

Specialist, Business Intelligence, Consumer Electronics

Challenge

Data & Analytics

Bienvenido al Challenge de Data & Analytics de Mercado Libre. Agradecemos tu interés por formar parte de nuestro equipo, y nos gustaría que te enfrentarás a los desafíos con los que trabajamos en nuestro día a día. Por medio de esta prueba nos gustaría conocer cómo te desenvuelves en las diferentes etapas de la solución de un problema de negocio, desde el entendimiento del problema, el planteamiento de preguntas, la obtención y exploración de la información, el planteamiento y creación de la solución analítica y la presentación de tus resultados.

¡Te deseamos mucha suerte!

1. Instrucciones Generales

El objetivo de esta prueba es conocer la manera en que abordarías un caso de negocio, desde el entendimiento del problema hasta la creación de una solución analítica. No es necesario que tengas conocimiento particular de la industria en la que nos encontramos, pues lo que queremos evaluar es la forma en la que te desenvuelves en esta fase de Data Science. Te presentaremos unos casos de negocio y podrás escoger el que desees trabajar. Harás la extracción de la información que consideres necesaria para dar respuesta al problema, y construirás tu solución analítica. Finalmente debes realizar una presentación de 20 minutos en la que nos guíes a través del problema, la forma en que lo abordaste, tus análisis, presentación de resultados y conclusiones.

Adicionalmente, todos tus desarrollos deben estar contenidos en un repositorio público de GitHub

bajo tu usuario. Este repositorio debe poder ser clonado y utilizado para replicar tu código de extracción de información, análisis exploratorios y modelos. Es deseable que utilices Python en Jupyter Notebooks por presentación y organización.

2. Caso de Negocio

A continuación te presentamos un problema de negocio con el que nos enfrentamos día a día. Debes plantear una solución analítica de clusterización que te permita generar valor tanto a la unidad de negocio como a Mercado Libre en general.

- El equipo comercial quiere realizar estrategias focalizadas para los sellers, pero en este momento no existe una clasificación que permita identificar a aquellos que tienen un buen perfil y son relevantes para el negocio. ¿Cómo podrías ayudar al equipo comercial a identificar estos sellers?

3. Extracción de Información

La información de MercadoLibre es pública, de forma que cualquier persona puede consultar vía APIs detalles de ventas, envíos y estado de productos y sellers del Marketplace. A continuación te vamos a presentar algunos ejemplos que podrían serte de mucha utilidad para extraer la información que consideres pertinente.

API de Categorías

Es muy importante conocer las categorías de los ítems pertenecientes al marketplace. Para ello puedes consultarlas con el siguiente request:

```
import requests  
  
cats = requests.get('https://api.mercadolibre.com/sites/MLA/categories')  
  
cats.json()
```

Con esto vas a obtener las categorías de los productos pertenecientes a Mercado Libre Argentina. Así como la forma de identificar a Mercado Libre Argentina es con el id MLA, cada país tiene su id que puedes consultar [aquí](#). Para ver más sobre las categorías puedes ver más [aquí](#).

API de Ítems

Por otra parte, una vez conozcas las categorías de los productos puedes explorar los ítems de alguna categoría y para ello te recomendamos tener en cuenta un parámetro muy importante: el offset. Tomemos el siguiente request:

```
import requests  
  
cat_id = 'MLA1000'  
offset = 50  
url = f'https://api.mercadolibre.com/sites/MLA/search?category={cat_id}&offset={offset}'  
request = requests.get(url)  
items = request.json()
```

En este caso estamos usando el api para hacer una búsqueda de la categoría con id MLA1000 correspondiente a Electrónica, Audio y Video. El parámetro offset lo fijamos en 50 lo que nos devuelve los ítems de esta categoría que se encuentran en la posición 51 a 100 en el buscador del marketplace. Puedes ver más [aquí](#). Chequea el output resultante en la variable items y verás toda la información que podrás obtener de cada ítem.

Para extraer los datos que consideres pertinentes podrías iterar sobre las ids de las categorías de el(los) país(es) de tu interés y después iterar el parámetro offset de la siguiente manera: 0, 50, 100,..., 950 con lo que obtendrías los 1000 primeros ítems por búsqueda de cada categoría en el marketplace.

Sin embargo, este no es el único camino, si tienes otra u otras ideas que consideres pertinentes para extraer información de los ítems con el api de Meli, te invitamos y te animamos a que la implementes.

Nota: Puedes acceder a estas APIs directamente en el navegador para explorar la información. Por ejemplo, para la categoría anteriormente mencionada, puedes hacer click en la siguiente url y ver sus contenidos:

<https://api.mercadolibre.com/sites/MLA/search?category=MLA1000&offset=50>

Dentro de results encontrarás un json con la información de los 50 primeros ítems de la categoría. La descripción de cada atributo del ítem la puedes consultar en la referencia de la API o haciendo click en <https://api.mercadolibre.com/items/MLA1100359664#options>

4. Puntos Clave de Presentación

Una vez planteado y resuelto tu caso de negocio debes realizar la socialización de tus resultados. Algunos puntos a tener en cuenta son:

- *¿Qué problema se quiso resolver con el desarrollo de la solución?*
- *¿Qué información buscaste y utilizaste para el desarrollo de la solución? Incluye las hipótesis que te llevaron a descargar dicha información.*
- *Muestra los insights más relevantes que hayas obtenido de tu análisis exploratorio. ¿Cómo estos te guiaron a la hora de plantear la solución?*
- *Danos un contexto de la solución que escogiste, y cómo la abordaste. ¿Qué metodologías aplicaste? ¿Qué métricas de evaluación utilizaste?*
- *¿Cuál es tu solución final? ¿Cómo se comporta? ¿Cómo soluciona o ayuda a solucionar el problema de negocio?*
- *Concluye y plantea próximos pasos que harías después de entregar esta solución.*

Es importante que estos análisis estén respaldados por cifras, y que estas puedan ser consultadas y replicadas en los notebooks del repositorio de Github.

5. Información Adicional

Si encuentras algún problema o dudas, no lo pienses dos veces y contáctanos por medio de correo electrónico o los canales que te hayamos proporcionado cuando te enviamos esta prueba.