

PORTFOLIO PRESENTATION

Bouasker Oussama

DATA ANALYST CONSULTANT

28/ JANUARY, 2026

P3-Rockbuster Stealth Analysis



Business Challenge

Rockbuster Stealth must shift from physical rentals to online streaming to stay competitive. The BI team needs clear, data-driven insights to guide this transition.

Objectives

The objective of this project is to load Rockbuster's data into a PostgreSQL database, explore it using SQL, and understand the structure through an ERD and data dictionary. I then analyze movie performance, customer behavior, and regional sales patterns to answer key business questions and support Rockbuster's online-platform strategy.

Data Overview

Database includes film inventory, customers, rentals, and payments. Several related tables were loaded into PostgreSQL for efficient querying and analysis.

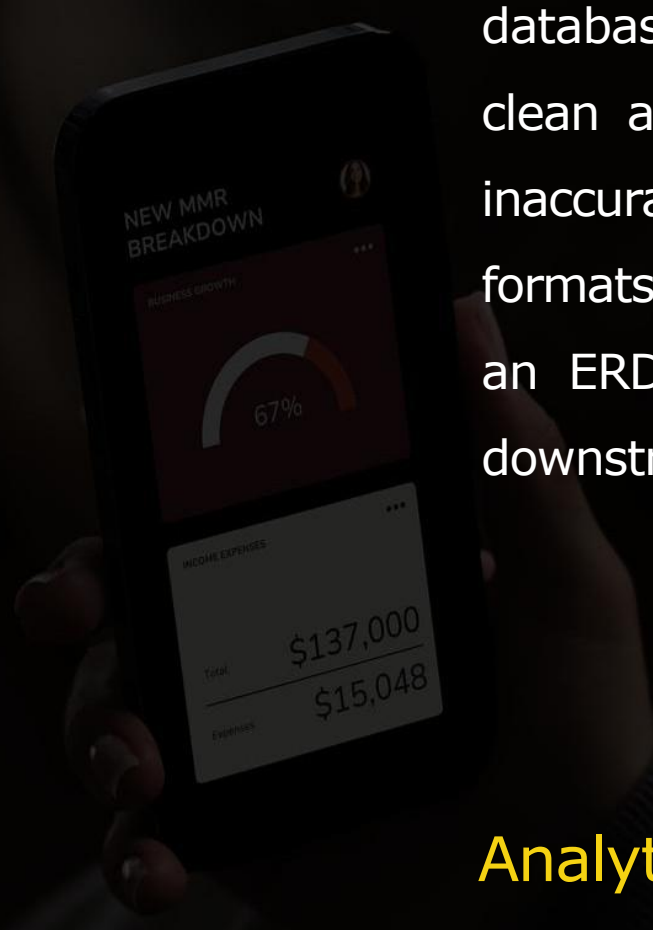
It contains 1000 films, 599 clients spanning 109 states

Data Cleaning

The Rockbuster Stealth dataset was imported into a relational database to ensure proper structure and integration. I used SQL to clean and standardize the data—removing duplicates, correcting inaccurate values, resolving inconsistencies, and harmonizing formats—so the database could support reliable analysis. I also built an ERD in DbVisualizer to clarify table relationships and guide downstream analytical work.

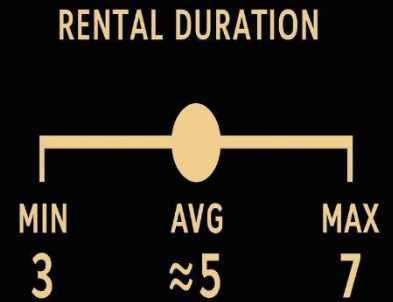
Analytical approach

Several SQL techniques were used to prepare the dataset for analysis, including JOINS to merge customer, film, rental, and payment tables; subqueries for targeted metric extraction; CTEs to streamline multi-step transformations; and aggregations to generate key analytical variables. These steps produced a clean, well-structured, and analysis-ready dataset that supports efficient insight generation.



ANALYSIS & METHODS

Descriptive statistics



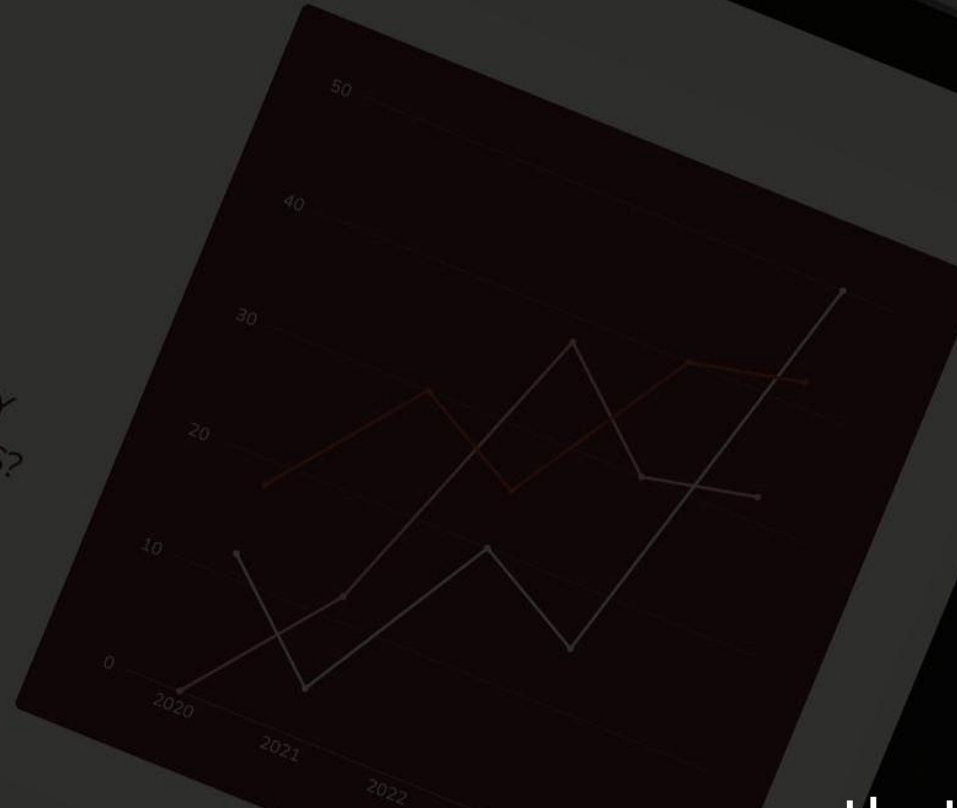
**ACADEMY
DINOSAUR
IS THE MOST
RENTED**

Most rented film is obtained by calculating the mode
Rental duration: Average of 5 days (min 3, max 7)

Popularity \neq Profitability: After searching, telegraph voyage was the most contributing film, but Academy Dinosaur is the most rented film.

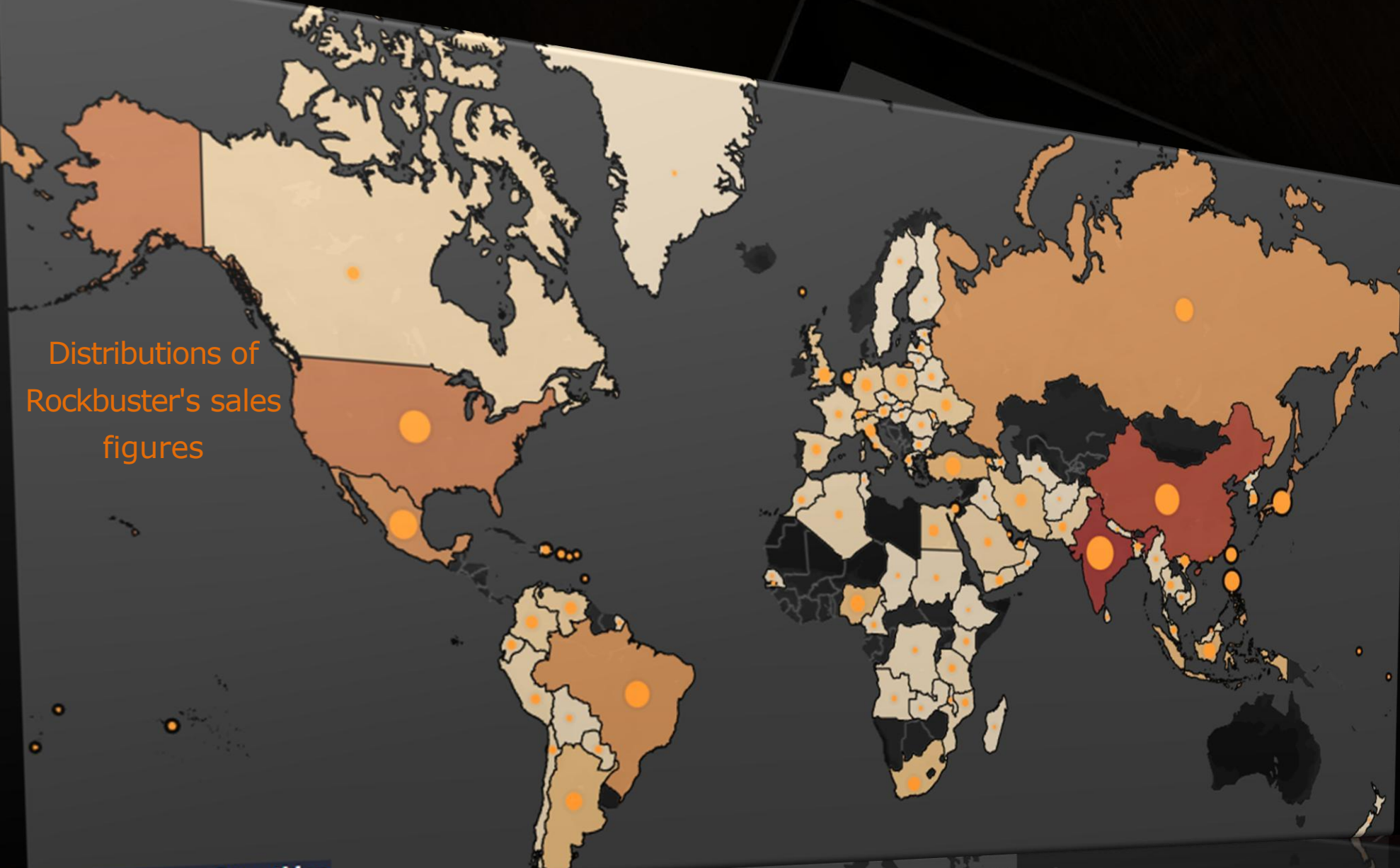
INDUSTRY BACKGROUND

WHAT IS THE INDUSTRY'S HISTORY AND WHAT ARE ITS USUAL TRENDS? DO YOU SEE NEW PATTERNS DEVELOPING? GIVE A PREDICTION OR OUTLOOK ABOUT WHERE THE INDUSTRY IS HEADED.



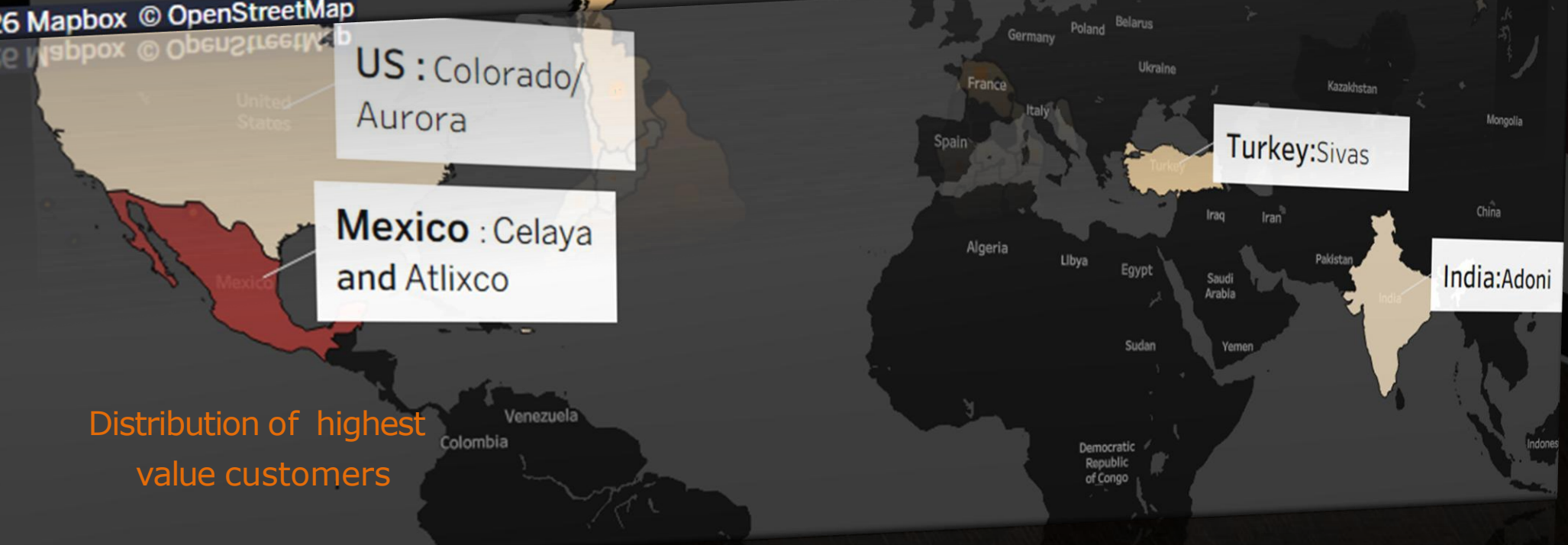
we can say that the most rented film is not the most profitable and frequency alone is not enough to generate revenue

METHODS AND ANALYSIS



After extracting the required data via SQL, we built a spatial analysis of sales and top-value customers using Tableau Public.

Rockbuster's sales are heavily concentrated in North America (United States), South America (Brazil), and Asia (India, China), where both customer volume and revenue are strongest. Europe shows fragmented performance, while Oceania (Australia) and Africa remain largely untapped markets. This uneven distribution highlights clear strategic growth opportunities.



High-value customers are also globally dispersed. Aurora (USA) reinforces the importance of the North American market; Celaya and Atlixco (Mexico) demonstrate Latin America's potential; Sivas (Turkey) reveals a profitable niche in Eurasia; and Adoni (India) confirms Asia's critical role in Rockbuster's customer base.

FINDINGS AND ACTION POINTS

The profitability of films and customers isn't solely based on their popularity or rental frequency. It depends on a balance between price, rental duration. Furthermore geographic distribution reveals a high concentration of revenue in certain regions (North America, Asia, Latin America), while other markets remain under exploited.

Target high-potential areas

strengthen offerings in countries like the United States, India, and Mexico, where high-value customers are already present.

Optimize prices and durations: adjust rates and rental conditions to maximize profitability, especially in high-activity areas.

Explore untapped markets : launch test campaigns in Australia or in some country in Africa to assess untapped potential.





Strengths Demonstrated

I was able to extract highly relevant insights while strengthening my SQL skills, especially by learning to handle more complex queries such as CTEs.



Moment of struggle

The SQL coding became increasingly complex, and at one point I lost track of the path to reach the results I wanted. But thanks to a methodical approach and steady persistence, I managed to overcome this high-level challenge and achieve the outcomes I was aiming for.



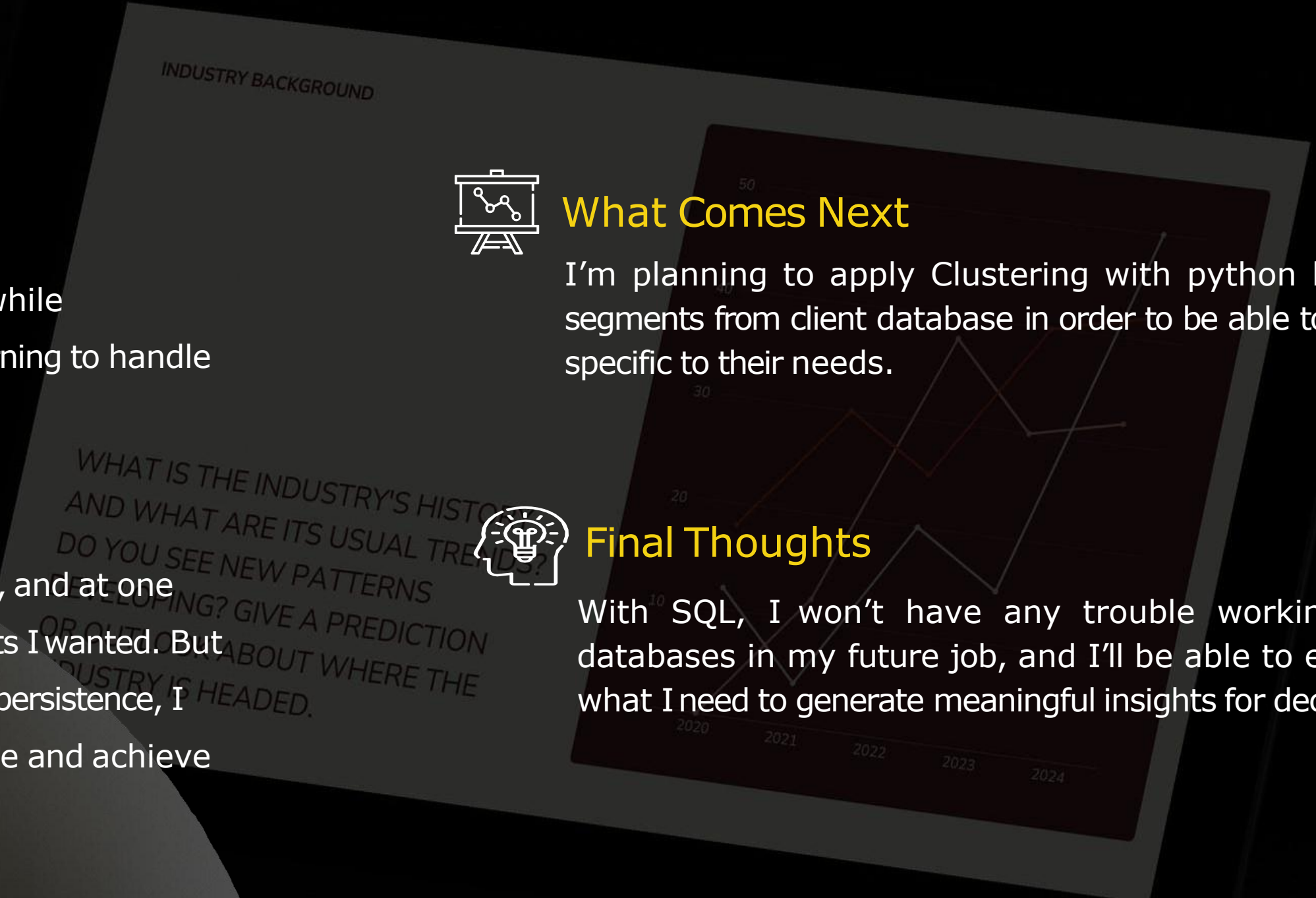
What Comes Next

I'm planning to apply Clustering with python by Creating new segments from client database in order to be able to serve customers specific to their needs.



Final Thoughts

With SQL, I won't have any trouble working with large databases in my future job, and I'll be able to extract exactly what I need to generate meaningful insights for decision-making.



P-4 Instacart Grocery Basket Analysis

Business Challenge

Instacart wants to better understand customer behavior to improve targeted marketing and sales performance. With thousands of products and diverse customer profiles, the company needs clear insights on when people order, what they buy, and how different customer groups behave.

Objective

- Analyze Instacart's large datasets using Python (Pandas, NumPy, os, Matplotlib, Seaborn, Scipy)
- Perform EDA to uncover ordering patterns, spending habits, and product trends.
- Identify customer segments based on loyalty, region, age, income, and family status.
- Provide insights to support marketing segmentation and sales optimization while relying on consumer behavior

Data overview

This project uses several Instacart datasets, including:

- Orders (timing, frequency)
- Products (names, departments, prices)
- Order-product relationships
- Customer demographics

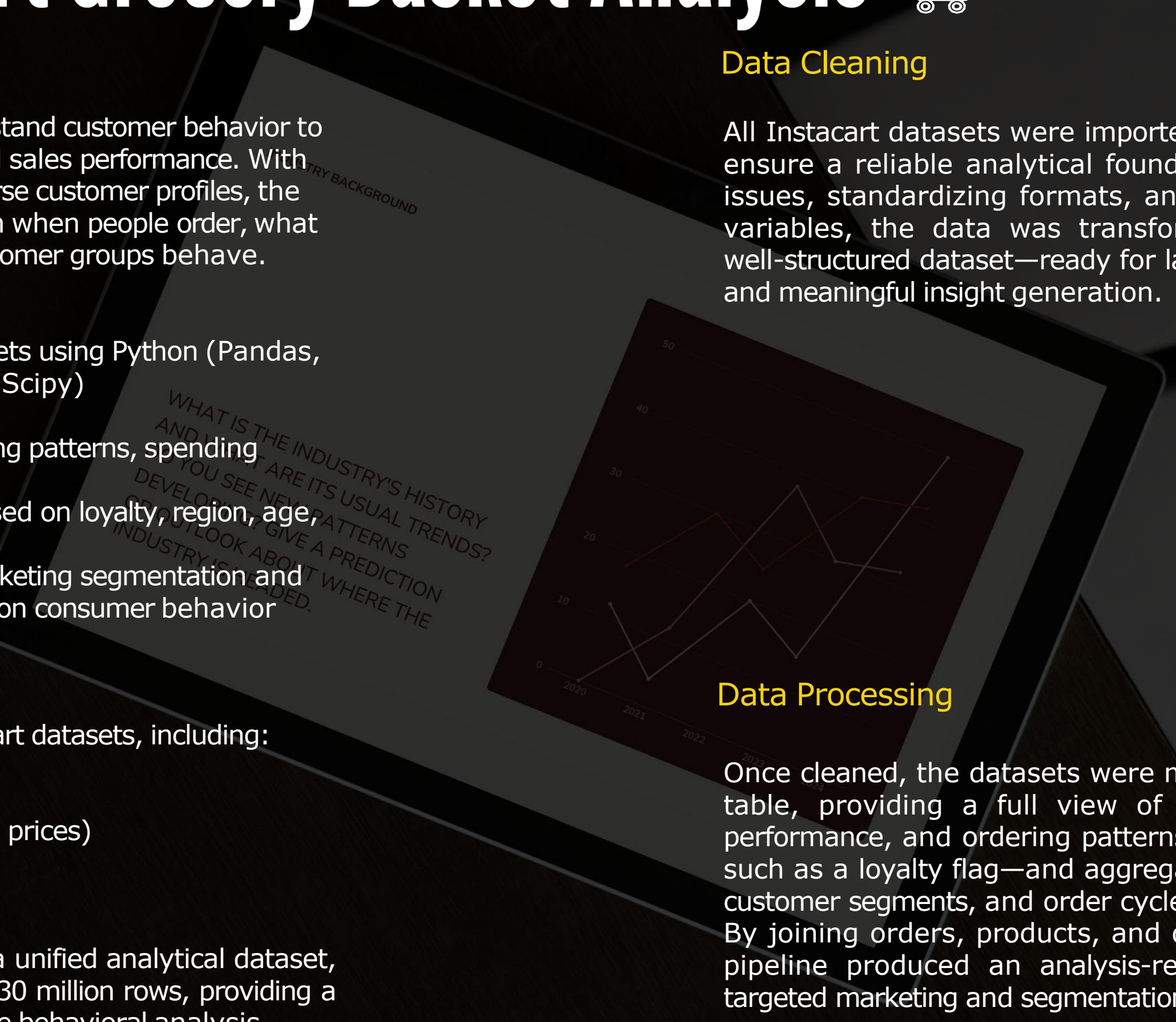
After merging all sources into a unified analytical dataset, the final table expands to over 30 million rows, providing a robust foundation for large-scale behavioral analysis.

Data Cleaning

All Instacart datasets were imported into Python and cleaned to ensure a reliable analytical foundation. After resolving quality issues, standardizing formats, and isolating the most relevant variables, the data was transformed into a consistent and well-structured dataset—ready for large-scale exploratory analysis and meaningful insight generation.

Data Processing

Once cleaned, the datasets were merged into a single analytical table, providing a full view of customer behavior, product performance, and ordering patterns. I engineered key features—such as a loyalty flag—and aggregated data across departments, customer segments, and order cycles to reveal meaningful trends. By joining orders, products, and demographic information, the pipeline produced an analysis-ready structure that supports targeted marketing and segmentation strategies.



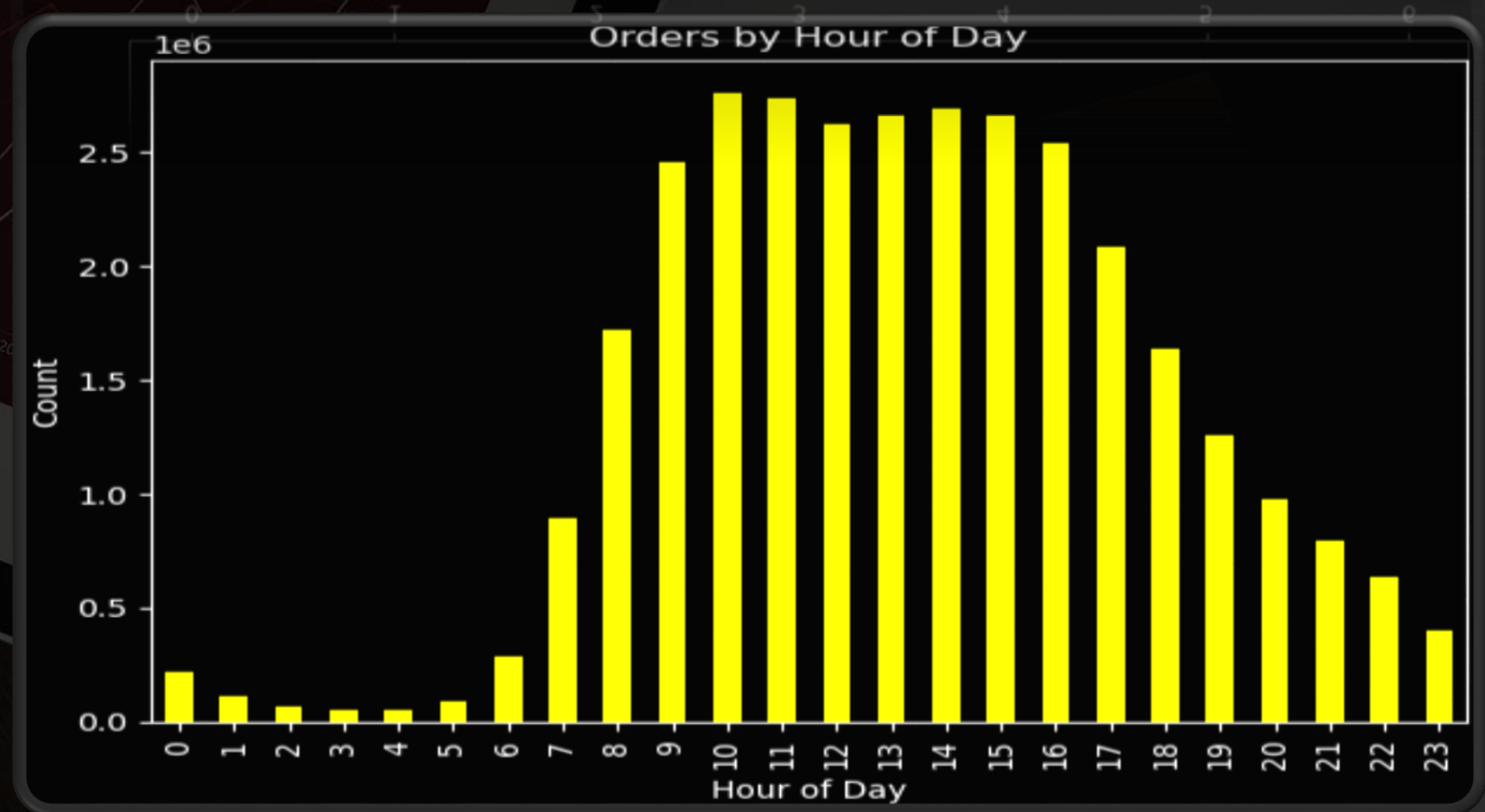
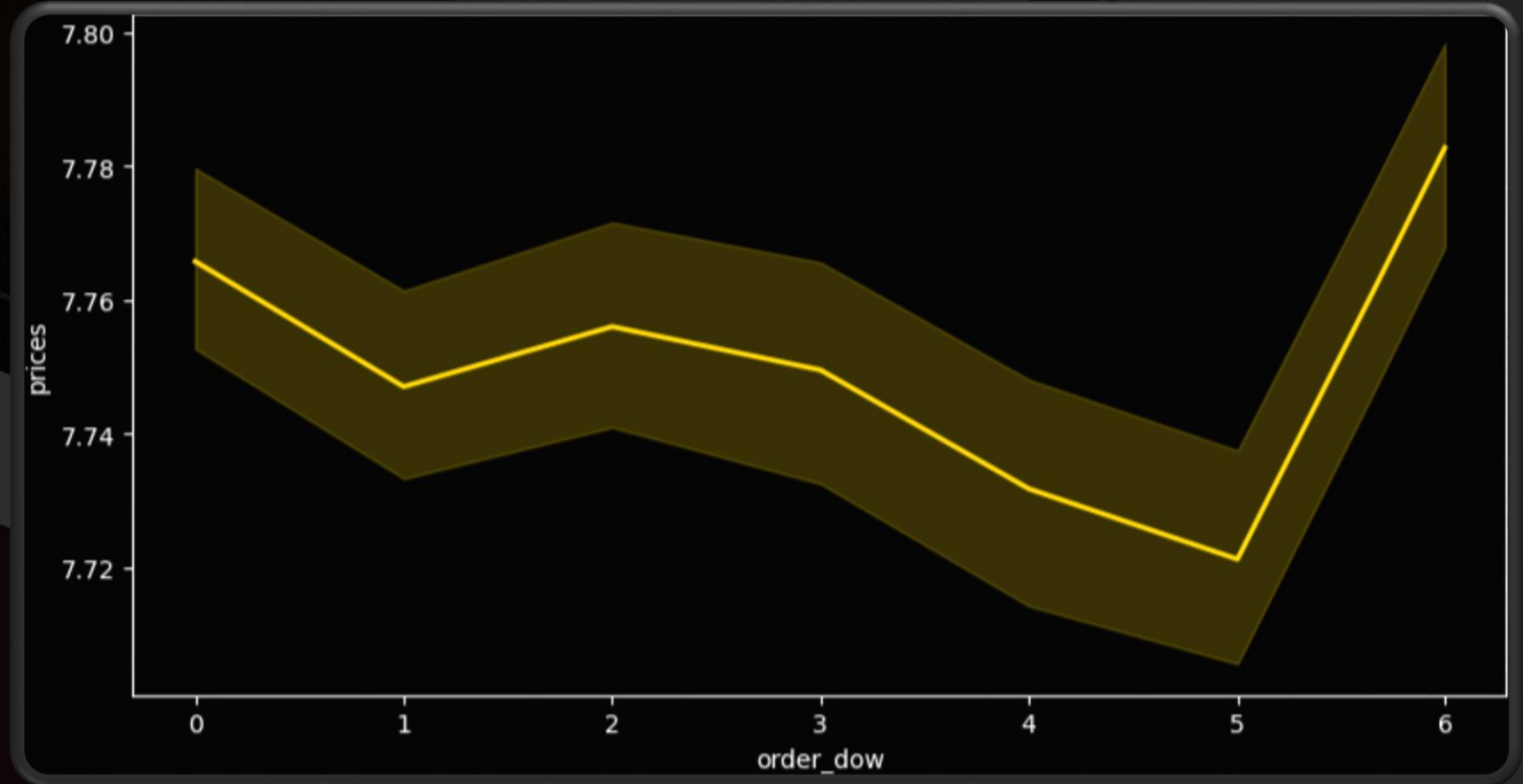
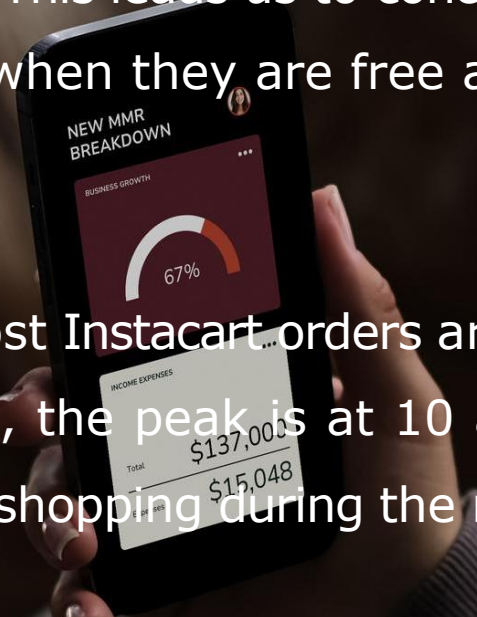
ANALYSIS & INSIGHTS

During the week, prices tend to be downward, rising over the weekend. This is explained economically by the fact that when demand increases, prices also increase. This increase is slight, and its variation increases during the week and decreases over the weekend.

During the day, prices rise in the morning but fall during peak demand. This can be explained by the promotions Instacart runs during peak times to boost sales.

Demand increases on Saturdays and Sundays and decreases during the week. This leads us to conclude that customers prefer to shop when they are free and have fewer commitments.

Order Time Distribution: Most Instacart orders are placed between 6 AM and 23 PM, the peak is at 10 am. This suggests customers prefer shopping during the middle of the day

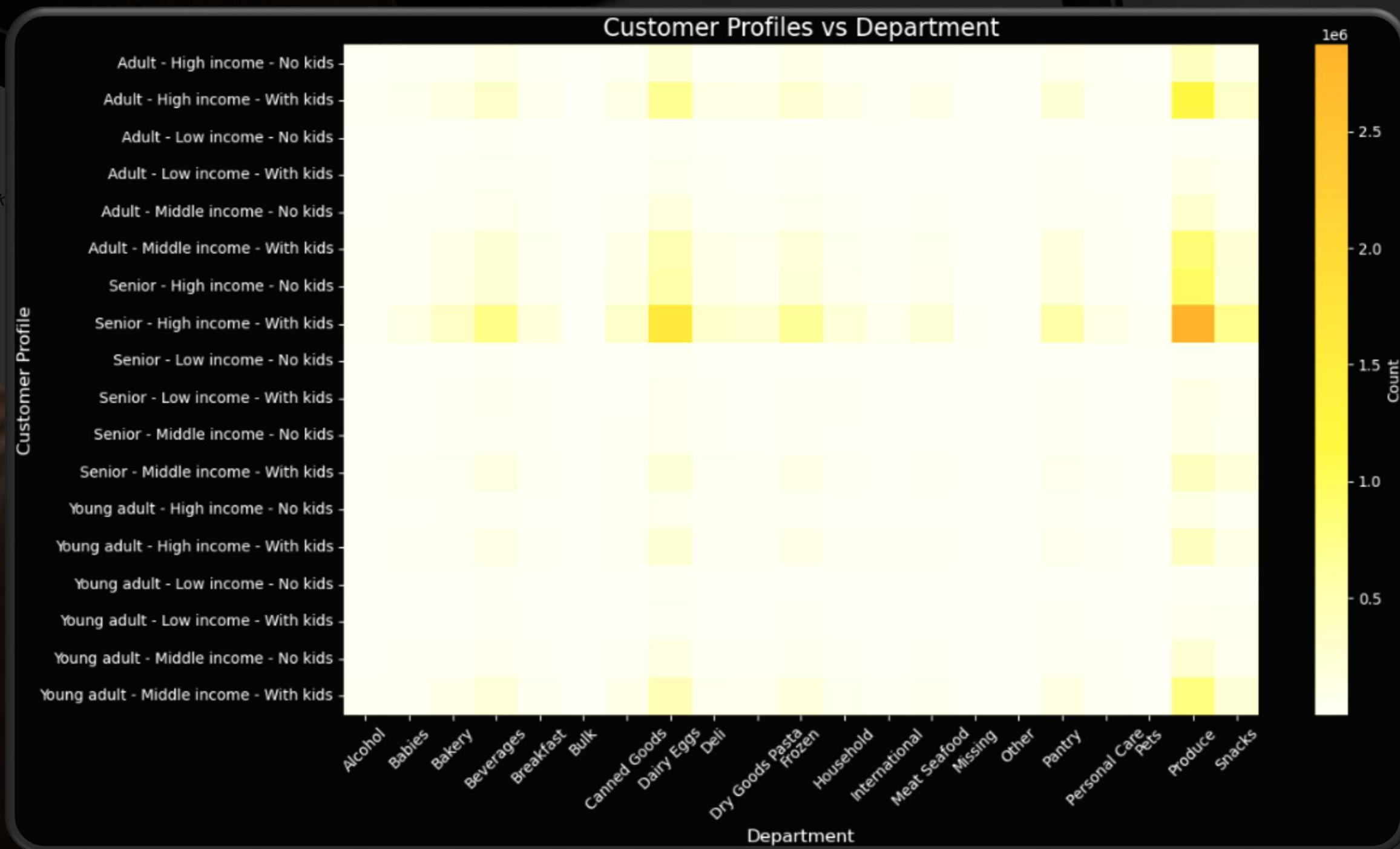


METHODOLOGY , ANALYSIS & INSIGHTS

Our analysis shows that most Instacart purchases fall under \$15, highlighting a strong preference for low-cost, everyday items. The most active departments—Produce, Dairy & Eggs, and Snacks—confirm that customers primarily buy fresh and quick-to-consume products, while categories like Pets, Alcohol, Babies, and International remain far less in demand.

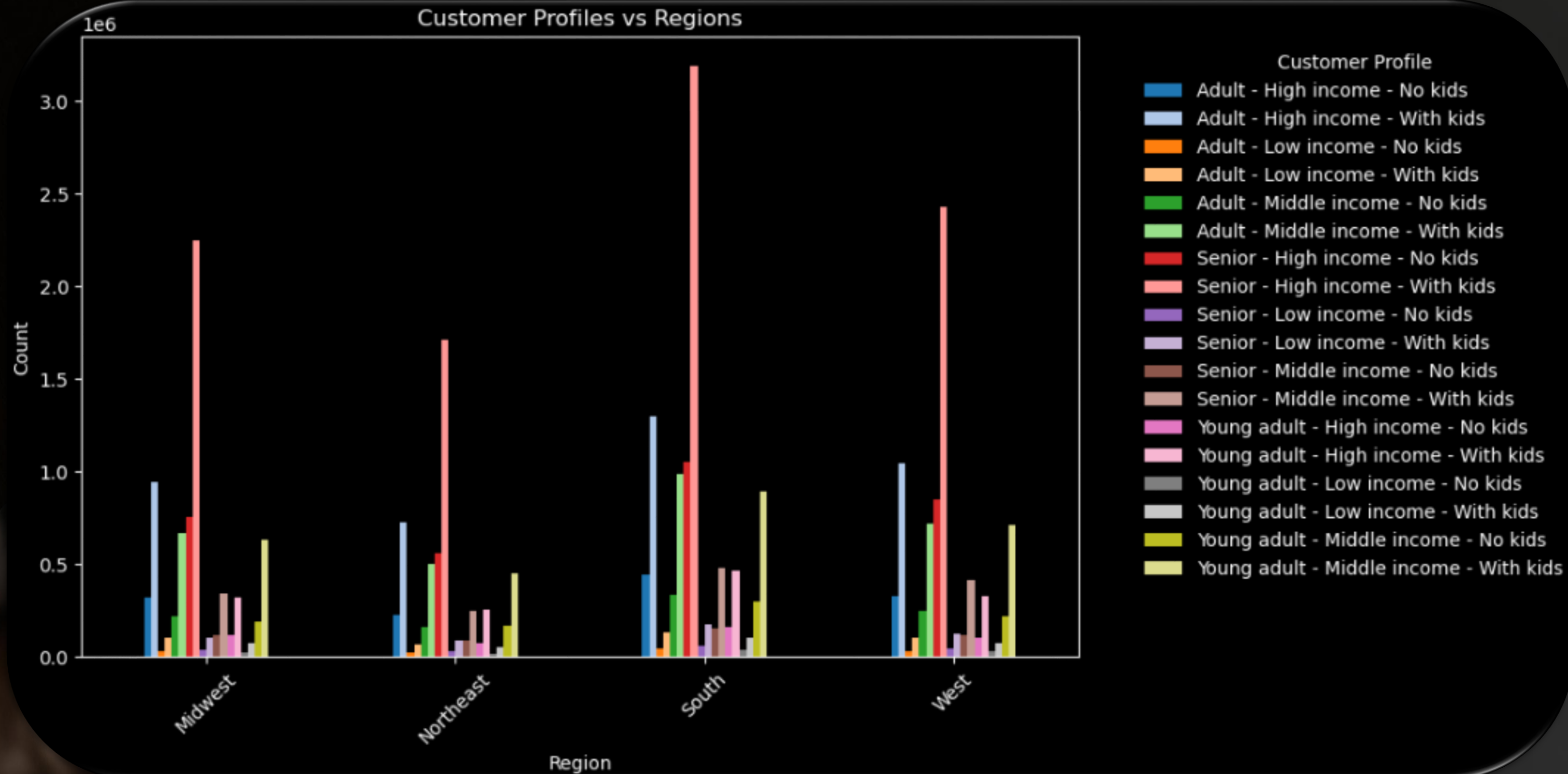
Regular customers generate the highest order volumes, forming Instacart’s most loyal and valuable segment. Building on these insights, we developed detailed customer profiles based on age, income, household size, purchasing patterns, and geographic distribution to better understand behavioral differences across U.S. regions and support more targeted marketing strategies.

High Income With Kids” customers concentrate most of their spending in Produce, Dairy & Eggs, and Snacks, reflecting fuller baskets and broader family needs. In contrast, “Low Income With Kids” customers show a smaller, more dispersed purchasing pattern, with a stronger focus on essential categories such as Pantry, Beverages, and Frozen Foods—indicating more selective and lower-volume shopping behavior.



when comparing profiles with departments, these same groups show a clear preference for essential and high-volume categories such as produce, dairy eggs, and snacks, while lower-income or no-kids profiles display more selective purchasing patterns focused on basic or lower-cost departments.

ANALYSIS & INSIGHTS

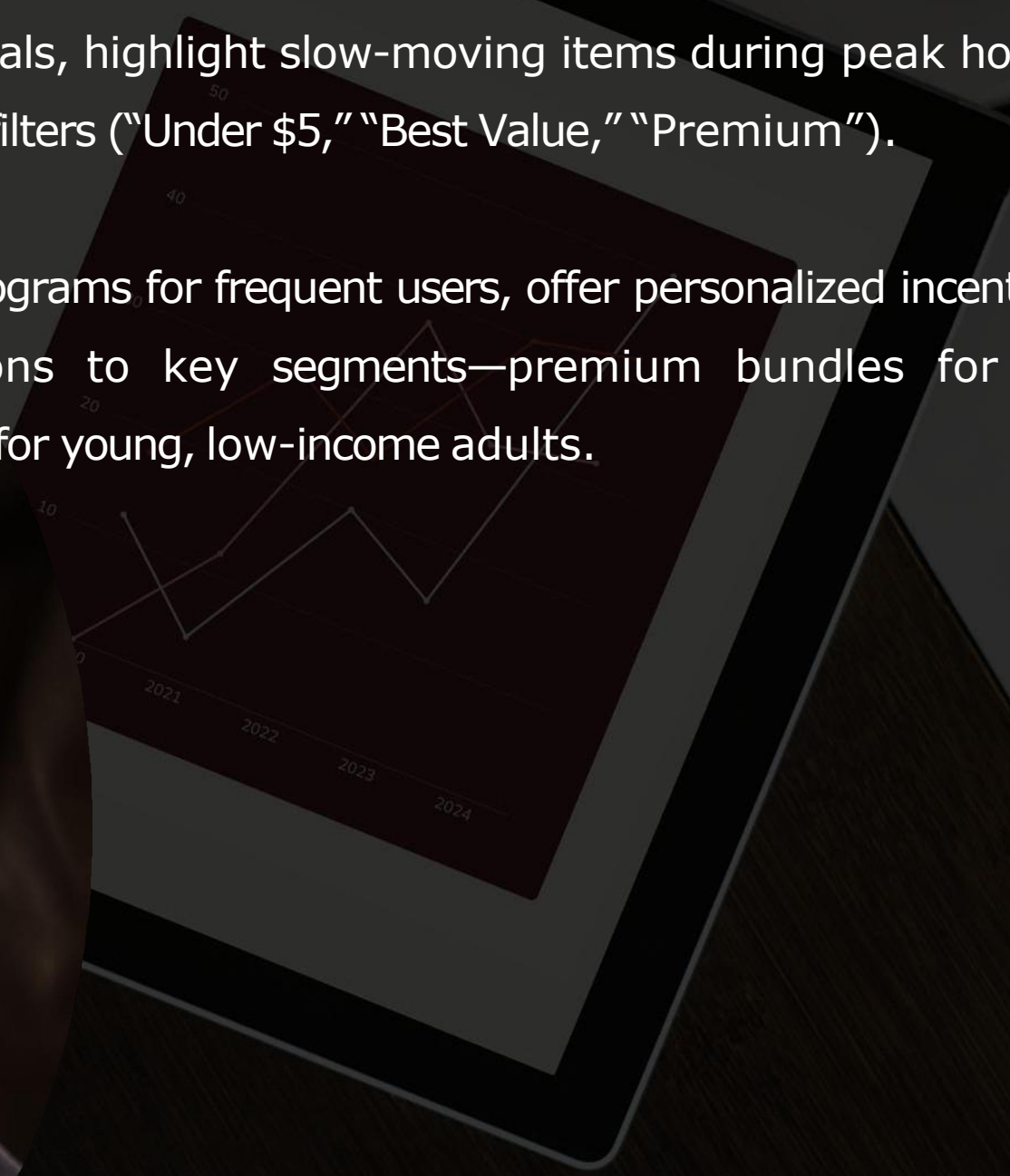


The "Senior - High income - With children" profile is overrepresented in the South, indicating a high population concentration. The other regions show a more balanced distribution between profiles, with less overall intensity.

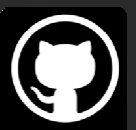
Across regions, the most represented customer profiles (notably High Income – With Kids) appear consistently dominant, showing that affluent family households form a strong core of Instacart's user base regardless of geography.

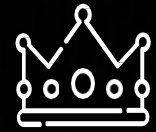
RECOMMENDATIONS

- Focus marketing efforts on high-potential regions (South, West, Midwest) and tailor campaigns to local profiles. Boost off-peak sales with targeted weekday ads and evening promotions. Promote fresh essentials, highlight slow-moving items during peak hours, and introduce clear price tiers with simple filters (“Under \$5,” “Best Value,” “Premium”).
- Strengthen loyalty programs for frequent users, offer personalized incentives for new customers, and adapt promotions to key segments—premium bundles for affluent families and budget-focused offers for young, low-income adults.



To see more





Strengths Demonstrated

With Python, I have mastered the fundamental techniques required for the daily work of a data analyst, such as merging datasets, data cleaning, grouping, and data visualization



Moment of struggle

Despite the high performance of my computer, it frequently froze due to insufficient memory to handle a dataset of more than 30 million rows. Under this constraint, I initially felt stressed about not being able to produce convincing results or visualizations.

However, I managed to overcome this challenge and deliver meaningful visualizations without compromising the quality of my decisions. This was possible thanks to my knowledge in computer science, data analysis, statistics, and the skills I developed throughout the course.



What Comes Next

As next steps, I would focus on applying clustering techniques to create new customer segments from the client database, allowing the company to tailor its services more precisely to customer needs.

I would also use regression models to predict the likelihood of customer churn and to estimate Customer Lifetime Value helping the business understand how much each customer is expected to contribute over the course of their relationship with us



Final Thoughts

I'm really happy that I now master the famous Python language—the one my friends in IT are always talking about. What makes me even happier is that I've learned to use a strong and reliable Python ecosystem that enables me to extract meaningful insights.