

CAPSTONE PROJECT

2025B Data Science – Diretcho Trabajo

Your Role:

Junior Data Analyst at a large department store marketplace connecting small businesses from all over Brazil to channels without hassle and with a single contract.

Business Challenge:

The Marketing Director needs to optimize campaign spending. Currently, the team sends the same promotions to ALL customers, which is expensive and ineffective.

The team wants to launch targeted campaigns but does not know which customers to prioritize.

Goal: Build an RFM customer segmentation model and create a Dashboard for the stakeholders.

Deliverables:

By the end of this project, you will submit:

1. **Jupyter Notebook** (*groupnumber_rfm_analysis.ipynb*) with the ff:

- Data Loading and exploration
- Data Cleaning, validation and preparation
- RFM Calculation and Scoring
- Customer Segmentation
- Static Dashboard
- Business Insights and Recommendations

2. **Deck**

- Technical
 - Data and Process Done
- Business
 - Insights/Findings
 - Recommendations
 - Dashboard

3. **Clean Data Export** (*groupnumber_customer_segments.csv*)

- Final RFM Scores + Segment Labels for all customers

Learning Objectives

- **Data Wrangling:** loading multiple CSV, exploratory data analysis, data preparation and data validation
- **Feature Engineering:** Calculate Recency, Frequency, Monetary metrics
- **Business Logic:** Implement RFM scoring methodology
- **Analytical Thinking:** Justify segmentation decisions with data
- **Data Visualization:** Create an executive dashboard
- **Communication:** Translate technical findings into business recommendations

Final Note:

The dataset is a REAL data from a real company. It's messy, imperfect, and challenging - just like data you'll work with in the real world. Don't expect perfection. Your job is to make it clean enough to extract insights!

Good luck! 