TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

HIMALAYA COLLEGE OF ENGINEERING

A FINAL PROJECT REPORT

ON

"Identifying Product Bundles from Sales Data using Market Basket"

[CT 755]

SUBMITTED BY: -

Ashish Neupane (89856)

Bibek Dhakal (89858)

Bijay Aryal (89860)

Nabin Kandel (89870)

SUBMITTED TO: - DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

March, 2021

[Identifying Product Bundles from Sales Data using Market Basket]

[CT 755]

A FINAL PROJECT REPORT

"A report submitted in partial fulfillment of the requirements for bachelor's degree in computer engineering"

SUPERVISOR

Ramesh Tamang

SUBMITTED BY: -

Ashish Neupane (89856)

Bibek Dhakal (89858)

Bijay Aryal (89860)

Nabin Kandel (89870)

SUBMITTED TO: -

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

CHYASAL, LALITPUR

March, 2021

ACKNOWLEDGEMENT

The success and final outcome of this project requires a lot of guidance and assistance from many people and we are extremely privileged to have got this along the starting of the project. This project is possible only due to such supervision and assistance of Teaching staffs of Electronics and Computer engineering department and we would not forget to thank them. We respect and thank our department for providing us such opportunity and helping us directly or indirectly to do this project. We are thankful to our project coordinator **Er. Narayan Adhikari Chettri** and supervisor **Er. Ramesh Tamang** for providing us constant encouragement, support and guidance.

We would like to express my deepest appreciation to all those who saw the possibility in us for doing this project. A special thanks to them who contributed in stimulating suggestions and encouragement and helped us to coordinate our project especially in writing this report.

We are thankful to our college administration for equipping us with all the resources and providing us with the pleasant environment to work on. Also, we would like to thank our seniors and colleagues for their valuable comments and suggestion for this project.

ABSTRACT

Bundling has become a prevalent promotion strategy rapidly since it is capable of raising values to buyers and generating profit to sellers, which perfectly matches the objective of a transaction process. From the consumer's perspective, they will be able to save up to certain percent on average through purchasing a bundle package with a discounted price, which is a key driver of bundling. Also, consumers can save search cost, which will increase their willingness to purchase since they can easily find all wanted products and services in a bundle package provided by the seller. Some people also prefer bundles because they can reduce compatibility risk among components. From the seller's perspective, adopting bundling can help increase the number of buyers and thus increase sales. Moreover, a newly released product will be noticed and accepted by consumers if it is bundled with an existing product. The seller's cost, like packaging cost and distribution cost, can also be saved by offering several products as a bundle.

Here in this project in titled "Identifying Product Bundles from Sales Data using Market Basket" is a model which is determined by its performance, that is helping to identify product bundle and determine the market basket. As we know there are so many prediction problems during sales of products. With reference to old sales data, the system predicts product bundles. Having a sample of data from Instacart, our project is capable of going through some clustering and analyzing process. After analyzing of data, the system will predict the product bundle which will help the selling company to increase its sales by determining the product bundle for E-commerce site by using k-means as clustering process. This project is specially designed for the E-commerce websites who want to solve problem of product bundling and market basket analysis. This provides the platform to use some technique to solve problems like product bundling. It helps to generate sufficient theory and practice material for research-based study. This project helps us to predict future in more reliable way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

MSPL	Most Sold Product Listing
PCA	Principal Component Analysis
WSSSE	Within Set Sum of Squared Error

1. INTRODUCTION

What's your daily routine that keeps you busy the whole day? Go to the gym? Shop your groceries? Instacart is a same-day grocery delivery service that can save yourself that trip to the market. It will connect you with personal shoppers in your area to shop and deliver groceries from your favorites stores in as little as an hour. Using the Instacart Public Datasets, we would like to explore solving the following three problems: recommend new products to customers, recommend products to be bought together, and predict which products will be in a customer's next order.

The project is of great business value. Instacart, as a grocery delivery startup, has several main competitors like Amazon Fresh and Shipt. In order to acquire more customers and increase customer retention, Instacart needs to provide more delightful shopping experience by making it easier to fill customer's refrigerator and pantry with personal favorites. With the right recommendation and prediction, Instacart can boosts sales and profits, and attracts more customers by helping customers save shopping time.

1.1 Background

Price optimization is the concept of offering goods at different prices which varies according to customer's demand. Fierce competition always makes business adopt various promotion strategies to attract more consumers and outperform other competitors. To meet consumer's needs and expectation is the basic principle to survive in this competitive business environment. A transaction is a two-sided process from which both buyers and sellers want to get the most benefit. Thus, it is crucial to keep balance between them. With consumer's desire for buying related products at the same time, seller have to provide combinations of products in order to facilitate the purchasing process.

Bundling is a promotion strategy in which sellers provide multiple products or events as a single package with an attractive price. Bundling has become a prevalent promotion strategy rapidly since it is capable of raising value of buyers and generating profit to sellers, which perfectly matches the objective of a transaction process. Also, consumers can save search cost, which will increase their willingness to purchase since they can easily find all wanted products and services in a bundle package provided by the seller. From the seller perspective, seller's cost, like packaging cost and distribution cost, can also be saved by offering several products as a bundle. A bundle of durable item with some nondurable ones is the most common format of bundling, like bundling a printer and ink or selling a computer with some accessories. Without bundling, under which the reputation mechanism is inoperative, and sellers may have chances to reduce the quality of durable and infrequently purchased goods. But for frequently purchased items which has known quality, firms need to keep their quality at a high level in order to hold the advantages over their competitors.

1.2 Problem Statement

There is different strategy for sales data out of which product bundling can be considered as better strategy. There are different traditional methods in which consumers can only purchase products or services separately with their original prices. It allows buyers to see the sales process clearly and pick up exactly the products they want.

On the contrary, in product bundling strategy, we recommend several products to the customer for discovering items that are most likely to be purchased together(bundle) by mining historical data of transactions. Bundling strategy is a more flexible one to offer both individual products and the whole bundle, and a buyer can make a choice between purchasing the entire bundle or one part of the bundle package and help system to predict which product will be in user's next order.

1.3 Objectives

The general objective of our project is to implement sales using product bundling strategy.

The specific objectives are:

- Recommend new products to customers
- To discover items that are most likely to be purchased together(bundle) by mining historical data of transactions
- To predict which product will be in user's next order

1.4 Scope of the Project and Application

The scope of this project is to provide an efficient, easy mechanism for selling of product in bundle form. Individual can take recommendation of different bundles of product. The project will help for increasing of sales and also fulfilment of product in efficient way.

The project has following application:

- It helps customers to get recommendation with their favorite products
- It helps recommending the bundles of products to customers
- It helps to predict which product the customer will buy
- It assists department to clear dead stock and increase sales

2. LITERATURE REVIEW

Companies offer product bundles with special discounts in order to sell more products. The main purpose is to propose a novel model for product bundling in e-commerce websites by using market segmentation variables and customer loyalty analysis. RFM model is employed to calculate customer loyalty. The product bundles are determined for each market segment via clustering algorithms. Apriori algorithm is also used to determine the association rules for each product bundle. Classification models are applied in order to determine which product bundle should be recommended to each customer. The results demonstrate that the silhouette coefficient, support, confidence, and accuracy values are higher when both customer loyalty level and market segmentation variables are used in product bundling. Accordingly, the proposed model increases the chance of success in direct marketing and recommending product bundles to customers [1].

Bundling, the strategy of marketing products in particular combinations, is growing in significance given the boom in high technology and e-commerce. Two main objectives: First, sought to draw a set of key guidelines for bundling and pricing from a large body of 'traditional' literature rooted in stylized economic models. By considering factors such as the nature of heterogeneity in consumers' reservation prices, the extent of the underlying correlation in reservation prices, the degree of complementarity or substitutability, and the nature of competition. The key conclusion is that no one form of bundling is always the best. Second, an attempt to showcase the extant methodologies for bundle design and pricing. The studies consider here have an empirical character and pertain to issues of a 'marketing' nature [2].

Product bundling is a marketing strategy that has been widely studied in research literature and extensively used in practice. It is necessary to develop algorithmic approaches to determine which items should be in a profitable bundle, which bundling strategy is most profitable, and what the proper price is for a bundle.

Consumer' behaviors may be not in accordance with their statements in a survey, thus the transaction data is a more reliable source to predict their purchase behaviors. As consumers' demand and market supply will fluctuate continuously, fail to consider price elasticity of demand (PED) will cause biases for prediction, where PED is used to measure consumers' abilities and willingness to pay for certain products. In this thesis, they propose a data mining framework which incorporates the time value of money in data mining tasks, and it is capable of determining the product combination and price of a bundle in order to maximize the revenue. Applying association mining to generate meaningful candidate bundles and reduce computation cost. This framework analyzes consumer and product data, taking demand and inflation factors into consideration, to fill in the gaps as mentioned [3].

Bundling is a very popular sales-promotion tool, in which a critical issue is to decide what products should be sold together in order to improve sales. Traditionally, this decision is based on the order data collected from the points of sale. However, Internet marketing now allows marketers to efficiently collect not only order data but also browsing and shopping-cart data, which provide marketers with information on the consumers' decision-making processes, rather than only the final shopping decisions. The present study aimed to determine the value of this newly available information by comparing the performance of decision-making on product bundling based on three types of data on online shopping behaviors. The results from a field experiment reveal that significantly better decisions are made on the bundling of products when browsing and shopping-cart data are integrated than when only order data or browsing data are used [4].

Bundling is an efficient method to achieve business objectives in many industries. However, decisions of bundle selection and pricing are complicated when multiple products are involved. In this paper, they investigate a bundle-pricing decision model for multiple products. With the objective to maximize the retailer's profit, an integrated bundle-pricing model for multiple commodities is formulated as a Non-Linear Mixed Integer Program based on the framework of Stackelberg game. By adding auxiliary decision variables, this model is converted into a Mixed Integer Linear Program and solved by Cplex. Numerical experiments and sensitive analysis are conducted to provide managerial insights for bundling multiple products. It indicates that low consumption level consumers prefer bundles composed of more commodities with lower prices. The products with higher cost level should be bundled with smaller bundle size and higher prices [5].

Over half a century old and showing no signs of aging, *k*-means remains one of the most popular data processing algorithms. As is well-known, a proper initialization of *k*-means is crucial for obtaining a good final solution. The recently proposed *k*-means++ initialization algorithm achieves this, obtaining an initial set of centers that is provably close to the optimum solution. A major downside of the *k*-means++ is its inherent sequential nature, which limits its applicability to massive data: one must make *k* passes over the data to find a good initial set of centers. In this work we show how to drastically reduce the number of passes needed to obtain, in parallel, a good initialization. This is unlike prevailing efforts on parallelizing *k*-means that have mostly focused on the post-initialization phases of *k*-means. We prove that our proposed initialization algorithm *k*-means|| obtains a nearly optimal solution after a logarithmic number of passes, and then show that in practice a constant number of passes suffices. Experimental evaluation on real-world large-scale data demonstrates that *k*-means|| outperforms *k*-means++ in both sequential and parallel settings [6].

Offering online personalized recommendation services helps improve customer satisfaction. Conventionally, a recommendation system is considered as a success if clients purchase the recommended products. However, the act of purchasing itself does not guarantee satisfaction and a truly successful recommendation system should be one that maximizes the customer's after-use gratification. By employing an innovative associative classification method, they are able to predict a customer's ultimate pleasure. Based on customer's characteristics, a product will be recommended to the potential buyer if our model predicts his/her satisfaction level will be high. The feasibility of the proposed recommendation system is validated through laptop Inspiron 1525 [7].

Gradient Boosting Decision Tree (GBDT) is a popular machine learning algorithm, and has quite a few effective implementations such as XGBoost and pGBRT. Although many engineering optimizations have been adopted in these implementations, the efficiency and scalability are still unsatisfactory when the feature dimension is high and data size is large. A major reason is that for each feature, they need to scan all the data instances to estimate the information gain of all possible split points, which is very time consuming. To tackle this problem, they proposed two novel techniques: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). With GOSS, we exclude a significant proportion of data instances with small gradients, and only use the rest to estimate the information gain. They prove that, since the data instances with larger gradients play a more important role in the computation of information gain, GOSS can obtain quite accurate estimation of the information gain with a much smaller data size. With EFB, they bundle mutually exclusive features (i.e., they rarely take nonzero values simultaneously), to reduce the number of features. They prove that finding the optimal bundling of exclusive features is NP-hard, but a greedy algorithm can achieve quite good approximation ratio (and thus can effectively reduce the number of features without hurting the accuracy of split point determination by much). They call our new GBDT implementation with GOSS and EFB LightGBM. Our experiments on multiple public datasets show that, LightGBM speeds up the training process of conventional GBDT by up to over 20 times while achieving almost the same accuracy [8].

Market Basket Analysis is a technology to observe patterns of purchase and this technology depends on the extent to which products relate to each other and the quantities sold. Mechanism of Action: they used the Rapid Miner program to analyze the data and after the analysis results were found showing the relationship between the products and what products are associated with each other are explained in the report. These results help the employer arrange the rows of products and connect them to each other in specific packages [9].

3. REQUIREMENT ANALYSIS

3.1 Software Requirement

3.1.1 Python Libraries

- Pandas: - Pandas is a python library that is used to read different csv files and merge csv files. It is used for different preprocessing which includes cleaning the dataset and transferring the datasets used. Finally working data set is generated. Pandas is used in the project for exploratory data analysis.

- Numpy: - Numpy is used for working with arrays and multidimensional matrix in the project. The project includes dimensionality reduction and generating vector space and for this numpy is used.

- PySpark: - PySpark is used for data preprocessing in complex data structure for large datasets. It is achieved using Py4j library as it provides the bridge between python and java. PySpark speed up computational time for the project by creating different Java Virtual Machine.

3.1.2 Django Framework

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so it can focus on writing your app without needing to reinvent the wheel. It’s free and open source. Django helps to take applications from concept to completion as quickly as possible. Django takes security seriously and helps to avoid many common security mistakes.

3.1.2 JavaScript

JavaScript is used as client scripting language in the project that provide mainly local storage feature for cart system and for handling different AJAX request for validating input at frontend with Django as backend.

3.2 Functional Requirements

Functional requirements may involve calculations, technical details, data manipulation and processing, and other specific functionality that define what a system is supposed to accomplish. Various functional requirements of our system are as follows:

- Create a new and valid account
- Validate the login credentials
- Display the list of products
- Add products to cart
- Confirmation of checkout by sending email

3.3 Nonfunctional Requirements

The non-functional requirements like performance, information, economy, control and security efficiency and services are very essential for successful project completion. The non-functional requirements of the project are as follows:

- The system should be easy to use, user friendly in operation
- The system should perform with efficient throughput and response time
- The system should provide good accuracy

4. System Design

4.1 Use Case Diagram

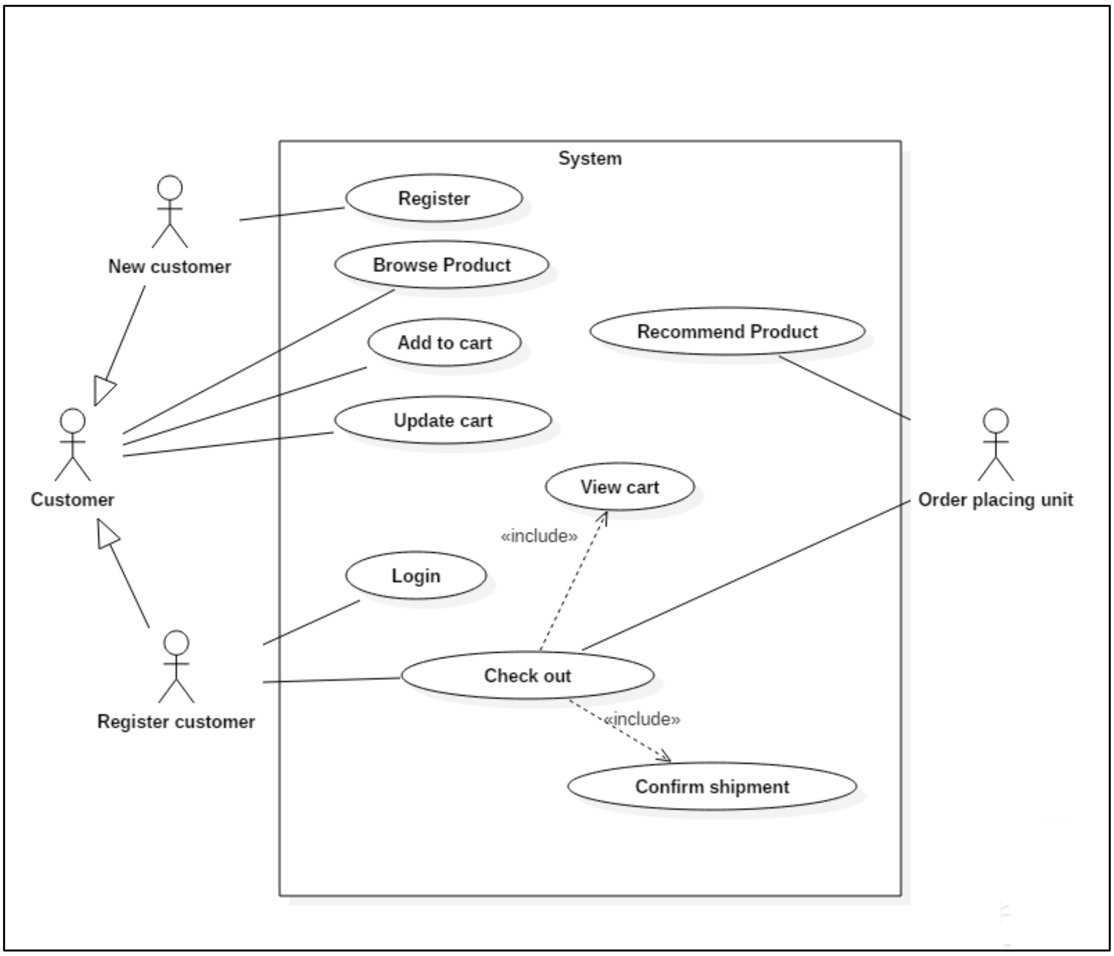


Figure 4. 1: Use Case Diagram.

In Use Case Diagram, new customer can browse product, register, add product to cart and update cart. Customer has additional features to login and checkout. While doing checkout, user must be logged in which includes viewing the cart (items, quantity and price of products) and shipping confirmation includes shipping details (customer detail, location, email). While adding to cart, new products that are likely to be purchased will be given by order placing unit.

4.2 ER Diagram

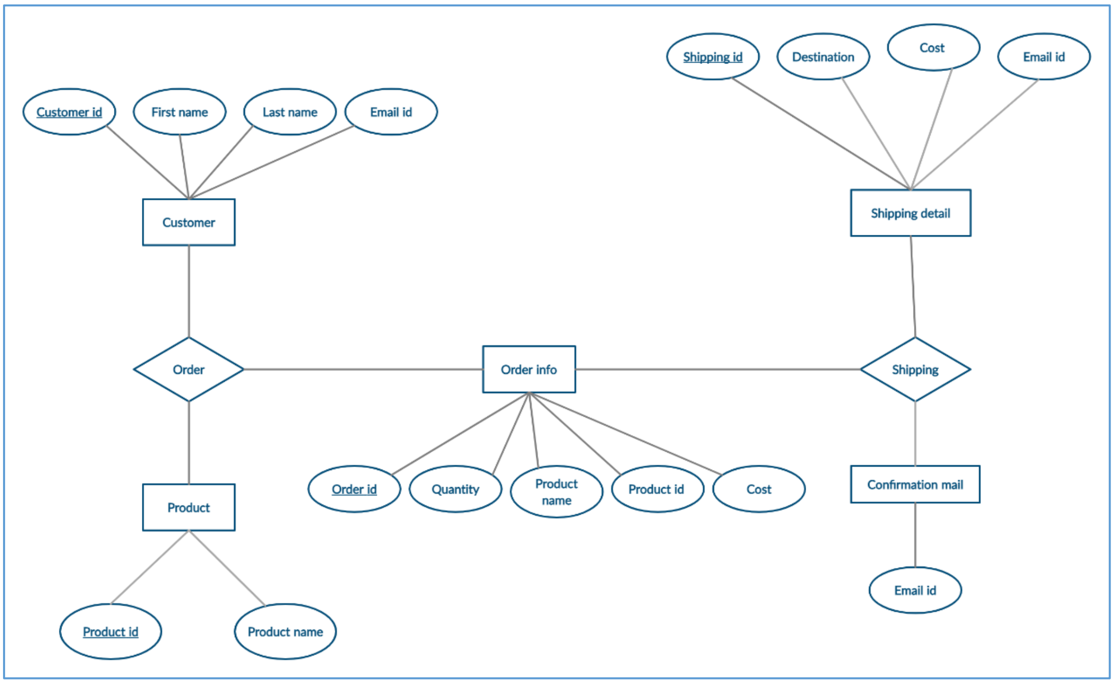


Figure 4. 2: ER Diagram.

Entity-Relationship (ER) diagram in this project include different entity, attributes and

relationship set. In this diagram we have entities as customer, product, order info, shipping details and confirmation. The entity customer has attributes first name, last name and Email Id. Entity product has attributes product name and product id. Two entities customer and product have relationship as customer can order the product. Entity order info has attributes quantity, product id, product name and cost. The relationship between the entities where different order may be appearing on order info. Similarly, entities shipping details and confirmation mail has attributes destination, cost, and email id respectively. Entities shipping details and confirmation has the relationship of shipping where multiple shipping details can be confirming through email.

4.3 Class Diagram

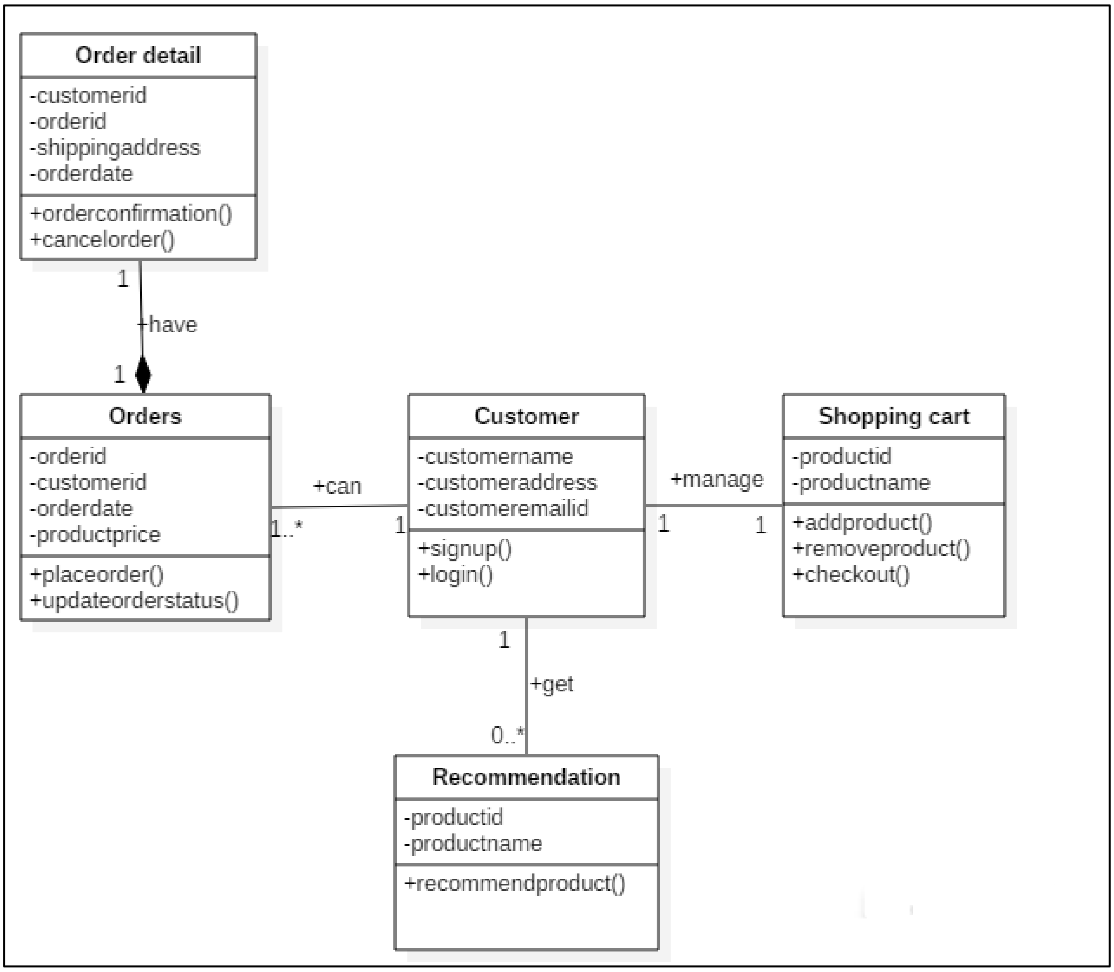


Figure 4. 3: Class Diagram.

This project class diagram has five different classes in which each of the classes are divided into three section. Top section includes name of the class. The second one is used to show the attributes and final one is used to describe the operations performed by the classes. The class name customer has the attributes customer name, customer address and customer email id in which - means that the attributes customer name, customer address and customer email id should be private. The operation performed on a customer class are sign up and login in which + means that the operations are public. Recommendation class have attributes product id and product name set as a private attribute which also perform the operation recommend_ product set as a public operation. The customer class and recommendation class have a relation in which the single customer can get zero to many recommendations of the product. The class order consists of attributes order id, customer id, order date and product price. All of those attributes are set as private. The operation performed are update order status and place order in which both are set as public operation. The class order details have a relation with the class order which consists of attributes id, order id, shipping address and order date with the operations cancel order and order confirmation. Those classes have the relationship as a single order can provide the detail of single order. The final class of our class diagram is shopping cart having attributes product id and product name include operations as add product to cart, remove product from cart and checkout where attributes are set as private and operations as a public respectively. The class customer has multiplicity relationship with the class shopping cart in which a single customer can handle a single shopping cart. The class orders and order detail have the relationship of composition that order detail can only be exist if the class orders are present.

5. METHODOLOGY

5.1 Workflow Diagram

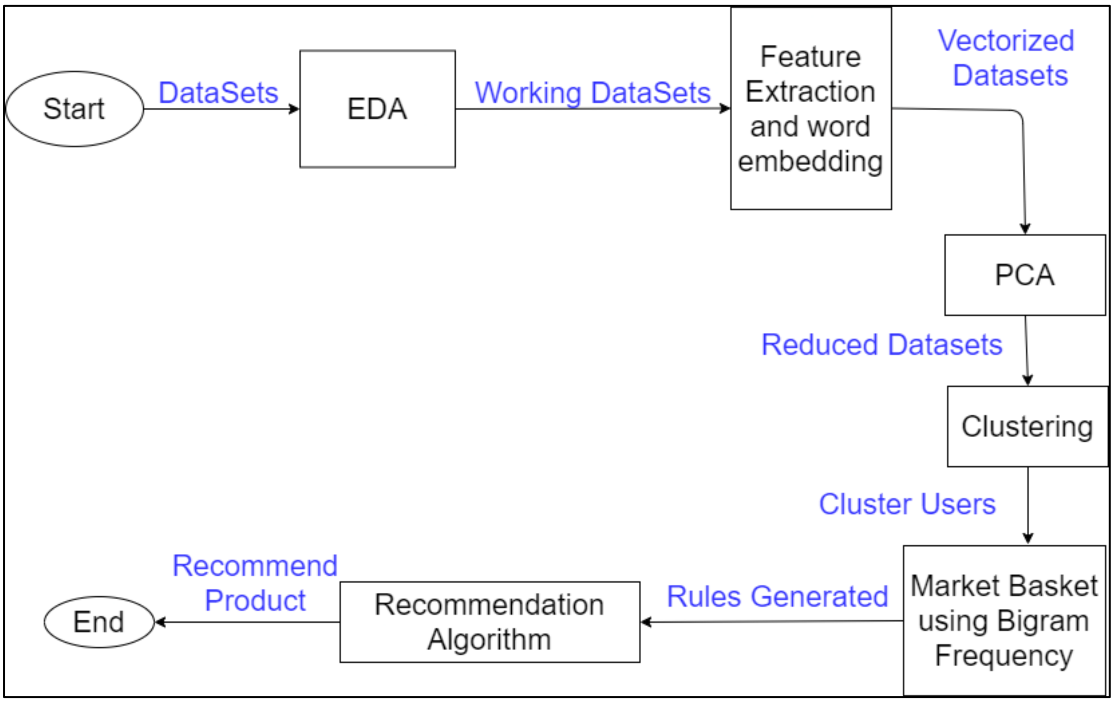


Figure 5. 1: Workflow Diagram.

5.1.1 Exploratory Data Analysis

5.1.1.1 Dataset Description:

The dataset contains sample of over 3 million grocery orders from more than 200,000 instacart users.

There are basically 5 datasets provided:

a) Orders.csv: this file contains all information about given order like: user who has made the order, when was the product purchased and days since prior order.

b) Order_products.csv: this file has detailed information about the product that has been brought in the given order along with reordered status

c) Products.csv: this file contains the name of products with their corresponding product_id

d) Aisles.csv: this file contains the different aisles name (passage in department store) with aisle_id

e) Departments.csv: this file contains the different department name (which contains aisle and products) along with the department_id

5.1.1.2 Exploration:

Exploring those datasets, there are exactly 3,421,082 orders made by 206,209 instacart users. Here a single customer has made from 4 to 100 orders where 4 is minimum no. of orders and 100 is maximum no. of order a customer has made. If looking at the ordering habits of customer in each day, orders are high during Saturday and Sunday and 10am to 11 am has highest order done in a day followed by 12pm to 4pm and from 5pm no. of order seems to be decreasing. Combining orders made in each day and each hour, it seems that, Saturday evening and Sunday mornings are the prime time for orders.

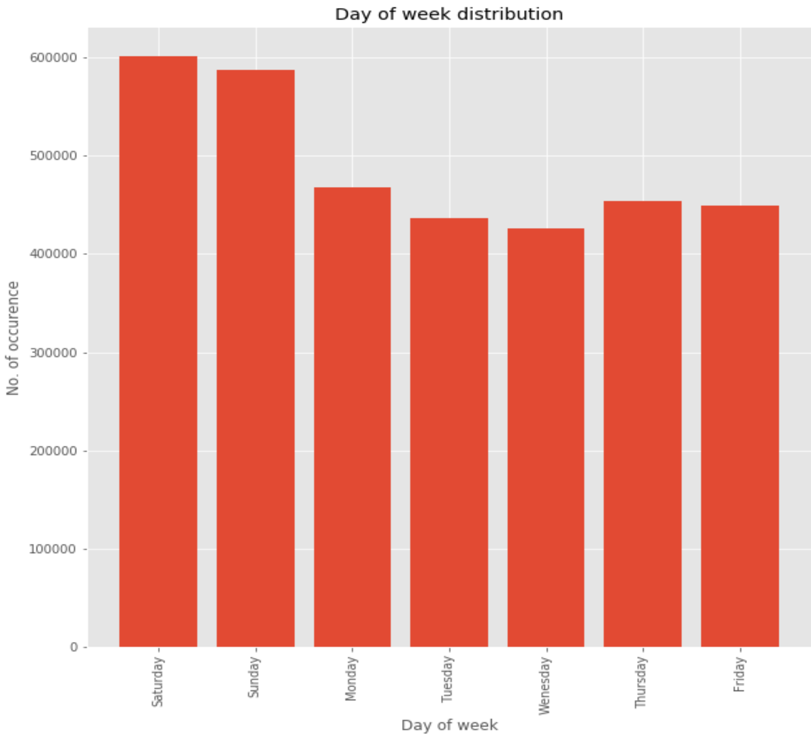


Figure 5.1.1.2. 1: Weekly Distribution

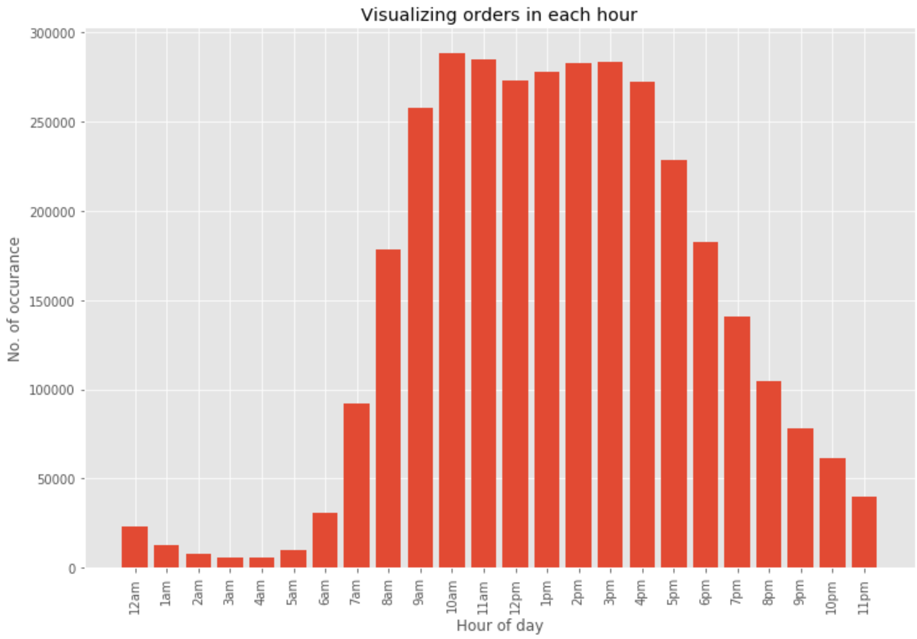


Figure 5.1.1.2. 2: Order Visualization

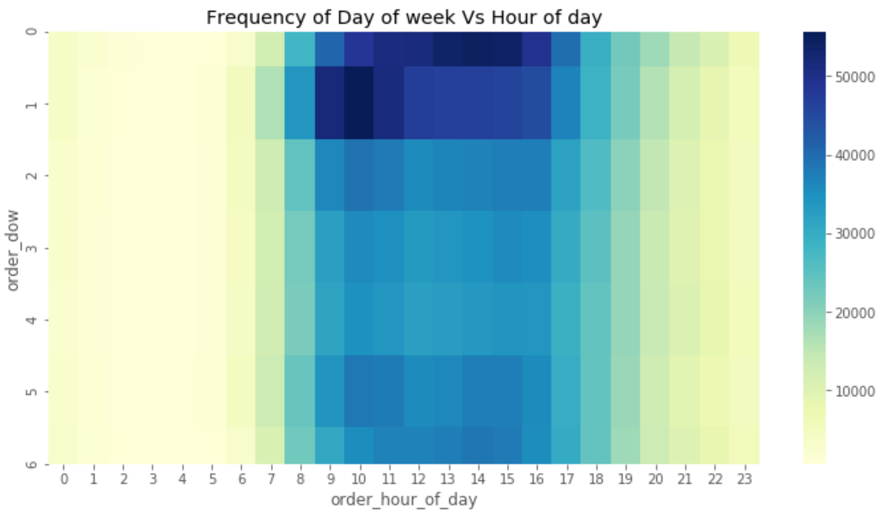


Figure 5.1.1.2. 3: Frequency Comparison

Looking at the best sold product or top product that customers has made clearly mentioned that banana is best sold product. Around 18,726 bananas have been ordered by user. Looking at the top 10 best sold products:

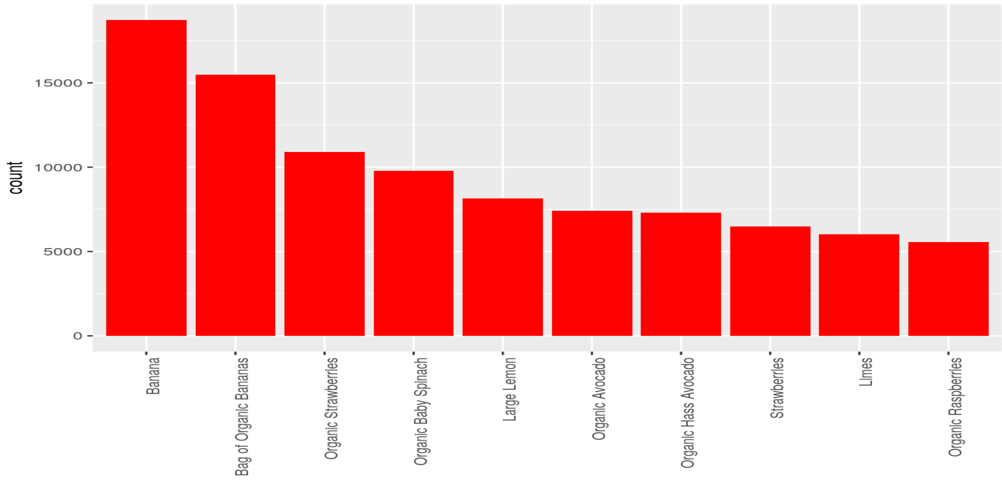


Figure 5.1.1.2. 4: Top 10 Selling Product

From the result, it is also clear that most of the products are sold from produce department, followed by dairy eggs and snacks departments. There were 21 departments containing 134 aisles.

Finally, the last stage of EDA involves merging and combining only those columns that will be needed for the further work. For a specific transaction i, if an item j is purchased then the matrix position (i,j) is made as 1. If the item j is not purchased in the transaction i then the matrix position (i,j) will be made as 0. Some dummy transactions will be there with no items, it should be rejected. The goal is to find the frequent items which occur together and so transactions with one or two items is rejected for effectiveness. Transactions with many numbers of items will provide useful information about customers' behavior. So, all these datasets were merged based on order_id, product_id, department_id and user_id and incomplete records were removed so that final working dataset is ready and cleaned.

5.1.2 Feature extraction

Once the final working dataset is ready some of the features are directly extracted while other are done using word2vec analysis by which working dataset is reduced to more manageable groups for processing. Some of the features that are directly extracted includes:

- Mean of order placed day of week
- Mean of order placed hour of day
- Mean of days since last order
- Total number of orders made
- Total number of products brought

5.1.2.1 Word2vec Analysis: Save customer based on their buying habits

Word2vec is one of the most popular technique to learn word embedding using shallow neural network. Word embedding are text converted into numbers and there may be different numerical representations of same text that means word embedding is the vector representation of a particular word so that this system do not need to process plain text or string during clustering.

For word2vec analysis, we combine all the product name into one row per user and the customer buying habit includes:

- Which day of week?
- Which hour?
- Days since last order
- Number of total orders

So, input data for word2vec is each row per user considering each row is a bag of words as:

user_id	product_name
1	Soda Original Beef Jerky Pistachios Organic St...
2	Artichoke Spinach Dip Chipotle Beef & Pork Rea...
5	Uncured Genoa Salami Plain Whole Milk Yogurt W...
7	85% Lean Ground Beef Organic Apple Slices Appl...
8	Organic Baby Spinach Michigan Organic Kale Bag...

Figure 5.1.2. 1: Input data for word2vec analysis

Such input is trained using word2vec model. The model maps each word to a unique fixed-size vector. The word2vec model transforms each row into vector using average of all words available in that word. This vector can then be later on used as input for clustering. The main advantage of doing this is that similar words are close to vector space which makes generalization to novel patterns easier, dimensionality reduction and model estimation more robust.

In our implementation of word2vec, we used skip-gram model. The training objective of this model is to learn word vector representation that are good at predicting its context in the same sentence. Mathematically, given a sequence of training words w1,w2,w3….wr, the objective of skip-gram model is to maximize average of log-likelihood.

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j=-k}^{j=k} log\, p\left(wt + j \middle| wt\right) \ldots\ldots\ldots\ldots\ldots\ldots \;(\, i)$$

Where k =size of training window

In skip gram model, every word w is associated with two vectors uw and vw which are vector representation of w as word and context respectively. The probability of correctly predicting word wi give word wj is determined by the softmax model, which is

$$\frac{\exp(u_{wi}^{T} * V_{wj})}{\sum_{l=0}^{v} \exp(u_{l}^{T} * V_{wi})} \quad \text{.....................................(ii)}$$

Where V is vocabulary size.

After learning a mapping from each row to vectors and saving the result as:

```
+--------------------+--------------------+
|        product_name|              result|
+--------------------+--------------------+
|[Hass, Avocados, ...|[0.35486072038398...|
|[Dulce, de, Leche...|[-0.4102908031394...|
+--------------------+--------------------+
only showing top 2 rows
```

Figure 5.1.2. 2: Output from word2vec Analysis

In order to use it or apply as input for k-mean clustering we will save all the vectors value by reshaping all the values into 5 vectorized features as:

	vectorized_feature_1	vectorized_feature_2	vectorized_feature_3	vectorized_feature_4	vectorized_feature_5	user_id
0	0.354861	0.275951	-0.142631	0.043631	-0.104238	160334
1	-0.410291	0.491902	-0.108292	-0.270677	-0.421957	88565
2	0.023476	0.175154	-0.221996	-0.103921	-0.421835	79012
3	0.196822	0.120681	-0.193381	-0.111424	-0.134239	169304
4	0.345206	0.037274	-0.129840	-0.088376	-0.154438	55886

Figure 5.1.2. 3: Vectorized Dataset

5.1.2.2 PCA: Dimensionality Reduction

The obtained result after word2Vec analysis is further reduced to 2-Dimension using PCA (principal component analysis) while preserving as much information as possible.

Step1: Standardization

In the first step we standardize or normalize the range of continuous initial variables so that each one of them contributes equally to the analysis.

Mathematically, this can be done by subtracting each value with the mean and dividing by standard deviation for each value of each variable.

$$Z = \frac{value - mean}{standard} deviation \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (iii)$$

Once the standardization is done, all the variables will be transformed to the same scale which will solve the problem of dominancy that means if there are large difference between the ranges of initial variables, those variables with larger ranges will dominate over the small range which is not acceptable.

Step2: Calculation of covariance matrix

The aim of this step is to understand how the variables of input dataset are varying from the mean with respect to each other or in other words, to see if there is any relationship between them.

Since the dataset we took is 2-Dimension, this will result in a 2*2 covariance matrix.

$$Matrix(convariance) = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (iv)$$

It is actually the sign of covariance matrix that matters:

- If positive then: two variables increase or decrease together (correlated)
- If negative then: one increases when the other decreases (inversely correlated)

Step3: Computing the Eigen vectors and corresponding Eigen values of covariance matrix to identify the principal components.

Eigen vectors and Eigen values are linear algebra concepts that we need to compute from the covariance matrix in order to determine principal components of data. Principal components are new variables that are constructed as linear combinations or mixtures of initial variables. These combinations are done in such a way that new variables or principal components are uncorrelated and most of the information within initial variables is compressed into first components that means PCA tries to put maximum possible information in first component then maximum remaining information in the second and so on.

For 2-Dimensional dataset there are 2 variables, therefore there are 2 Eigen vectors with 2 corresponding Eigen values. Eigen vector of covariance matrix are actually the directions of the axes where there is the most variance (most information) which we call principal components and Eigen values are simply the coefficients attached to Eigen vectors.

γ is an Eigen value for a matrix A if it is a solution of the characteristic equation:

$$\left|\gamma * I - A\right| = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots \textit{(v)}$$

where I is an identity matrix of same dimension as A. For each Eigen value γ, a corresponding Eigen vector v can be found by solving:

$$\left(\gamma * I - A\right) * v = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \textit{(vi)}$$

Step4: Choosing components and forming a feature vector.

We order Eigen value from largest to smallest so that it gives us components in order of significance. To reduce the dimensions, we choose first p eigenvalues and ignore the rest. We lost some information in the process but if Eigen values are small, we do not loss much information.

Next, feature vector is formed which is Eigen vectors. For 2-Dimension case:

$$Featurevector = (eigenvector1, eigenvector2) \ldots\ldots\ldots\ldots\ldots\ldots \text{(vii)}$$

Where, $eigenvector1$ has more significance or carry more information than $eigenvector2$.

Step5: Forming principal components.

In order to form principal components or new variable, we take transpose of feature vector and multiply it with transpose of original dataset.

$$NewData = Featurevector^T * StandardizedOriginalData^T \ldots\ldots\ldots\ldots\ldots \text{(viii)}$$

Where $NewData$ is matrix consisting of principal components and $StandardizedOriginalData$ is scaled version of original dataset.

From above steps, we reduced large dataset into 2-Dimension as:

```
pca_df.head()
```

	0	1
0	30.851828	-5.844095
1	-10.814946	1.914741
2	-111.061957	4.096510
3	-77.806557	0.083442
4	-22.860360	2.186698

Figure 5.1.2.2. 1: Reduced Datasets after PCA

5.1.3 Clustering: cluster user based on their buying habits and preference

After dimensionality reduction, next step is to cluster user which is done based on their buying habits and preference.

The customer buying habit includes:

- Which day of week?
- Which hour?
- Days since last order
- Number of total orders

The preference includes:

- Number of total products
- Preferred products from word2vec analysis

5.1.3.1 k-mean clustering

k-means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem.

Algorithm:

BEGIN

Find the optimal number of cluster K using the elbow method and get the cluster centers m = {m1, m2,….. k}.

Repeat

For each object $p \in D$

set p to the cluster $Gi \leftarrow$ arg $min\ i\ |p - mi|$ ^2

For each cluster $Gi \in G$

$mi \leftarrow$ the mean value of all objects belongs to Gi

until no changes;

END

Before applying k-mean clustering algorithm we need to first find the optimal number of clusters. Either we go with random number of clusters initially or the best solution to make balance between maximum accuracy and maximum compression is to find optimal number of clusters. Here, the optimal number of clusters is found by calculating the within set sum of squared error (WSSSE). WSSSE is a metric that measure how good our clusters are. It works as:

- Step1: Error is calculated which is distance from each point of dataset to it is centroid (final centroid in each cluster)
- Step2: Square of that error is taken and finally summed up for entire dataset
- Step3: Plot the curve of WSSSE according to number of clusters

Following above step, WSSSE is just the measure of how far apart each point is from its centroid. Obviously, if there is lot of error in our model they will tend to be far apart from our centroid. So, the best choice to find the optimal number of clusters is to look at the elbow of WSSSE graph which is also known as the elbow method. The location of a bend (knee) in the plot is generally considered as an indicator of the appropriate number of clusters which is 40 as shown in the figure in our case.

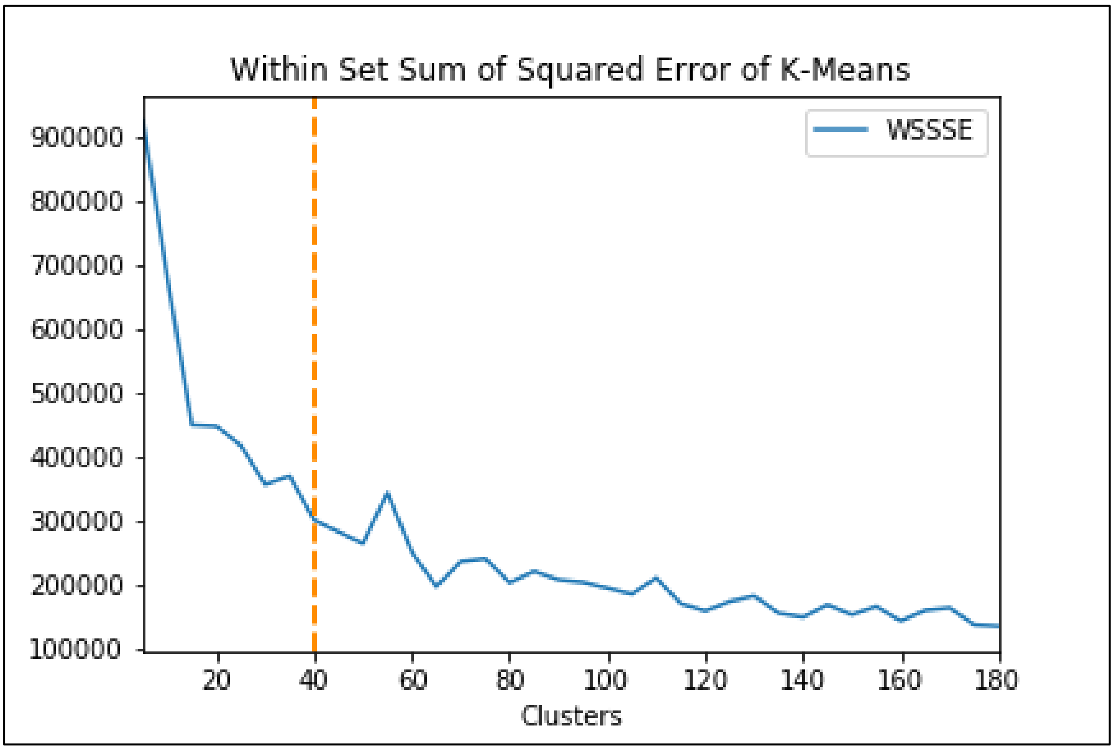


Figure 5.1.3. 1: WSSSE Of K-means

Once the optimal number of clusters for each user is calculated then the centers for each user is also calculated and visualizing the centers, it looks like this:

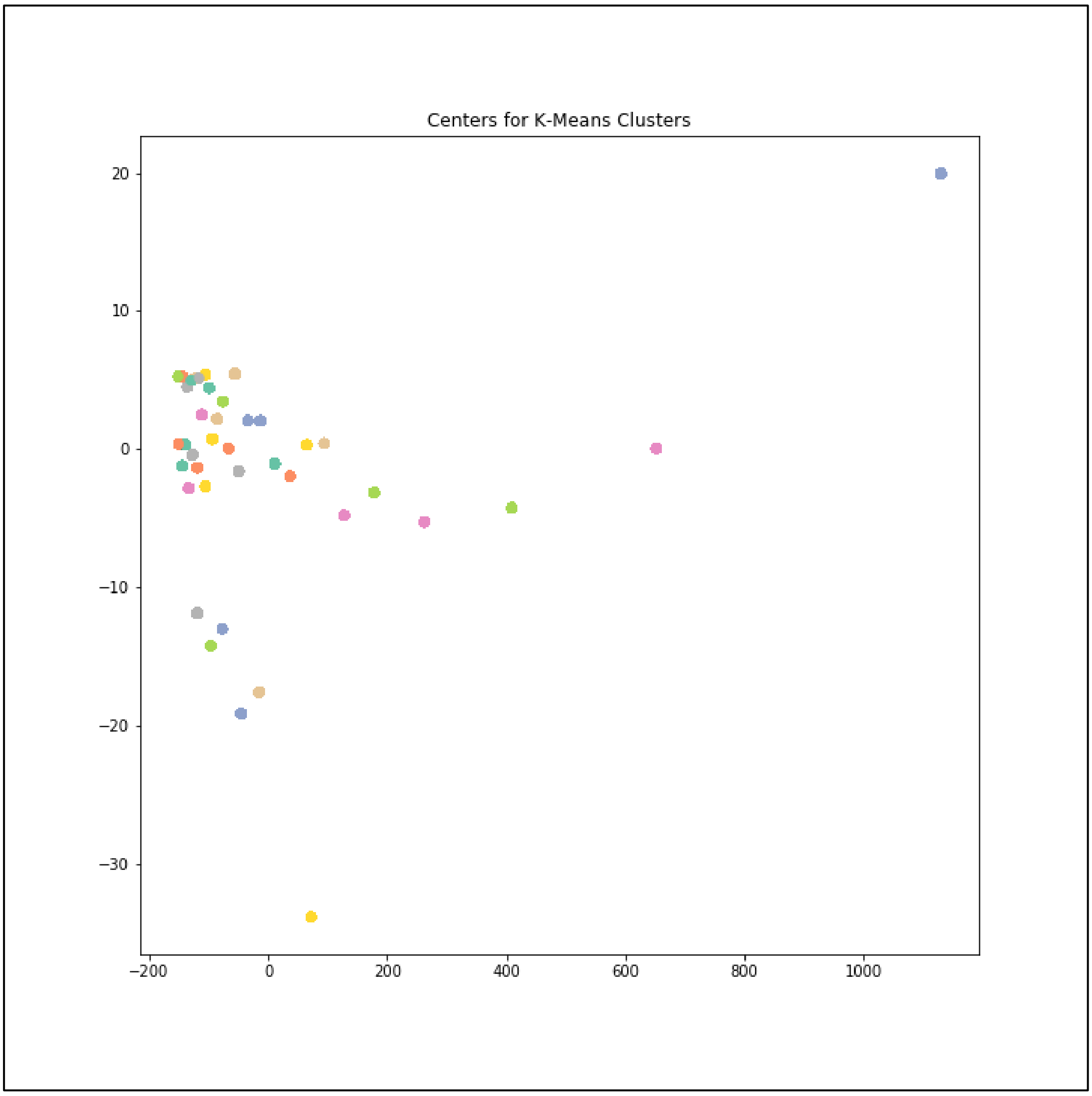


Figure 5.1.3. 2: Center Of K-means Clusters

After applying K-mean clustering, the result below shows the top common products in all cluster. Looking individually in all clusters, it was also found that banana, bag of organic banana, organic strawberries, organic hass avocado, Limes are top products in each cluster. In this way the most selling products is found.

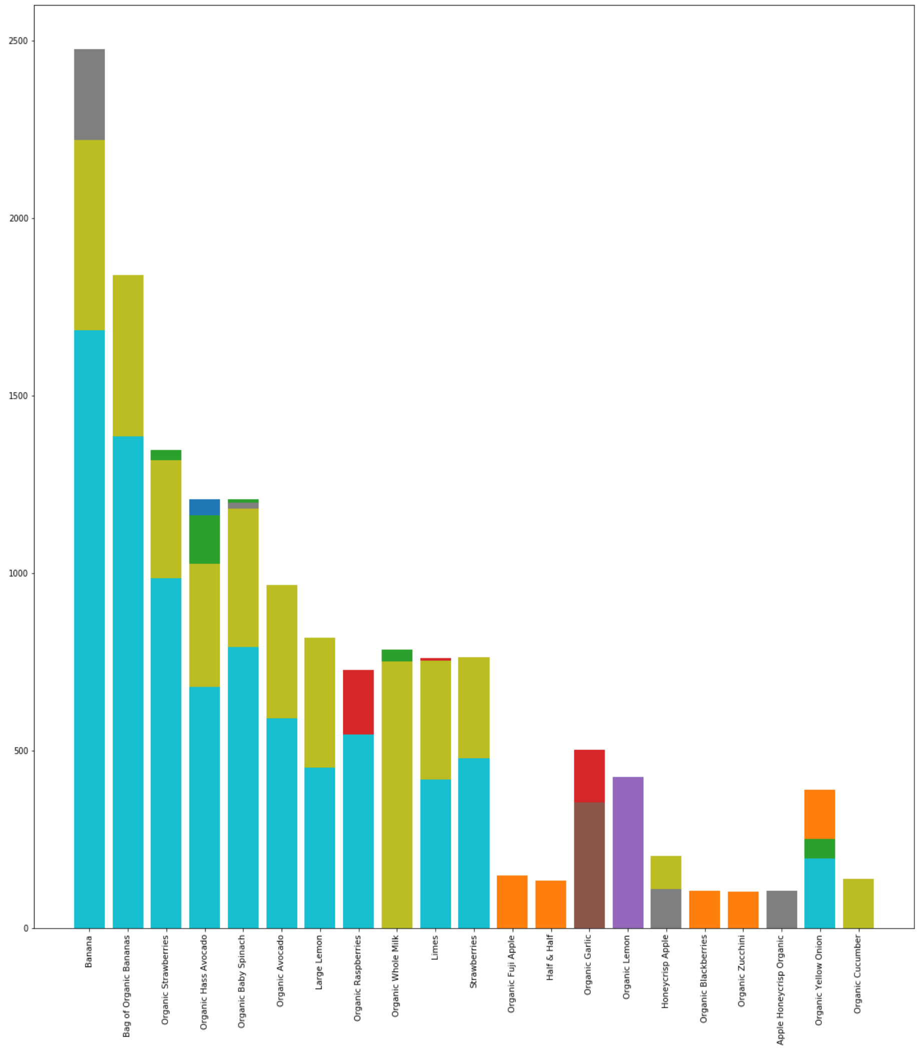


Figure 5.1.3. 3: MSPL using k-means clustering

5.1.4 Market Basket Using Bigram Frequency: To Generate Rules

The final task is to find out which products are frequently bought together and to generate rules. For this purpose, we will use bigram and count bigram frequency. A bigram is a sequence of two adjacent elements from a string of tokens, which are typically letters, syllabus or words. The frequency distribution of every bigram in a string is commonly used for simple statistical analysis of text in many applications, including in computational linguistics, cryptography, speech recognition, and so on (Collins, 1996). The probability of a token W_n given the preceding token W is equal to the probability of their n−1 bigram, or the co-occurrence of the two tokens P(W , W) divided by the probability of the preceding n−1 token.

Step1: Extract bigrams and calculate bigram frequency.

Bigram is extracted based on order id of each cluster that means in one order id of each cluster group which products have been bought together. So, in our case bigram will be in the format of

(

Order ID: [products bought in that order ID]

)

Once the bigram is extracted, the bigram frequency is calculated which is obtained by adding the same products bought in that order.

Step2: Save bigrams and bigram frequency to JSON file as rules.

Bigram are stored in a nested dictionary where

- First layer key is the first word in a bigram
- Second layer key is the second word in a bigram
- Second layer value is the frequency

Then, the dictionary is converted to JSON file as rules that will be used by recommendation algorithm to generate recommendations for each product.

The format for the rules are:

```
{
    Product Name: {
        [{Recommended product list: frequency}]
    }
}
```

5.1.5 Recommendation algorithm

We developed this algorithm for predicting product bundles; recommend a list of products that are likely to be bought together if a customer buy a certain product.

- Suppose we would like to recommend k products to be bought together after product S and the number of bigrams starting with S is denoted as n
- First, we sort the frequencies for each bigram starting with the product S in decreasing order, so we have n bigrams arranged in an order from the highest frequency to the lowest
- Second, we compare k with n, and fill the recommendation list with products in bigrams one by one based on the ordered bigram frequency. We consider the following two cases:
    - When $n \geq k$, if after sorting, there are some bigrams with the same frequency, we will pick each one with equal probability until the recommendation list reaches the total number of k
    - When $n \leq k$, we will not have enough product to fill the recommendation list of product S. In this case, we will first recommend all these n products. Then goes back to the product that is of highest frequency in bigrams of S, denoted as T, and fill the rest places in recommendation list with the products followed by T, adopting the same rule as before

Try an example: 15 Products recommended after "Organic_Mint_Bunch".

```
print(getRecommend("Organic_Mint_Bunch", 15))

['Coco_Crunch_Sprouted_Granola', 'Organic_Oat_Non-Dairy_Original_Beverage', 'Vine_Ripe_Tomatoes', 'Trilogy_Kombucha_Drink', 'Sherry_Vinegar', 'Banana', 'Organic_Navel_Orange', 'Organic_Sliced_Peaches', 'Banana', 'Peanut_Butter_Chocolate_Chip_Fruit_&_Nut_Food_Bar', 'Petite_Brussels_Sprouts', 'Flaky_Biscuits', 'Organic_Avocado', 'Original_Hummus', 'Cucumber_Kirby']
```

Figure 5.1.5. 1: Recommendation Algorithm Example

6. TESTING

Testing is advantageous in several ways. Firstly, the defects found help in the process of making the software reliable. Secondly, even if the defects found are not corrected, testing gives an idea as to how reliable the software is. Thirdly, over time, the record of defects found reveals the most common kinds of defects, which can be used for developing appropriate preventive measures such as training, proper design and reviewing.

6.1 Testing Plan

The testing sub-process includes the following activities in a phase dependent manner:

- Create Test Plans
- Create Test Specifications
- Review Test Plans and Test Specifications
- Conduct tests according to the Test Specifications, and log the defects
- Fix defects, if any
- When defects are fixed continue from activity

6.2 Testing Methods

6.2.1 Black Box and White Box Testing

In black box testing a software item is viewed as a black box, without knowledge of its internal structure or behavior. Possible input conditions, based on the specifications (and possible sequences of input conditions), are presented as test cases.

In white-box testing knowledge of internal structure and logic is exploited. Test cases are presented such that possible paths of control flow through the software item are traced. Hence more defects than black box testing is likely to be found.

The disadvantages are that exhaustive path testing is infeasible and the logic might not conform to specification. Instrumentation techniques can be used to determine the structural system coverage in white box testing. For this purposed tools or compilers that can insert test probes into the programs can be used.

6.3 Testing Principles

- While deciding on the focus of testing activities, study project priorities. For example, for an on-line system, pay more attention to response time. Spend more time on the features used frequently
- Decide on the effort required for testing based on the usage of the system. If the system is to be used by a large number of users, evaluate the impact on users due to a system failure before deciding on the effort
- A necessary part of the test case is a definition of the expected result
- Write test cases for invalid and unexpected as well as valid and expected input conditions
- Thoroughly inspect the results of each test

We have performed both Unit Testing and System Testing to detect and fix errors. A brief description of both is given below.

Unit Testing

Objective

The objective of Unit Testing is to test a unit of code (program or set of programs) using the Unit Test Specifications, after coding is completed. Since the testing will depend on the completeness and correctness of test specifications, it is important to subject these to quality and verification reviews.

Input

- Unit Test Specifications
- Code to be tested.

Testing Process

- Checking for availability of Code Walk-thru reports which have documented the existence of and conformance to coding standards
- Review of Unit Test Specifications
    - Verify the Unit Test Specifications conform to the program specifications
    - Verify that all boundary and null data conditions are included

Techniques

Test Walk-through

This method of reviewing modules, a check for testability, is done by mentally executing the code with example test cases. The advantage is that the programmer can ensure that the path taken is always the one intended and that the values of variables are always sensible. This is not a proven method and can be used only by persons with knowledge of the particular language or application. It is recommended that a number of test cases from the Unit Test Specifications be utilized for Test Walk-thru.

Two strategies are defined for testing modules by actual execution: the top down approach and the bottom-up approach.

In the top-down approach, the main routine is tested first, with the subroutines being substituted by dummy or null routines called stubs. As testing progresses, the stubs are replaced by real routines that in turn may call further stubs representing lower-level modules in the module hierarchy.

In the bottom-up approach, testing starts at the lowermost level and then proceeds to higher-level programs or modules. The testing team directly calls the module to be tested, which may call other pre-tested modules.

Integration Test

After completion of our module along with testing, modular coding strategy was used. After integrating the module with the complete application, time was given to our team to test their part of module completely and thoroughly.

As the whole application is divided into several modules, there were a lot of variable names and function names, which were common to all the modules. There existed a lot of session variables, which we had to incorporate into our module, but as different modules were being developed simultaneously, we had to hard code things in place of the session variables in our module. So, at the time of integration a lot of hard coded things had to be removed and session variables were replaced.

6.4 Test Cases

Sr. No	Test Case Description	Expected Result	Actual Result
1.	No Username and Password	• The Home page should be displayed • Able to add products to cart • Unable to go to checkout page	Same as expected.
2.	Correct Username and password	• Home page should be displayed with logged in Username • Able to add products to cart • Able to go to checkout page	Same as expected.
3.	Incorrect username and password	• Checkout button disabled • Error message telling incorrect username and password	Same as expected.
4.	Fill in field of search for products and then click on submit.	• Give suggestion for similar product available • Give recommendation for next five products to buy • Add product to cart	Same as expected.
5.	Fill in field of search for product by selecting n number of recommendation and then click on add to cart	• Give recommendation for next n products that are likely to be purchased next • Add product to cart	Same as expected.

6.	Select product from recommendation or product available list and click on add to cart button	• Add product to cart • Give recommendation for 10 next product that are likely to be purchased next.	Same as expected.
7.	Click on cart icon	• Popup shopping cart window • Display "No items added to cart ", message if cart is empty • Display products added to cart	Same as expected.
8.	Click on cross icon from product displayed on shopping cart windows	• Delete product from cart • Display product deleted from cart message	Same as expected.
9.	Fill in quantity field from product displayed on shopping cart windows.	• Update product to cart with filled numbers of quantity.	Same as expected.
10.	Clicked on checkout button	• Display shipping address form • Display products added to cart with quantity • Display total price for products	Same as expected.
11.	Fill in shopping address form and click continue to checkout button	• Save User Details and Product to database	Same as expected.

		• Send email to user with product they ordered along with quantity and price	
12.	Click on Log Out button	• Logout user and go to home page	Same as expected.

7. RESULT

- Recommending customers with favorite products
- Recommending the bundles of products to customers
- Predicting which product will the customer buy
- Helping E-commerce website on increasing sales

8. CONCLUSION AND ENHANCEMENT

8.1 Conclusion

Word2vector, k-means algorithm and different algorithmic processes have helped to generate market basket of the product very effectively. We implemented three major functions - recommend new products to customers, recommend product bundles, build predictive modeling on a user's next order - on the Instacart public datasets. In terms of recommendation, we are predicting accuracy around 20%, that means, 1 in 5 products we recommend using our algorithm is actually purchased by the customer. We do think it can create business value for Instacart and same techniques may transfer to other industries.

Future Improvements could be made in this aspects:

- Create more correlated features

8.2 Limitation

- Since the datasets doesn't contain the price tag of products, we have to randomly tag price to product

8.3 Future Enhancement

- None of payment method is integrated, Integration of different payment method will enhance the system.

9. REFERENCES

[1] M. Gholamian, "A novel model for product bundling and direct marketing in e-commerce based on market segmentation," Decision Science Letters, vol. 7, pp. 39-54, january ,2018.

[2] M. Venkatesh, "The design and pricing of bundles. A review of normative guidelines and practical approaches.," Handbook of Pricing Reasearch in Marketing., pp. 232-257, January,2009.

[3] P. C. Dumez Herve, "The Role of Bundlling in Firms' Marketing Strategies: A Synthesis," Recherche et Application en Marketing, vol. 27, pp. 91-105, 2012/06.

[4] T.-C. Y. Hsiangchu Lai, "Comparison of product bundling strategies on different online," Science Direct, p. 295–304, 12 June 2006.

[5] L. YanFang, "Bundle-Pricing Decision Model for Multiple Products," Procedia Computer Science, vol. 112, no. 1, pp. 2147-2154, 2017.

[6] B. M. V. k. Bahman Bahmani, "Scalable k-means++," Proceedings of the VLDB Endowment, vol. 5, no. 7, pp. 622-633, March 2012.

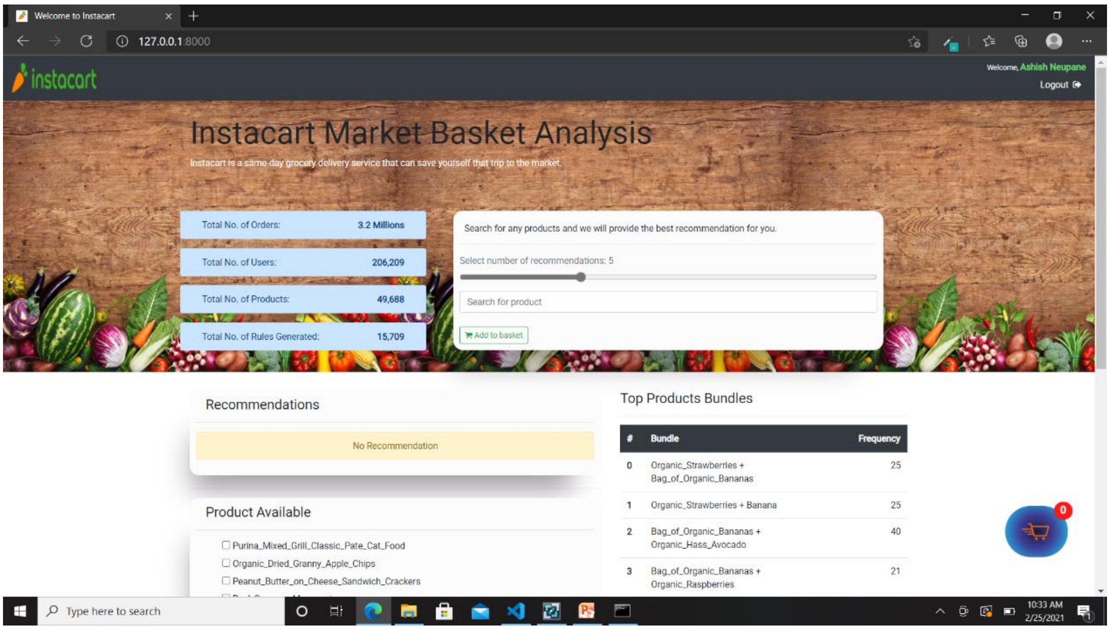
[7] J. S. L. Yuanchun Jiang, "Maximizing customer satisfaction through an online recommendation system: A novel associative classification model," Decision Support Systems, vol. 48, no. 3, pp. 470-479, February 2010.

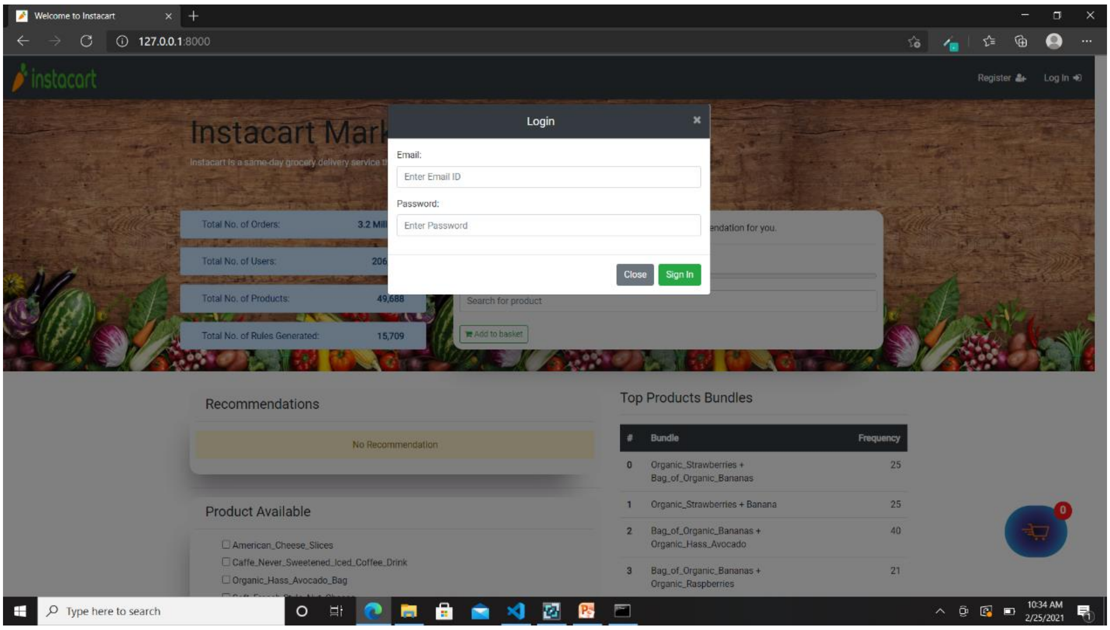
[8] T. W. M. K. Y. F. Wei Chen, "NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems," LightGBM: a highly efficient gradient boosting decision tree, vol. 12, no. 1, p. 3149–3157, December 2017.

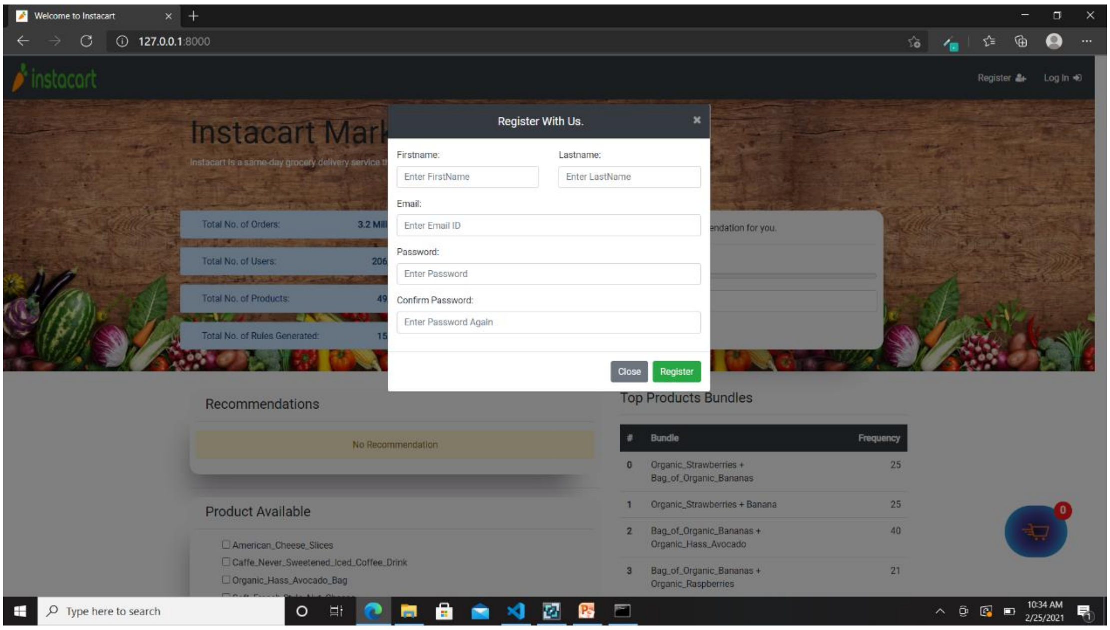
[9] Q. N. B. S. Owies Mais, "Market Basket Analysis," Management Information Systems, vol. 6, no. 1, pp. 50-62, 2017.

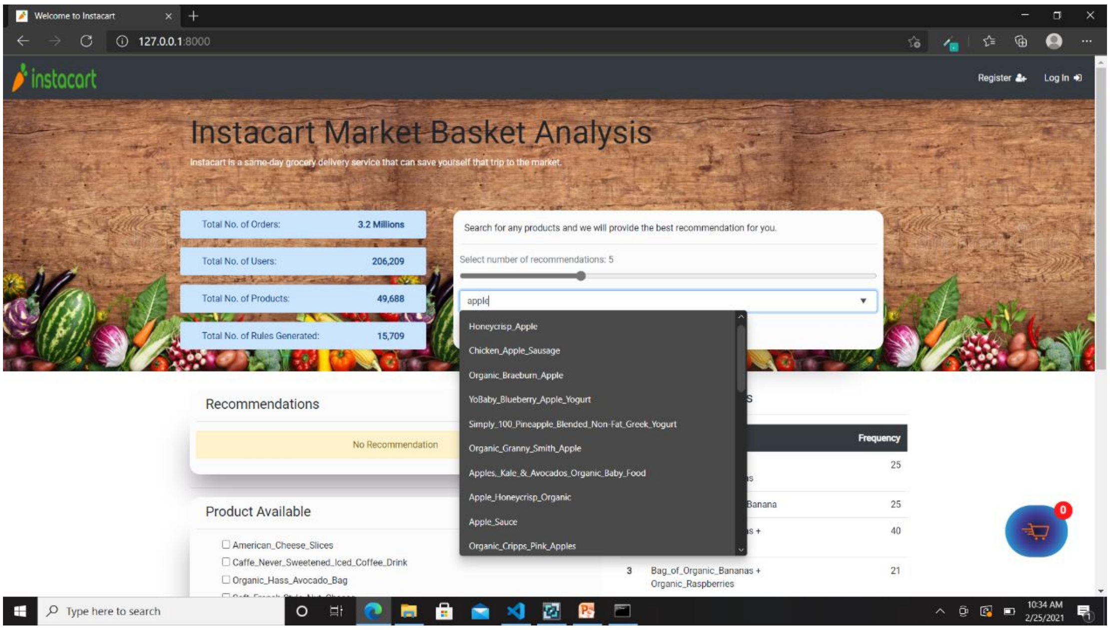
[10] M. M. B. L. Apoorva Mehta, "Instacart," Instacart.com, 2012. [Online]. Available: https://www.instacart.com/datasets/grocery-shopping-2017. [Accessed 25 Feburary 2020].

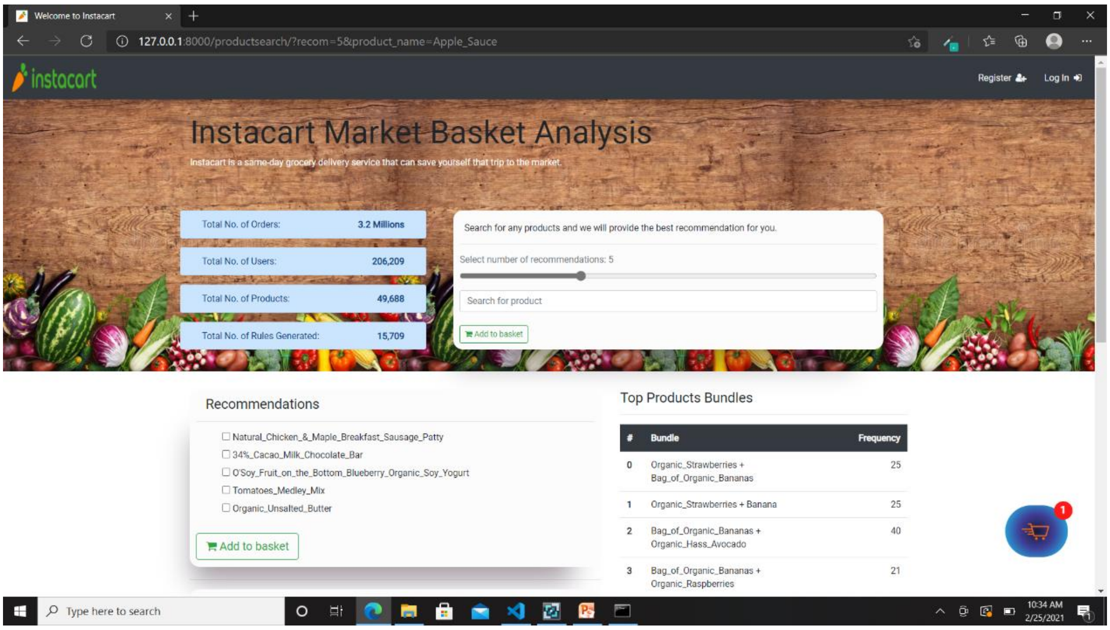
APPENDIX

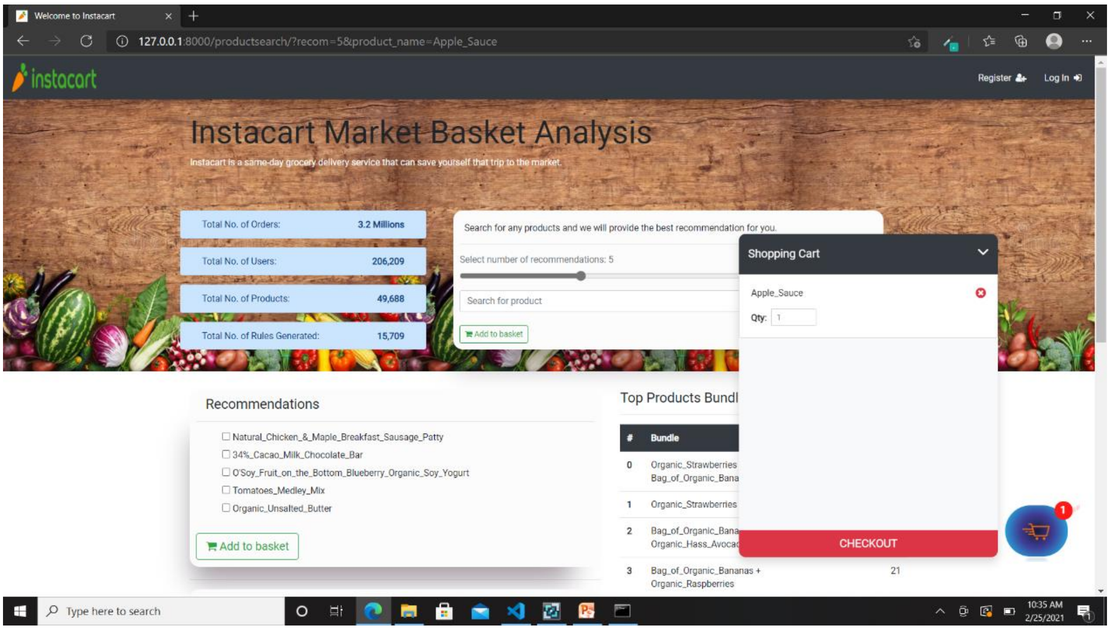

Welcome to Instacart
127.0.0.1:8000
instacart
Welcome, Ashish Neupane
Logout
Instacart Market Basket Analysis
Instacart is a same-day grocery delivery service that can save yourself that trip to the market.
Total No. of Orders: 3.2 Millions
Total No. of Users: 206,209
Total No. of Products: 49,688
Total No. of Rules Generated: 15,709
Search for any products and we will provide the best recommendation for you.
Select number of recommendations: 5
Search for product
Add to basket
Recommendations
No Recommendation
Product Available
Purina_Mixed_Grill_Classic_Pate_Cat_Food
Organic_Dried_Granny_Apple_Chips
Peanut_Butter_on_Cheese_Sandwich_Crackers
Top Products Bundles
Bundle Frequency
0 Organic_Strawberries + Bag_of_Organic_Bananas 25
1 Organic_Strawberries + Banana 25
2 Bag_of_Organic_Bananas + Organic_Hass_Avocado 40
3 Bag_of_Organic_Bananas + Organic_Raspberries 21
Type here to search
10:33 AM
2/25/2021

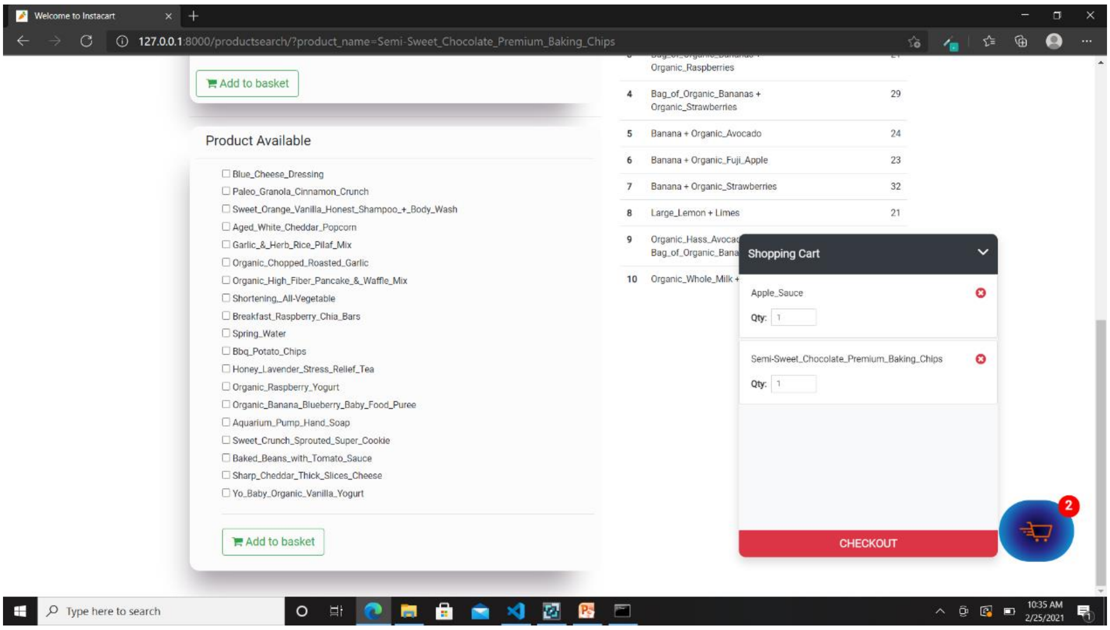

Welcome to Instacart
127.0.0.1:8000
instacart
Register
Log In
Login
Email:
Enter Email ID
Password:
Enter Password
Close
Sign In
Recommendations
No Recommendation
Product Available
American_Cheese_Slices
Caffe_Never_Sweetened_Iced_Coffee_Drink
Organic_Hass_Avocado_Bag
Top Products Bundles
Bundle Frequency
0 Organic_Strawberries + Bag_of_Organic_Bananas 25
1 Organic_Strawberries + Banana 25
2 Bag_of_Organic_Bananas + Organic_Hass_Avocado 40
3 Bag_of_Organic_Bananas + Organic_Raspberries 21
Type here to search
10:34 AM
2/25/2021

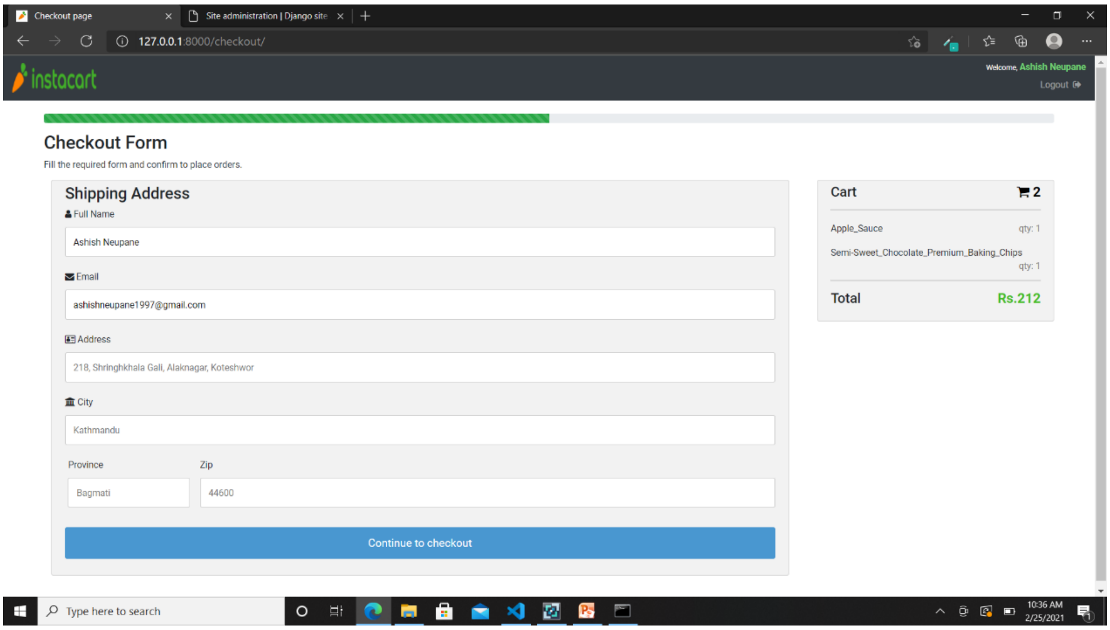

Welcome to Instacart
127.0.0.1:8000
instacart
Register
Log In
Register With Us.
Firstname:
Enter FirstName
Lastname:
Enter LastName
Email:
Enter Email ID
Password:
Enter Password
Confirm Password:
Enter Password Again
Close
Register
Recommendations
No Recommendation
Product Available
American_Cheese_Slices
Caffe_Never_Sweetened_Iced_Coffee_Drink
Organic_Hass_Avocado_Bag
Top Products Bundles
Bundle Frequency
0 Organic_Strawberries + Bag_of_Organic_Bananas 25
1 Organic_Strawberries + Banana 25
2 Bag_of_Organic_Bananas + Organic_Hass_Avocado 40
3 Bag_of_Organic_Bananas + Organic_Raspberries 21
Type here to search
10:34 AM
2/25/2021

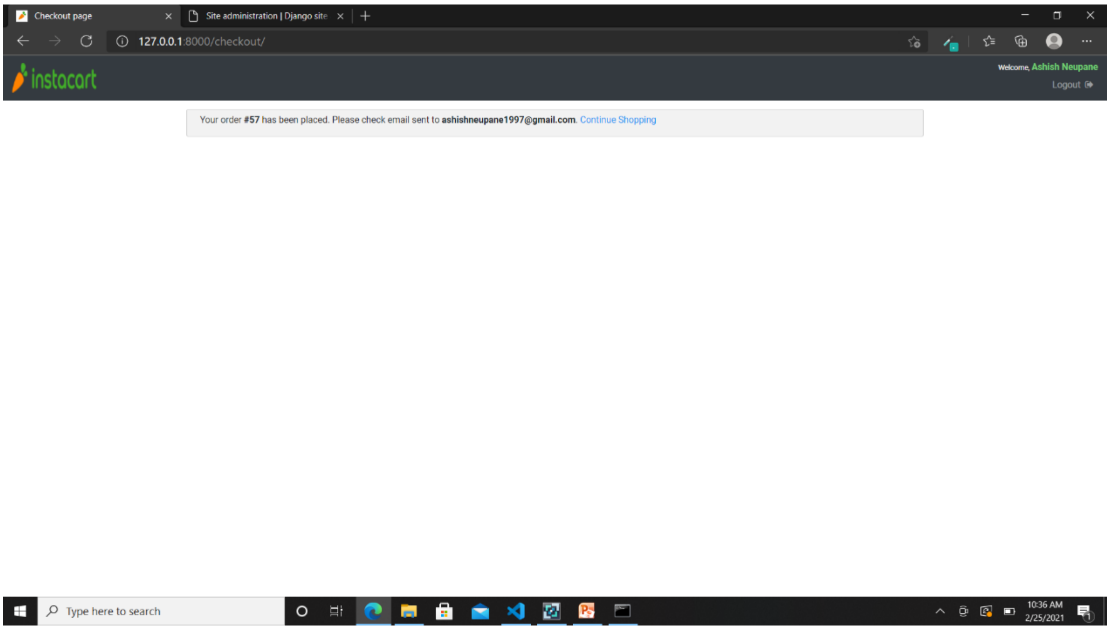

Welcome to Instacart
127.0.0.1:8000
instacart
Register
Log In
Instacart Market Basket Analysis
Instacart is a same-day grocery delivery service that can save yourself that trip to the market.
Total No. of Orders: 3.2 Millions
Total No. of Users: 206,209
Total No. of Products: 49,688
Total No. of Rules Generated: 15,709
Search for any products and we will provide the best recommendation for you.
Select number of recommendations: 5
apple
Honeycrisp_Apple
Chicken_Apple_Sausage
Organic_Braeburn_Apple
YoBaby_Blueberry_Apple_Yogurt
Simply_100_Pineapple_Blended_Non-Fat_Greek_Yogurt
Organic_Granny_Smith_Apple
Apples,_Kale_&_Avocados_Organic_Baby_Food
Apple_Honeycrisp_Organic
Apple_Sauce
Organic_Cripps_Pink_Apples
Recommendations
No Recommendation
Product Available
American_Cheese_Slices
Caffe_Never_Sweetened_Iced_Coffee_Drink
Organic_Hass_Avocado_Bag
Frequency
3 Bag_of_Organic_Bananas + Organic_Raspberries 21
Type here to search
10:34 AM
2/25/2021

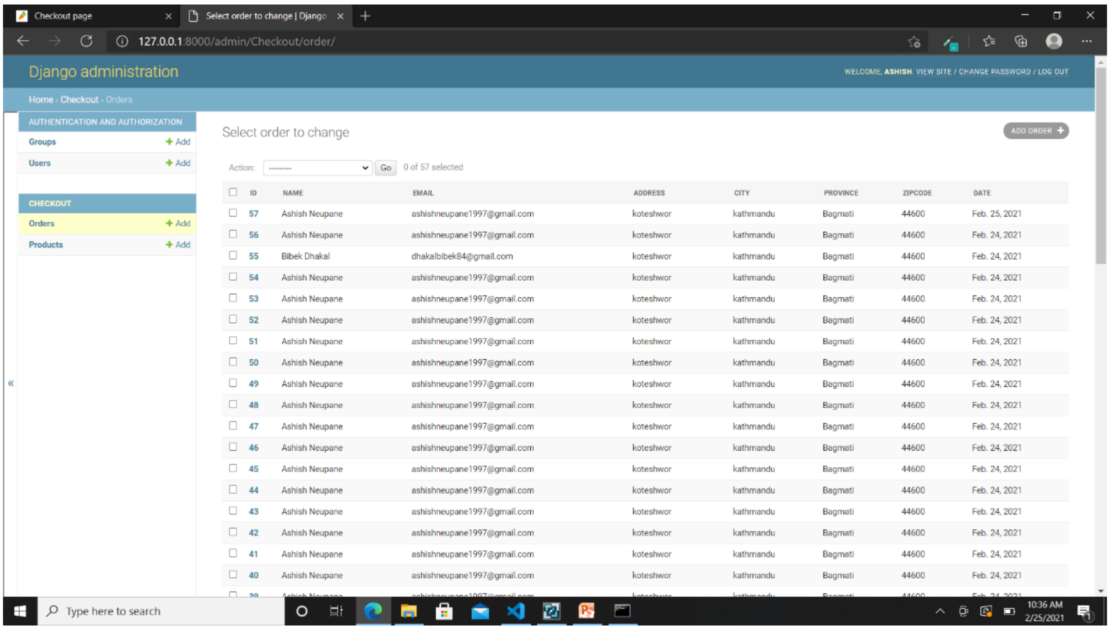

Welcome to Instacart
127.0.0.1:8000/productsearch/?recom=5&product_name=Apple_Sauce
instacart
Register
Log In
Instacart Market Basket Analysis
Instacart is a same-day grocery delivery service that can save yourself that trip to the market.
Total No. of Orders: 3.2 Millions
Total No. of Users: 206,209
Total No. of Products: 49,688
Total No. of Rules Generated: 15,709
Search for any products and we will provide the best recommendation for you.
Select number of recommendations: 5
Search for product
Add to basket
Recommendations
Natural_Chicken_&_Maple_Breakfast_Sausage_Patty
34%_Cacao_Milk_Chocolate_Bar
O'Soy_Fruit_on_the_Bottom_Blueberry_Organic_Soy_Yogurt
Tomatoes_Medley_Mix
Organic_Unsalted_Butter
Add to basket
Top Products Bundles
Bundle Frequency
0 Organic_Strawberries + Bag_of_Organic_Bananas 25
1 Organic_Strawberries + Banana 25
2 Bag_of_Organic_Bananas + Organic_Hass_Avocado 40
3 Bag_of_Organic_Bananas + Organic_Raspberries 21
1
Type here to search
10:34 AM
2/25/2021

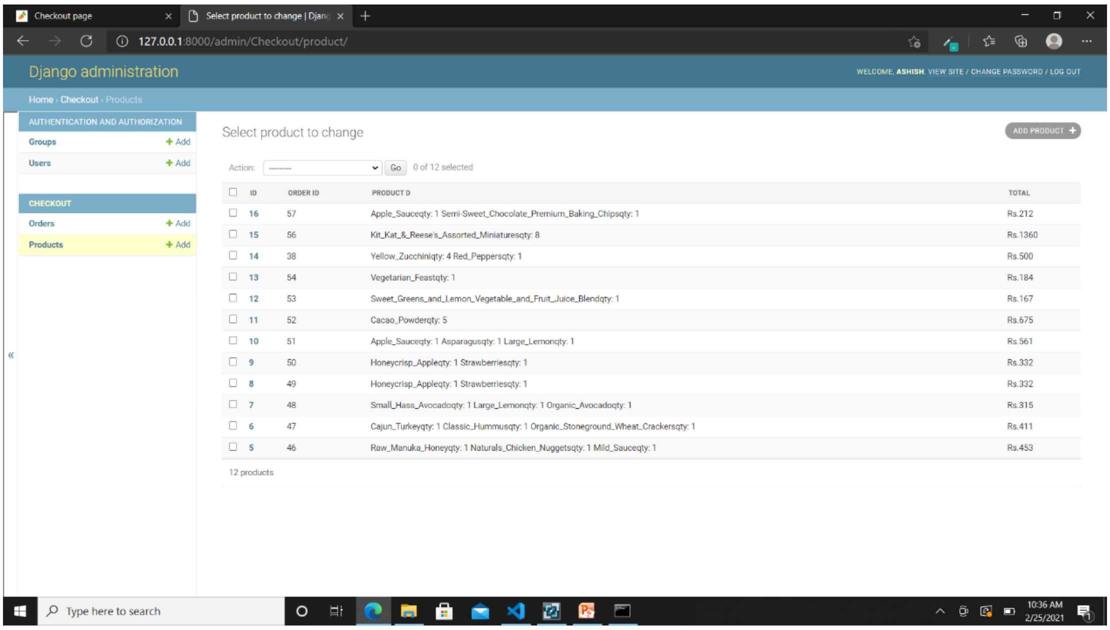

Welcome to Instacart
127.0.0.1:8000/productsearch/?recom=5&product_name=Apple_Sauce
instacart
Register
Log In
Instacart Market Basket Analysis
Instacart is a same-day grocery delivery service that can save yourself that trip to the market.
Total No. of Orders: 3.2 Millions
Total No. of Users: 206,209
Total No. of Products: 49,688
Total No. of Rules Generated: 15,709
Search for any products and we will provide the best recommendation for you.
Select number of recommendations: 5
Search for product
Add to basket
Shopping Cart
Apple_Sauce
Qty: 1
CHECKOUT
Recommendations
Natural_Chicken_&_Maple_Breakfast_Sausage_Patty
34%_Cacao_Milk_Chocolate_Bar
O'Soy_Fruit_on_the_Bottom_Blueberry_Organic_Soy_Yogurt
Tomatoes_Medley_Mix
Organic_Unsalted_Butter
Add to basket
Top Products Bundl
Bundle
0 Organic_Strawberries Bag_of_Organic_Bana
1 Organic_Strawberries
2 Bag_of_Organic_Bana Organic_Hass_Avocad
3 Bag_of_Organic_Bananas + Organic_Raspberries 21
1
Type here to search
10:35 AM
2/25/2021

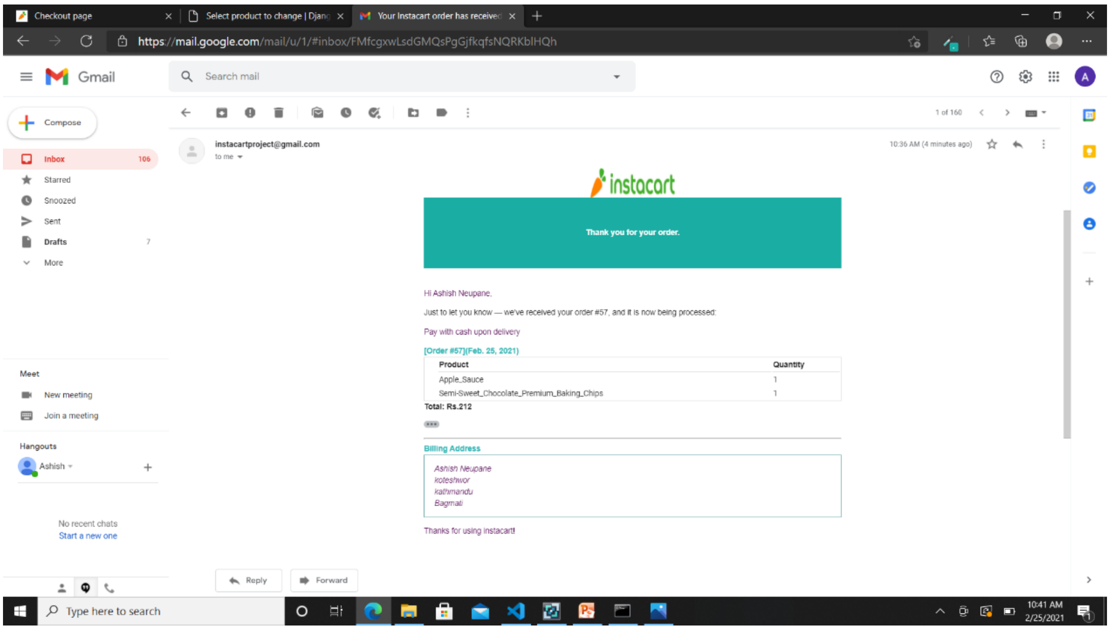

Welcome to Instacart
127.0.0.1:8000/productsearch/?product_name=Semi-Sweet_Chocolate_Premium_Baking_Chips
Add to basket
Product Available
Blue_Cheese_Dressing
Paleo_Granola_Cinnamon_Crunch
Sweet_Orange_Vanilla_Honest_Shampoo_+_Body_Wash
Aged_White_Cheddar_Popcorn
Garlic_&_Herb_Rice_Pilaf_Mix
Organic_Chopped_Roasted_Garlic
Organic_High_Fiber_Pancake_&_Waffle_Mix
Shortening_All-Vegetable
Breakfast_Raspberry_Chia_Bars
Spring_Water
Bbq_Potato_Chips
Honey_Lavender_Stress_Relief_Tea
Organic_Raspberry_Yogurt
Organic_Banana_Blueberry_Baby_Food_Puree
Aquarium_Pump_Hand_Soap
Sweet_Crunch_Sprouted_Super_Cookie
Baked_Beans_with_Tomato_Sauce
Sharp_Cheddar_Thick_Slices_Cheese
Yo_Baby_Organic_Vanilla_Yogurt
Add to basket
Organic_Raspberries
4 Bag_of_Organic_Bananas + Organic_Strawberries 29
5 Banana + Organic_Avocado 24
6 Banana + Organic_Fuji_Apple 23
7 Banana + Organic_Strawberries 32
8 Large_Lemon + Limes 21
9 Organic_Hass_Avocad Bag_of_Organic_Bana
10 Organic_Whole_Milk +
Shopping Cart
Apple_Sauce
Qty: 1
Semi-Sweet_Chocolate_Premium_Baking_Chips
Qty: 1
CHECKOUT
2
Type here to search
10:35 AM
2/25/2021


Checkout page
Site administration | Django site
127.0.0.1:8000/checkout/
instacart
Welcome, Ashish Neupane
Logout
Checkout Form
Fill the required form and confirm to place orders.
Shipping Address
Full Name
Ashish Neupane
Email
ashishneupane1997@gmail.com
Address
218, Shringhkhala Gali, Alaknagar, Koteshwor
City
Kathmandu
Province
Bagmati
Zip
44600
Continue to checkout
Cart
2
Apple_Sauce
qty: 1
Semi-Sweet_Chocolate_Premium_Baking_Chips
qty: 1
Total
Rs.212
Type here to search
10:36 AM
2/25/2021


Checkout page
Site administration | Django site
127.0.0.1:8000/checkout/
instacart
Welcome, Ashish Neupane
Logout
Your order #57 has been placed. Please check email sent to ashishneupane1997@gmail.com. Continue Shopping
Type here to search
10:36 AM
2/25/2021


Checkout page
Select order to change | Django
127.0.0.1:8000/admin/Checkout/order/
Django administration
WELCOME, ASHISH. VIEW SITE / CHANGE PASSWORD / LOG OUT
Home › Checkout › Orders
AUTHENTICATION AND AUTHORIZATION
Groups + Add
Users + Add
CHECKOUT
Orders + Add
Products + Add
Select order to change
ADD ORDER +
Action: --------- Go 0 of 57 selected
ID NAME EMAIL ADDRESS CITY PROVINCE ZIPCODE DATE
57 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 25, 2021
56 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
55 Bibek Dhakal dhakalbibek84@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
54 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
53 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
52 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
51 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
50 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
49 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
48 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
47 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
46 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
45 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
44 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
43 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
42 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
41 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
40 Ashish Neupane ashishneupane1997@gmail.com koteshwor kathmandu Bagmati 44600 Feb. 24, 2021
Type here to search
10:36 AM
2/25/2021


Checkout page
Select product to change | Django
127.0.0.1:8000/admin/Checkout/product/
Django administration
WELCOME, ASHISH. VIEW SITE / CHANGE PASSWORD / LOG OUT
Home › Checkout › Products
AUTHENTICATION AND AUTHORIZATION
Groups + Add
Users + Add
CHECKOUT
Orders + Add
Products + Add
Select product to change
ADD PRODUCT +
Action: Go 0 of 12 selected
ID ORDER ID PRODUCT D TOTAL
16 57 Apple_Sauceqty: 1 Semi-Sweet_Chocolate_Premium_Baking_Chipsqty: 1 Rs.212
15 56 Kit_Kat_&_Reese's_Assorted_Miniaturesqty: 8 Rs.1360
14 38 Yellow_Zucchiniqty: 4 Red_Peppersqty: 1 Rs.500
13 54 Vegetarian_Feastqty: 1 Rs.184
12 53 Sweet_Greens_and_Lemon_Vegetable_and_Fruit_Juice_Blendqty: 1 Rs.167
11 52 Cacao_Powderqty: 5 Rs.675
10 51 Apple_Sauceqty: 1 Asparagusqty: 1 Large_Lemonqty: 1 Rs.561
9 50 Honeycrisp_Appleqty: 1 Strawberriesqty: 1 Rs.332
8 49 Honeycrisp_Appleqty: 1 Strawberriesqty: 1 Rs.332
7 48 Small_Hass_Avocadoqty: 1 Large_Lemonqty: 1 Organic_Avocadoqty: 1 Rs.315
6 47 Cajun_Turkeyqty: 1 Classic_Hummusqty: 1 Organic_Stoneground_Wheat_Crackersqty: 1 Rs.411
5 46 Raw_Manuka_Honeyqty: 1 Naturals_Chicken_Nuggetsqty: 1 Mild_Sauceqty: 1 Rs.453
12 products
Type here to search
10:36 AM
2/25/2021


Checkout page
Select product to change | Django
Your Instacart order has received
https://mail.google.com/mail/u/1/#inbox/FMfcgxwLsdGMQsPgGjfkqfsNQRKblHQh
Gmail
Search mail
Compose
Inbox 106
Starred
Snoozed
Sent
Drafts 7
More
Meet
New meeting
Join a meeting
Hangouts
Ashish
No recent chats
Start a new one
instacartproject@gmail.com
to me
10:36 AM (4 minutes ago)
1 of 160
instacart
Thank you for your order.
Hi Ashish Neupane,
Just to let you know — we've received your order #57, and it is now being processed:
Pay with cash upon delivery
[Order #57](Feb. 25, 2021)
Product Quantity
Apple_Sauce 1
Semi-Sweet_Chocolate_Premium_Baking_Chips 1
Total: Rs.212
Billing Address
Ashish Neupane
koteshwor
kathmandu
Bagmati
Thanks for using instacart!
Reply
Forward
Type here to search
10:41 AM
2/25/2021