

## 2023 MMA Datathon

### Case Challenge: Optimizing Grocery Delivery

#### Background

On-demand grocery delivery and pickup service has grown rapidly to become one of the major ways in which Canadians shop for groceries. Popular services such as Instacart have revolutionized the way people shop for groceries by providing a convenient and efficient solution to the traditional brick-and-mortar shopping experience. These services typically operate through both mobile apps and websites, connecting customers of the retailer with personal shoppers employed/contracted by the service who handpick and deliver their orders from local grocery stores.

These services are either run by the retailers themselves or through a third-party service (e.g., Instacart) who partner with major chains such as Loblaws, Walmart or Costco. They allow customers to choose from a vast selection of groceries, including fresh produce, and household essentials, all from the comfort of their homes.

To maintain a competitive edge and improve customer satisfaction, these services need to focus on optimizing the speed at which its shoppers pick items from grocery retailers. Efficient and timely picking plays a crucial role in ensuring prompt deliveries and meeting customer expectations. Technology solutions such as barcode scanning devices or mobile apps provide personal shoppers with real-time inventory information and optimized routes within the store. This enables shoppers to locate items quickly, reducing the overall picking time.

During the process, the personal shopper navigates the store to locate and select the items on the customer's shopping list. They may follow an optimized route based on the app's suggestions or their knowledge of the store layout. As they find each item, the personal shopper scans or adds it to the virtual shopping cart within the app. Once the personal shopper has gathered all the items on the customer's list, they proceed to a designated area for checkout, which is typically separate from the regular store checkout counters.

When a shopper is unable to find a specific item during the shopping process, they will typically explore options for [substitution](#) or contact the customer to inform them about the unavailability. If a suitable substitute is available, the shopper will confirm with the customer whether they would like to proceed with the substitution. In cases where a substitution is not feasible or the customer prefers not to substitute, the shopper will communicate the unavailability to the customer and inquire about their preference for a refund or an alternative solution. Depending on the customer's choice, a refund for the missing item can be processed, either as a credit for future purchases or returned to the customer's payment method. Product substitutions are a major pain point since they tend to consume a lot of personal shopper time. Unfortunately, no data on product substitutions is recorded by personal shoppers.

## **MM&A's Supermarket**

You and your team have been brought in as consultants of MM&A's Supermarket, a national supermarket chain serving mostly urban and suburban areas in all ten provinces with leading market positions in Ontario and British Columbia. Most of MM&A's stores use the traditional [supermarket](#) or [hypermarket](#) format, offering a full line of groceries, meat and produce with additional departments for general merchandise, drug, and health & beauty, depending on the store format.

In recent years, MM&A has been partnering with an on-demand grocery delivery service Instabasket, to increase sales. To date, the integration with Instabasket has involved connecting MM&A's inventory systems, setting new price levels for Instabasket, and details on store locations and layouts. As volume from Instabasket sales have increased, the regular customer shopping experience in MM&A's most popular stores have been impacted because the Instabasket personal shoppers are rapidly moving between aisles to maximize their shopping efficiency. This has led to complaints by regular customers and simultaneously has slightly reduced in-store traffic.

MM&A management thinks that their stores are not optimized for the high demand from Instabasket personal shoppers. They would like to test a store format that includes a special Instabasket aisle that can be a one-aisle stop for the personal shoppers to fulfill their orders. This aisle would contain at most 1000 products due to space constraints and, at most, 100 refrigerated and 100 frozen items. The expected outcome is that most Instabasket orders could be primarily fulfilled by the top 1000 items in the designated aisle. Instabasket management also wonders if some suggested set of substitutable products can be identified and used by personal shoppers. To encourage customers to accept the substitutes, a special 5% discount would be offered to compensate for not fulfilling the original request. Selecting and identifying products that fulfill a large number of orders but that are also good substitutes for other products is desirable to maximize the productivity of this Instabasket aisle. The overall effect of this aisle should be to minimize the number of trips to other aisles that would disrupt regular in-store customers.

Your task is to analyze the given data and recommend the top 1000 products to place in the "Instabasket" aisle, given the above refrigeration and freezer constraints. The two metrics to optimize are (a) the number of orders that utilize the in-aisle items, and (b) the average % of items in each order that utilize in-aisle items, accounting for any identified substitutes.

For your final presentation, summarize your product selections (including refrigeration and freezer constraints), provide your selection reasoning and methodology, and provide any additional insights into the data. If recommending substitutes, please clearly indicate how you chose them and the reasoning behind it. The audience for the presentation will be the senior management team from MM&A.

## **Dataset**

Dataset based upon “*The Instacart Online Grocery Shopping Dataset 2017*”,  
<https://www.kaggle.com/competitions/instacart-market-basket-analysis/>.

## **Deliverables**

- A presentation of up to 15 slides in PPT or PDF format with appendix. While the Appendix is not limited in size, the relevance of every page in the Appendix and how it supports the main body of your presentation should be very clear to the audience (reader). Note that your presentation should be self-contained: the only reason a reader would open the appendix is to get more details on some of the results/visualizations presented in the main body; the reader should not need to open the appendix to understand the statements made in the main body of the presentation.
- You can use analytical tools of your choice.
- Submit presentation via Quercus website.

## Appendix A: Dataset Notes

mma\_mart.csv (100k orders, 987k rows)

| Column Name   | Description           | Type          | Example          |
|---------------|-----------------------|---------------|------------------|
| order_id      | order identifier      | numeric       | 1                |
| product_id    | product identifier    | numeric       | 49302            |
| product_name  | product name          | alpha-numeric | Bulgarian Yogurt |
| aisle_id      | aisle identifier      | numeric       | 120              |
| aisle         | aisle name            | alpha-numeric | yogurt           |
| department_id | department identifier | numeric       | 16               |
| department    | department name       | alpha-numeric | dairy eggs       |