



Mercado Libre Data Engineering Challenge

You have to gather information about items published in an ecommerce site, save it in a DB and launch alerts if certain criteria are met. This pipeline should be implemented in airflow and run daily.

To do this you will need to interact with the public API from the site. Here is some information you will need:

List of categories for a given site: <https://api.mercadolibre.com/sites/MLA/categories>

Specific category information: <https://api.mercadolibre.com/categories/MLA1577>

Search api for a given category:

<https://api.mercadolibre.com/sites/MLA/search?category=MLA1577#json>

Specific item information: <https://api.mercadolibre.com/items/MLA830173972>

Task 1: Create a data pipeline to gather items information and save it in a database
From the MercadoLibre site, get the 50 most relevant published items, for a particular category,

"MLA-MICROWAVES" (category id MLA1577).

For each item get the following info:

- "id"
- "site_id"
- "title"
- "price"
- "sold_quantity"
- "thumbnail"

Store all this data with an extra field "created_date" in a Database. This simple Data pipeline must be implemented using an Airflow Dag.

Task 2: Send an alert vía mail

In the same Dag developed for the previous task, each time the data gathering task runs, check

if any item has earned more than 7.000.000 \$ (price x sold_quantity) and if so send an email with all the gathered data for every item that meets the criteria. The mail can have any format as long as the info is there.

Notes:

- The output format is not specified.
- Any libraries or tools needed are also accepted.

Bonus:

- *When your code is deployable or runs locally
- *Any form of unit/E2E testing
- *Additional metadata or data lineage information
- *Any form of automation
- *Design and documentation