

AUGUST 12, 2023



Assignment m10.e1.1

ASHLEY LAU

Table of Contents

1	EXECUTIVE SUMMARY.....	2
---	------------------------	---

Table of Figures

Figure 1	Relationship diagram for Instacart metadata.....	2
Figure 2	Popularity by Item Overview dashboard.....	3
Figure 3	Days Since Prior Order Trendline.....	3
Figure 4	Reorder Items trend dashboard.....	4

1 EXECUTIVE SUMMARY

This document will explain the detailed approach towards analyzing popularity of items bought from Instacart users and interpret key insights based on overall popularity, time series analysis, and reordered item analysis. From these results, Instacart shoppers and grocery stores will have a better understanding of strategic actions that can be taken, whether it is restocking certain items ahead of time or ensuring more shoppers are available at certain hours of the day, to ensure customer service requirements are satisfied.

Before conducting any initial analysis, we first need to understand the data points provided. Fundamental metadata related to Instacart orders included order information, product data, previous customer order history, reference knowledge of respective aisle and department ids as well as a trained order dataset to compare with our test data. Another important aspect to complete this analysis is the usage of tools. Databricks was chosen Spark engine used to preprocess metadata, stored as csv files, to configure into a master spreadsheet for easier compute. Tableau was chosen as a visualization tool to analyze key data columns that allowed us to gain insights from metadata. A relationship diagram highlighting metadata relationships between data files is highlighted in Figure 1.

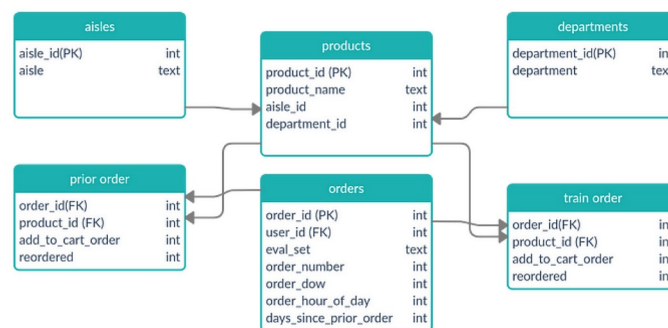


Figure 1 Relationship Diagram for Instacart metadata

First thing, some introductory metrics about the metadata to note from the Instacart data shows counts of 206,209 Instacart unique users placing over 3 million orders, amounting over 32 million items paid from a pool of 49,676 products coming from 21 departments and 134 aisles. In terms of order amount, the top 5 popular categories of items bought by customers were produce, dairy eggs, beverages, snacks, and frozen foods. The top ranked aisle, from respective department in Instacart orders, were fresh fruits, yogurt, sparkling water, chips/pretzels, and frozen produce. Not many customers were interested in buying pet supplies, or bulk packaged items. When organizing ranks of top 50 bestsellers from reference data, the products are mainly comprised of fruit or vegetables. Bananas were the number 1 selling item, followed by bag of organic bananas and organic strawberries. From the list of top 50 bestselling products, 64% of products are organic items. Additionally, Sundays are the busiest days customers stock up on items for their household. When looking at a view of products available across departments in stores, personal care, snacks, and pantry items have the broadest product selection; however, customers are majorly purchasing produce, dairy and

beverages. The least broad product selection were bulk packaged items, meat/seafood, and pet supplies. Interestingly, the lowest ranked department categories from order amount perspective were the same ranking as the least broad product selection. A snapshot of dashboard displaying item and time popularity is shown in Figure 2.

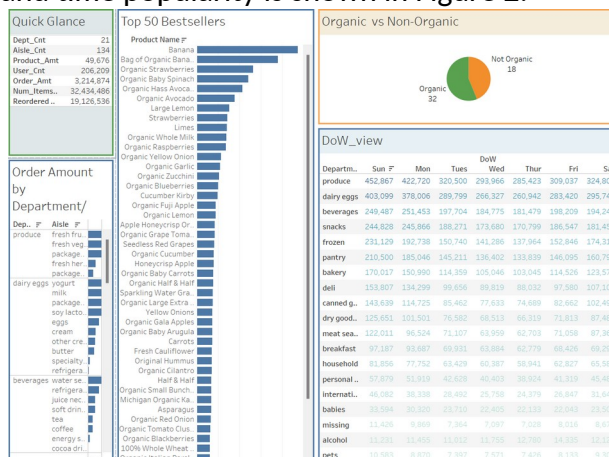


Figure 2 Popularity by Item Overview dashboard

From an hourly analysis, it can be found that customers, across the board, commonly place Instacart orders the most between 10am – 4pm every day. As previously mentioned, Sunday is the most popular shopping day followed by Mondays. Within this universal time, produce, dairy, and beverages continue to be the top 3 departments customers are purchasing items from in their orders. From a volume perspective, Fridays and the weekend are days where the broadest number of products are selected in customer's baskets; Sundays/Mondays attract the most number of orders (around 555,000 orders). Thursdays are the least popular day to place an Instacart order. A supplemental data point in our analysis covers days since prior order for customers that placed Instacart order more than once. Amongst customers that have made a previous order including products from baked goods, beverages, breakfast foods, canned goods, dairy, deli items, dry goods, frozen foods, household items, meat/seafood, pantry products, produce or snacks, the highest level of orders spikes after 7 days or 30 days. A snapshot of a similar trend for customers reordering products after 7 days or 1 month period is shown in Figure 3.

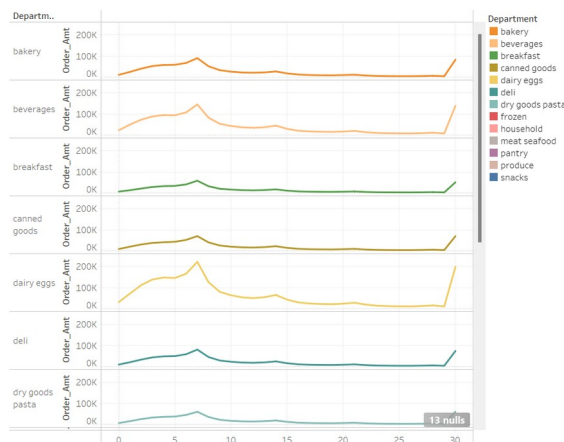


Figure 3 Days Since Prior Order Trendline

The last analysis conducted focused on customers that made more than one Instacart order. Out of 32 million plus items that were ordered, over 19 million items were ordered again. That is roughly 58% portion of complete orders in this database. Of all products available customers commonly restock on snacks, personal care items and pantry condiments. The least commonly restocked items were bulk packaged items, meat/seafood, and pet supplies; this dually aligns with least popular items from earlier analysis. The total number of items in a single Instacart reorder were bucketed in groups of 5; from this it is found that reorders are typically 10 items or less. A chart showing the trend of reordered number of items per order is shown in Figure 4.

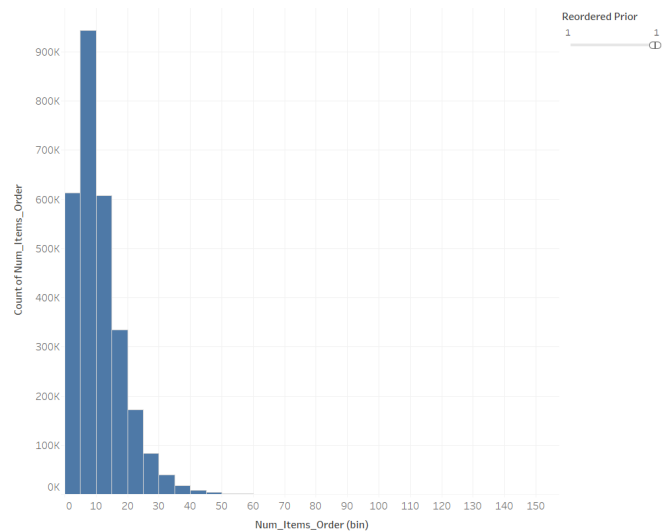


Figure 4 Reorder Items trend dashboard

In conclusion, the insights collected via Tableau offered invaluable observations of product preferences and time periods when Instacart users would place an order. To get a deeper understanding of these popularity stats, possible follow up steps to tie relationships of these popular product purchases are finding more data attributes related to customer as well as dissecting customer segmentation and market basket analysis. What are the age range and demographic of Instacart customers? What grocery store did user place order from; how close/far is grocery store relative to their marked address? Are customers placing requests for same day order and pickup, if not how many days in advance are customers placing their order? Sample questions, like these, can help Instacart's strategy team brainstorm ways to attract and retain customer base.

The subsequent parts of the project will go deeper into customer segmentation and leverage metrics of lift, support, and confidence to understand association rules that mathematically determine likelihood of customers purchasing select products in sequence.