Instacart - Customer Profiling and Regional Insights

Alexandru Cojocari | datadrivenalex@gmail.com | 16.08.2025

Part 1 - Data Enrichment and Analysis

- Import the merged 4.9 dataset into a fresh Jupyter notebook; treat PII before analysis.

- Create US regions from the State column (regions only) and segment customers.

- Assess spending differences by region by crossing Region with the spending flag.

- Create an exclusion flag for low-activity customers (< 5 orders) and export the filtered sample.

- Build a customer-profile variable using age, income, department_id goods, and number of dependents; optionally include orders_day_of_week and order_hour_of_day.

- Create a visualization that shows the distribution of profiles.

- Aggregate max, mean, and min for usage frequency and expenditure at the profile level.

- Compare profiles with regions and departments; produce suitable charts and export them.

- Tidy the script, export the final dataset, and save the notebook.

Part 2 - Final Report and Handoff

- Tab 1: include project details and cite data.

- Tab 2: complete the population flow.

- Tab 3: list the consistency checks (including the Crosstabs in Python check).

- Tab 4: document data wrangling and data-security measures addressed.

- Tab 5: list derived columns and their sources.

- Tab 6: include visualizations with commentary answering business questions.

- Tab 7: include results and recommendations for Instacart's marketing strategy.

- Tidy project folder and scripts; add contents lists; save report in Sent to client.

- Move the Data folder out, compress project to .zip for submission, then move Data back.

Thank You!

Alexandru Cojocari

datadrivenalex@gmail.com