

Code Challenge Senior Data Analyst

Adevinta

(this challenge is confidential, please don't share it externally)

Here attached you can find a dataset (`rating_per_user.json`), where every row represents the collection of received ratings related to transactions from a seller to a buyer (or the opposite), during the same time period.

Every row is a different user. You can find the description of the attributes of the dataset in the file `dataset characteristics`.

Exercise 1 - Analysis

Perform an exploratory analysis using Python to try to understand the dataset.

In order to reduce the scope of the exploration, you can focus on the following questions:

1. What is the period where these ratings were collected?
2. What is the distribution of the number of received ratings per user?
3. What is the average of the rating value?
4. What differences are there between ratings SellerToBuyer and BuyerToSeller?

Expectations

For this part, we are expecting a clear and reproducible notebook in Python, answering all the previous questions in a clear way.

Exercise 2 - Presentation

Prepare a presentation with a summary of all your findings and insights.

When preparing it, imagine presenting these results to non-technical stakeholders.

Expectations

For this part, we are expecting a presentation containing a summary of the findings and the suggested next steps.