### **SUMMARY**

### **ABSTRACT**

This project is to analyze the daily and hourly instacart orders to find out the amount of highest daily and hourly orders and finally predict the number of future orders; which will help instacart to estimate the amount of shoppers they need. This will also help them to estimate profit and to find a balance between their in-store shoppers and full-service-shoppers for areas depending on their predicted daily demands.

#### **DESIGN**

- Data selection and cleaning
- Exploratory data analysis
- Time series model building (future work)

### **DATA**

"The Instacart Online Grocery Shopping Dataset 2017" iS used for this project. The data\_set used for this project is taken from evaluation\_set data and the last almost 98,570 rows are used.

# **Data description**

- order id: order identifier
- product id: product identifier
- add to cart order: order in which each product was added to cart
- reordered: 1 if this product has been ordered by this user in the past, 0 otherwise
- user id: customer identifier
- eval set: which evaluation set this order belongs in
- order number: the order sequence number for this user (1=first, n=nth)
- order dow: the day of the week on which the order was placed
- order hour of day: the hour of the day on which the order was placed
- days since prior order: days since the last order, capped at 30, NA if order number=1
- product name: name of the product
- aisle id: aisle identifier
- department id: department identifier
- aisle: the name of the aisle
- department: the name of the department

### **METHODOLOGY**

- Remove duplicates and n/a values
- Produce columns: Order day, order hour, days since last order: using Vlookup

Number of orders: using CountIF function

• Produce visuals to see relationship between order number(Target) with possible features order\_days, order\_hour,

# **TOOLS**

EXCEL - Data organization and analysis TABLEAU - Data visualization

# **COMMUNICATION**

Tableau dashboard

