

Challenge Técnico

En este Challenge se tendrá en cuenta:

- Limpieza de datos.
- Gráficos usado
- Claridad de la información.
- Visualización y presentación.
- Análisis de los datos.
- Comentarios.
- Manejo Datos duplicados / faltantes.
- Funcionalidad del dashboard.
- Métricas utilizadas.

OBJETIVOS

Parte 1 Análisis de datos.

A partir de la siguiente fuente de datos desarrollar:

- https://cdn.buenosaires.gob.ar/datosabiertos/datasets/ministerio-de-economia-y-finanzas/buenos-aires-compras/bac_anual.c

El diccionario de datos del dataset propuesto es:

• https://cdn.buenosaires.gob.ar/datosabiertos/datasets/ministerio-de-economia-y-finanzas/buenos-aires-compras/Metadata_OCDS.xl

1. Se solicita al participante crear un notebook en python con el cual se normalicen, y limpien los datos.
2. Generar un archivo con una inserción inicial de los datos y un DAG en Airfl para las actualizaciones recurrentes. (Opcional)
3. Generar un script que suba a un repositorio el archivo un archivo en Excel con los datos que se solicitan en el punto anterior.
4. Responder de la forma que considere más adecuada:
 - a. Montos totales operados en el periodo, desglosado por rubro y repartición.
 - b. Montos comprometidos por empresas.
 - c. Montos por tipo de contratación.
 - d. Evolución de ejecución de presupuesto.
5. Crear un dashboard en <https://datastudio.google.com/> o similar mostrando los datos del punto 4.

Parte 2 Modelos.

Utilizando el siguiente dataset:

1. Filtre únicamente las propiedades en Argentina.
2. Utilice algún modelo para clasificar las propiedades en venta y alquiler q tome como variables de entrada las características de la propiedad.
3. Ajuste un modelo para describir el precio en función de las características de una propiedad.

ESPECIFICACIONES

- Formato Fecha: 2022-08-16. <YYYY-MM-DD>
- <https://datastudio.google.com/> (puede usar el sistema de visualización que desee siempre que se pueda acceder de forma pública.)