

Analysis of customer specific and store specific attributes to improve sales and increase loyal customers

Presented by

Vaishnavi Kommaraju

(Data Scientist)



Data Structure

(269720 rows, 75 columns)

Geographical variables

(city, province, country, store address, etc)

Demographic variables

(income, occupation, no. of children, occupation, education)

Financial variables

(Store cost, sales, unit sales, etc.)

Store specific variables

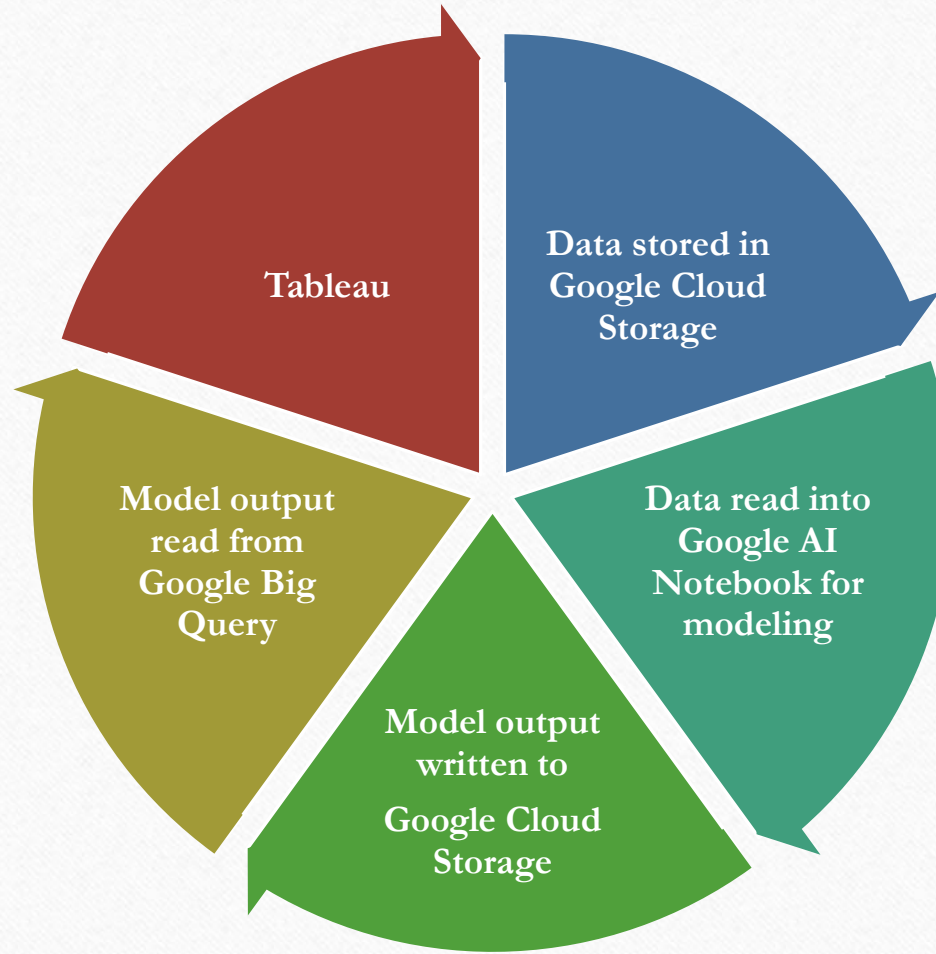
(Frozen area, Grocery sq. ft., Meat Sq. ft., etc)

Product specific attributes

(category, subcategory, recyclable packaging, etc)

Promotions

Data Workflow



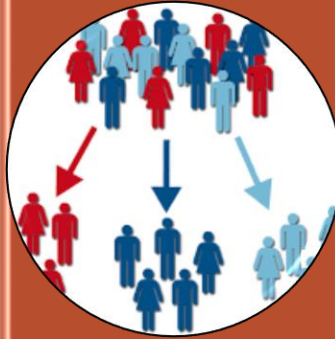
Business Questions



How does my business look like?



What items do customers buy together?



Are there any groups of customers with similar attributes?



Who are my Platinum and Gold customers?



How do I increase my Platinum and Gold customers?



How do I improve my store profit margins?

How can I apply Machine Learning to solve these problems?


```
def recommendations_using_Apriori(item):
    recommend = []
    for i in range(0,2646):
        if item == apriori_rules.iloc[i,0]:
            recommend.append(apriori_rules.iloc[i,1])

    return recommend
```

10 recommendations for Big Time Frozen Cheese Pizza

```
product_name = {'Big Time Frozen Cheese Pizza'}
recommendations=recommendations_using_Apriori(product_name)
print(*recommendations[0:10], sep = "\n")

frozenset({'Big Time Pancake Mix'})
frozenset({'Carrington Ice Cream'})
frozenset({'Fabulous Orange Juice'})
frozenset({'Framton City Map'})
frozenset({'Hermanos Squash'})
frozenset({'Hilltop Extra Moisture Shampoo'})
frozenset({'Hilltop Silky Smooth Hair Conditioner'})
```

15 recommendations for Special Wheat Puffs.

```
product_name = {'Special Wheat Puffs'}
recommendations=recommendations_using_Fpgrowth(product_name)
print(*recommendations[0:15], sep = "\n")

frozenset({'Better Chicken Noodle Soup', 'Moms Potato Salad'})
frozenset({'Sunset Large Sponge'})
frozenset({'Fabulous Apple Juice'})
frozenset({'Hilltop Silky Smooth Hair Conditioner'})
frozenset({'Hermanos Potatos'})
frozenset({'Tell Tale Lettuce'})
frozenset({'Steady 200 MG Acetaminifen', 'Hilltop Silky Smooth Hair Conditioner'})
frozenset({'Better Chicken Noodle Soup'})
frozenset({'Better Chicken Noodle Soup', 'ADJ Rosy Sunglasses'})
frozenset({'ADJ Rosy Sunglasses', 'Moms Potato Salad'})
frozenset({'Better Chicken Noodle Soup', 'ADJ Rosy Sunglasses', 'Moms Potato Salad'})
frozenset({'Nationeel Grape Fruit Roll'})
frozenset({'Hermanos Potatos', 'Nationeel Grape Fruit Roll'})
frozenset({'Carrington Beef TV Dinner', 'Hilltop Mint Mouthwash'})
frozenset({'Fort West BBQ Potato Chips'})
```

Market Basket Analysis

RFM Analysis



How recently did
the customer visit
the store?



How frequently
does the customer
buy from the store?



Rank the customers
based on the
monetary value of
their transactions.

Calculate the composite RFM score by combining the above features.

K – Means Clustering



Demographic Variables



Product Specific Preferences



Store Specific Preferences



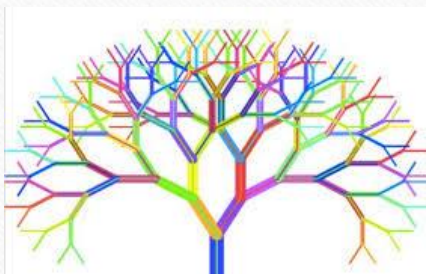
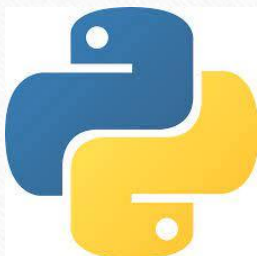
Time of purchase



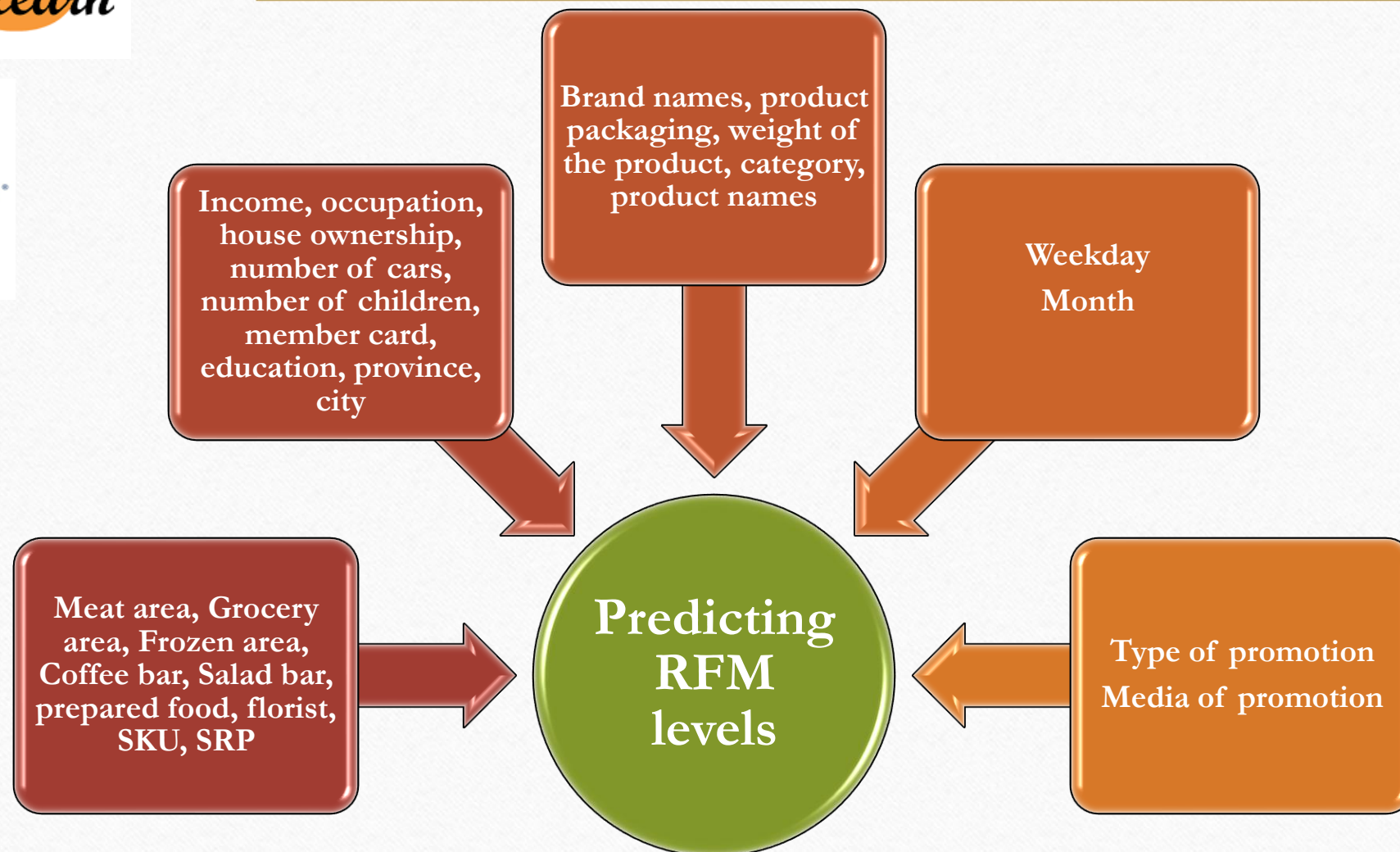
RFM Loyalty Scores

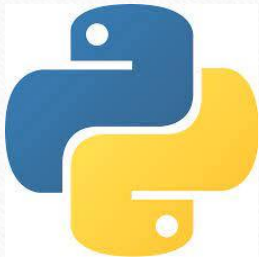
dmlc
XGBoost + 

 + a b l e a u



Model Accuracy: 78%





Model Accuracy: 69%

