Project Proposal

Jehan Joshi

5 July 2017

# *Instacart :- Market Basket Analysis (Kaggle Competition)*

## Company Description:

“Whether you shop from meticulously planned grocery lists or let whimsy guide your grazing, our unique food rituals define who we are. Instacart, a grocery ordering and delivery app, aims to make it easy to fill your refrigerator and pantry with your personal favorites and staples when you need them. After selecting products through the Instacart app, personal shoppers review your order and do the in-store shopping and delivery for you” – **Instacart**

## Problem Statement:.

Predict which products will a user buy again for the first time or add to the cart during a separate session.

**Link:**

[Kaggel](https://www.kaggle.com/c/instacart-market-basket-analysis)

## File Descriptions:

## Aisles.CSV :-

Contains a unique Aisle id and the description of the various aisles in a store. *Example :- [ 1, prepared soup salad ].* Where aisle id is one and prepared soup salad is the aisle description.

## Department.CSV :-

This file contains the department names and their unique id

## Order\_Products\_\_.csv :-

These files specify which products were purchased in each order. Order\_Products\_\_Prior.csv contains previous order contents for all customers. *'Reordered'* indicates that the customer has a previous order that contains the product.

## Note:-

That some orders will have no reordered items. You may predict an explicit 'None' value for orders with no reordered items.

## Orders.csv :-

This file specifies to which set (prior, train, test) an order belongs. Prediction is for **reordered** items only for the test set orders. 'order\_dow' is the day of week.

## Products.csv :-

This file contains the *product\_id*,*product\_name*,*aisle\_id*,*department\_id*

# Methods That Can Be Used

* Association :- In order to know which items can be paired togethere for more effective selling
* Frequently Reordered items :- Analyze frequently reorderd sets of items and their count
* Seasonality And Customer Trends :- Identify seasonality and customer trends for a longer period of time
* Algorithms That Can Be Used :- The **Apriori Algorithm** is a popular Association Algorithm which is used to find frequently ordered itemsets and generate rules regarding the same.