de los datos nos indica que parecen completos y relevantes.

- **¿Cómo representamos mejor esa información? ¿Cuáles son las mayores implicaciones de este proyecto? ¿Qué tan grande es el mercado? ¿Quiénes son los otros jugadores en este mercado?**

Para analizar el incremento de vistas en función del número de suscriptores, sugiero:

1. **Gráficos:**

- **Scatter plot:** Vistas por video (Vistas/Uploads) frente a suscriptores.
- **Boxplot:** Comparar la distribución de vistas entre grupos de canales según el tamaño de suscriptores (e.g., pequeño, mediano, grande).

2. **Resumen tabular:**

- Agrupaciones por país o rango de suscriptores.

Las implicaciones serán identificar prácticas exitosas en canales con pocos suscriptores, pero alta relación vistas/subida y ayudar a diseñar estrategias para maximizar ingresos pasivos con videos “faceless” en base a la relación entre número de suscriptores y crecimiento en vista.

El mercado de ingresos pasivos en YouTube es masivo, con ganancias mayores a \$25 mil millones anuales con un creciente número de canales “faceless”.

Los otros jugadores en el mercado son: Nuevos creadores que son miles de personas que saben generar contenido con IA y herramientas automáticas.

- **¿Qué herramienta de BI utilizarías?**

Power BI: Para crear dashboards visuales interactivos con métricas clave.

- **Se utilizará Power BI como herramienta de visualización**