Project Brief: Instacart Grocery Basket Analysis

Objective

You’re an analyst for an existing company, Instacart, an online grocery store that operates through an app. Instacart already has very good sales, but they want to uncover more information about their sales patterns. Your task is to perform an initial data and exploratory analysis of some of their data in order to derive insights and suggest strategies for better segmentation based on the provided criteria.

Context

The Instacart stakeholders are most interested in the variety of customers in their database along with their purchasing behaviors. They assume they can't target everyone using the same methods, and they’re considering a targeted marketing strategy. They want to target different customers with applicable marketing campaigns to see whether they have an effect on the sale of their products. Your analysis will inform what this strategy might look like to ensure Instacart targets the right customer profiles with the appropriate products. The stakeholders would like to be able to answer the following key questions:

Key Questions

- The sales team needs to know what the busiest days of the week and hours of the day are (i.e., the days and times with the most orders) in order to schedule ads at times when there are fewer orders.
- They also want to know whether there are particular times of the day when people spend the most money, as this might inform the type of products they advertise at these times.
- Instacart has a lot of products with different price tags. Marketing and sales want to use simpler price range groupings to help direct their efforts.
- Are there certain types of products that are more popular than others? The marketing and sales teams want to know which departments have the highest frequency of product orders.

Note: Instacart is a real company that’s made their data available online. However, the contents of this project brief have been fabricated for the purpose of this Achievement.

- The marketing and sales teams are particularly interested in the different types of customers in their system and how their ordering behaviors differ. For example:
    - What's the distribution among users in regards to their brand loyalty (i.e., how often do they return to Instacart)?
    - Are there differences in ordering habits based on a customer's loyalty status?
    - Are there differences in ordering habits based on a customer's region?
    - Is there a connection between age and family status in terms of ordering habits?
    - What different classifications does the demographic information suggest? Age? Income? Certain types of goods? Family status?
    - What differences can you find in ordering habits of different customer profiles? Consider the price of orders, the frequency of orders, the products customers are ordering, and anything else you can think of.

Stakeholders

- **Vice President of Marketing:** "We're always looking into improving our targeting for ad campaigns."

- **Senior Vice President of Sales:** "We need to know what part of our offering has the lowest market share and why. Based on this input, we could improve this sector and boost sales."

- **Instacart Customer:** "I want to receive ads, promotions, and recommendations that are relevant to the products I order regularly."

Note: Instacart is a real company that's made their data available online. However, the contents of this project brief have been fabricated for the purpose of this Achievement.

Data

Throughout this Achievement, you'll be using a number of open-source data sets from Instacart. You'll also receive a customer data set (created and included for the purpose of this project), on which you'll apply what you've learned to address the project's key questions. While each data set contains a different kind of information, they all include some kind of common identifier.

The project data you'll need is linked for reference below. However, you'll receive links to each data set in the Exercise content, as well.

Data Sets:

- Customers Data Set

Instacart Data Sets:

- Data Dictionary
- Citation (required in your final report): "The Instacart Online Grocery Shopping Dataset 2017", Accessed from www.instacart.com/datasets/grocery-shopping-2017 via Kaggle on <date>.

Note on Instacart "orders_dow" Variable

One of the variables in the data is "orders_dow", with "dow" meaning "days of the week". Each day corresponds to a number, as follows:

- 0 = Saturday
- 1 = Sunday
- 2 = Monday
- 3 = Tuesday
- 4 = Wednesday
- 5 = Thursday
- 6 = Friday

Note: Instacart is a real company that's made their data available online. However, the contents of this project brief have been fabricated for the purpose of this Achievement.