

CBT Data Analytics Challenge

Data & Analytics

Welcome to the Data Analytics Challenge of Mercado Libre. We appreciate your interest in being part of our team, and we would like you to face the challenges we work on in our day-to-day activities. Through this test, we would like to understand how you handle the different stages of solving a business problem, from understanding the problem, formulating questions, obtaining and exploring information, developing and creating the analytical solution, to presenting your results. We wish you the best of luck!

1. General Instructions

The objective of this test is to understand how you would approach a business case, from understanding the problem to creating an analytical solution. It is not necessary for you to have specific knowledge of the industry we are in, as what we want to evaluate is how you navigate each phase of the Data Science cycle. We will present you with several business cases, and you can choose the one you would like to work on. You will extract the information you deem necessary to address the problem and build your analytical solution. Finally, you must prepare a 20-minute presentation in which you guide us through the problem, how you approached it, your analysis, presentation of results, and conclusions. Prepare your presentation in a way that is easy to understand for anyone, regardless of their technical level.

Additionally, all your developments must be contained in a public GitHub repository under your user account. This repository should be clonable and usable to replicate your information extraction code, exploratory analyses, and models. It is preferable that you use Python in Jupyter Notebooks for presentation and organization.

2. Business Case

Below we present a business problem that we face on a daily basis. You must propose an analytical clustering solution that allows you to generate value for both the business unit and Mercado Libre as a whole.

- The commercial team wants to implement targeted strategies for sellers, but at this moment, there is no classification that allows them to identify those with a good profile and who are relevant to the business. How could you help the commercial team identify these sellers?

3. Information Extraction

The information from MercadoLibre is public, so anyone can access details of sales, shipments, and the status of products and sellers in the Marketplace via APIs. Below we will present some examples that could be very useful for extracting the information you deem relevant.

API de Categorías

It is very important to know the categories of the items belonging to the marketplace. You can consult them with the following request

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

With this, you will obtain the categories of the products belonging to Mercado Libre Argentina. Just as the way to identify Mercado Libre Argentina is with the id MLA, each country has its own id that you can consult [here](#). To see more about the categories, you can find more [information here](#).

API de Ítems

On the other hand, once you know the categories of the products, you can explore the items within a specific category, and for this, we recommend considering a very important parameter: the offset. Let's take the following 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()
```

In this case, we are using the API to search for the category with id MLA1000, which corresponds to Electronics, Audio, and Video. We set the offset parameter to 50 which returns the items in this category that are located in positions 51 to 100 in the marketplace search. You can view more [here](#). Check the resulting output in [the variable items](#), and you will see all the information you can obtain for each item.

To extract the data you consider relevant, you could iterate over the ids of the categories of the country or countries of your interest and then iterate the offset parameter as follows: 0, 50, 100,..., 950, which would give you the first 1000 items per search for each category in the marketplace.

However, this is not the only approach; if you have another idea or ideas that you consider relevant for extracting information about the items using the Meli API, we invite and encourage you to implement them.

Note: You can access these APIs directly in the browser to explore the information. For example, for the previously mentioned category, you can click on the following URL and view its contents:"

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

Within results, you will find a JSON with the information of the first 50 items in the category. You can consult the description of each item attribute in the API reference or by clicking on <https://api.mercadolibre.com/items/MLA1100359664#options>."

4. Key Presentation Points

Once you have presented and solved your business case, you must share your results. Some points to consider are:

- *What problems are you trying to solve with the development of the solution?*
- *What information did you search for and use in the development of the solution? Include the hypotheses that led you to download that information.*
- *Show the most relevant insights you obtained from your exploratory analysis. How did these guide you in proposing the solution?*
- *Give us context about the solution you chose and how you approached it. What methodologies did you apply? What evaluation metrics did you use?*
- *What is your final solution? How does it behave? How does it solve or help to solve the business problem?*
- *Conclude and outline the next steps you would take after delivering this solution."*

It is important that these analyses are supported by data, and that this data can be consulted and replicated in the notebooks of the GitHub repository.

5. Additional Information

If you encounter any problems or have questions, don't hesitate to contact us via email or the channels we provided to you when we sent you this test. te enviamos esta prueba.