

Data & Analytics Challenge



Challenge - Analytics Engineer



Candidato: Putrino, Rodrigo Nicolás

Fecha de inicio: 03/05/2024

Fecha de entrega: 07/05/2024

Índice

Índice.....	1
Challenge - Engineer - Segunda Parte - Python y APIs.....	2
Objetivo.....	2
1) Barrer una lista de más de 150 ítems ids en el servicio público:.....	2
2) Por cada resultado, realizar el correspondiente GET por Item_Id al recurso público:....	2
3) Escribir los resultados:.....	2
4) Elaborar el diseño y la documentación de la solución:.....	2
5) Análisis exploratorio:.....	2
Resolución.....	3
Diseño de la solución.....	3
Desarrollo y análisis exploratorio.....	4

Challenge - Engineer - Segunda Parte - Python y APIs

Objetivo

Realizar un análisis sobre la oferta/vidriera de las opciones de productos que responden a distintas búsquedas en el sitio Mercadolibre.com.ar utilizando el lenguaje Python y las bibliotecas que considere necesarias.

1) Barrer una lista de más de 150 ítems ids en el servicio público:

<https://api.mercadolibre.com/sites/MLA/search?q=chromecast&limit=50#json>

En este caso particular y solo a modo de ejemplo, son resultados para la búsqueda “*chromecast*”, pero deberás elegir otros términos para el experimento que permitan enriquecer el análisis en un hipotético dashboard (ejemplo Google Home, Apple TV, Amazon Fire TV, o afines para poder comparar dispositivos portátiles, o bien elegir otros 3 que te interesen para comparar).

2) Por cada resultado, realizar el correspondiente GET por Item_Id al recurso público:

https://api.mercadolibre.com/items/{Item_Id}

3) Escribir los resultados:

Escribir los resultados en un archivo plano delimitado por comas, desnormalizando el JSON obtenido en el paso anterior, en tantos campos como sea necesario para guardar las variables que te interesen modelar.

4) Elaborar el diseño y la documentación de la solución:

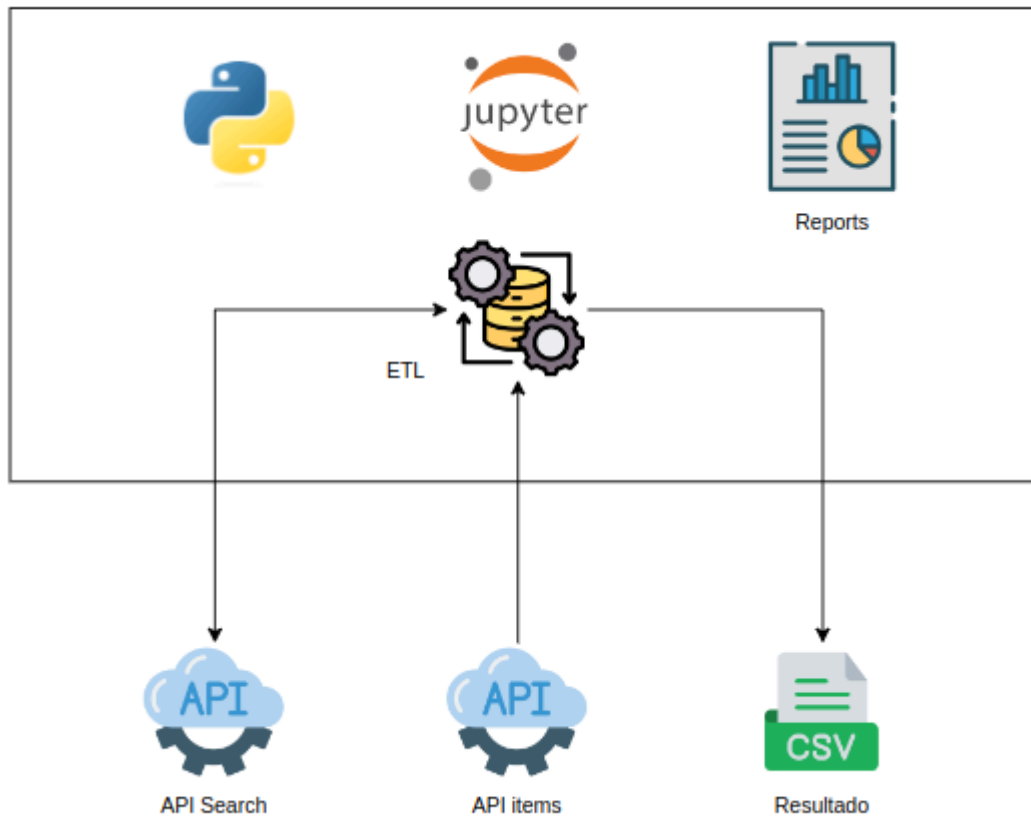
Presentar una solución para este escenario elaborando un diagrama de alto nivel de la solución y documentando los pasos necesarios para lograr este objetivo.

5) Análisis exploratorio:

Realizar un análisis exploratorio con las variables seleccionadas para el modelo a través de un notebook jupyter.

Resolución

Diseño de la solución



La solución fue desarrollada en Google Colab, a fines didácticos para agilizar su desarrollo y poder realizar el análisis exploratorio en el mismo.

Se compone de las siguientes etapas:

1. Ingesta y Limpieza de Datos
 - a. Se realizan llamados a la API de búsqueda para cada uno de los elementos de interés utilizando hilos para reducir los tiempos de ejecución.
 - b. Se consulta la API de ítems para obtener detalles específicos de cada ID obtenido en las llamadas anteriores.
 - c. Se agrega un campo que indica de qué elemento se trata.
 - d. Se crea un DataFrame de pandas con la información recolectada.
 - e. Se realiza una limpieza en el DataFrame, manteniendo solo las columnas de interés.
 - f. Finalmente, se escribe el DataFrame en un archivo .csv para su posterior análisis y visualización.
2. Análisis exploratorio

- a. Se manipulan los datos para realizar algunos análisis de interés.
- b. También se utiliza el módulo `ydata_profiling` para generar reportes automáticos que incluyan estadísticas resumidas, visualizaciones y análisis detallado de los datos.

Desarrollo y análisis exploratorio

Para acceder al .ipynb haga click [aquí](#).