Geographic Recommendation System for Supply Chain Management

Prashnim Seth, Harshita Singh, Wassnaa Al-Mawee
Department of Statistics, Western Michigan University

Abstract

An efficient supply chain means a robust and nimble commercial enterprise. Timely and cost-effective product delivery is a crucial part of the supply chain cycle. On the flip side, late, and failed deliveries would result in the company losing its customers. To keep the efficiency of the supply chain cycle intact, this paper aims to build a recommender system that analyses the demand of a selected product, groups it into regions and suggests a suitable location for a business or a warehouse as per the product. The paper follows a content-based filtering approach by applying an unsupervised learning algorithm – K-Means Clustering to develop a system that calculates the best locations for the selected product using different order locations from the dataset to generate a location recommendation for storing the product in a closer warehouse. By analyzing product demands as per the clustered order locations, this system offers valuable insights into optimal warehouse placement, resulting in potentially reduced delivery times.

Keywords: Recommender System, Clustering, K-Means Clustering, Content-Based Filtering, Warehouse, Unsupervised Learning.

I. **Introduction**

A Recommendation System is a tool that provides suggestions for a service or a product. These systems help with retaining customers, increasing sales, boosting cart value, speeding up the work pace and much more. A lot of big companies use recommendation systems with their products. For example, Amazon uses recommendation system to suggest similar or more products based on the customer's search pattern, Netflix uses it to suggest similar movies or series to keep their customers hooked, Spotify uses it to recommend songs of the same genre or a similar artist.

Usually, recommendation systems are used for online services, but they can be used in multiple traditional offline services too. Recommendation Systems can help suggest restaurants with specific cuisines, find locations with more customers or drive business value by suggesting in-store merchandizing strategies. The goal of this paper is to build a recommender system that will suggest a suitable location for the product warehouses to reduce the shipping time, delays and increase sales of the product. This will be done by identifying locations where most customers buy a specific product and its quantities and to suggest the best location for a warehouse.

Recommendation Systems are generally of three types,

1. Content-based filtering - uses the attributes or features of an item to recommend other items like the user's preferences. [1]
2. Collaborative filtering - recommend items based on preference information from many users. [1]
3. Hybrid recommender systems - combine the advantages of the types above to create a more comprehensive recommending system. [1]

Content based filtering technique will best suit this work. The technique will be used to compare orders, products, and delivery locations to suggest the best location for a warehouse for the products.

A recommendation system like this would benefit the business. It could save a lot of money and time spent delivering the order. It could also be developed further by adding more features like using profit per order to improve the system. The use case of this application can be extended to small stores by recommending them where they would find the most customers.

II. **Related Work**

A lot of work has been done in developing recommender systems and improving the supply chain cycle.

A Movie Recommender System: MOVREC [2] In this paper, the authors present a recommender system that uses collaborative filtering to recommend movies to the users. They took the weights to their selected features - genre, actor, director, year, and rating of the movies to develop an algorithm that uses the KNN model with Euclidean distance to suggest movies to their users.

Machine Learning Techniques for Recommender Systems – A Comparative Case Analysis [3] This paper compares the design techniques of recommender systems. They compared techniques are Memory based collaborative filtering, Model based collaborative filtering, Hybrid collaborative filtering, Content based filtering, Hybrid Filtering and Computational-Intelligence based. Furthermore, the representative algorithms of the mentioned techniques are compared along with their advantages, disadvantages, and their applications in different industries. They also discuss the challenges with recommender systems like Overspecialization problem, limited content analysis problem etc. The paper at last explains the various evaluation metrics for recommender systems.

A Survey Paper on E-Learning Recommender System [4] The author in this paper has talked about customizing e-learning using web mining techniques. They discuss giving users learning tasks based on their previous tasks, their previous success doing the tasks and identifying patterns between other similar users. The paper also explains various filtering models that could be used to develop recommendations for e-learning.

Application of Content-Based Approach in Research Paper Recommendation System for a Digital Library [5] This paper discusses the need for recommendation systems for digital libraries to reduce information overload. The project uses the content-based filtering approach since the content here is more than the users. An algorithm is designed that applies TF-IDF weighing scheme and cosine similarity measure to develop the recommendation system.

Food Recommender Systems: Important Contributions, Challenges and Future Research Directions [6] This paper speaks on the necessity and importance of food recommender systems. The authors compare various approaches for selecting the best approach for their model. They use preferential leaning to develop this recommender system by assigning weights to the user’s preference, health factors and allergies. The paper also discusses how the socio-economic factors can be used as another approach while developing this system.

The papers [2] - [6] describe different ways and use cases for which recommendation system can be built.

Data-Driven Machine Learning System for Optimization of Processes Supporting the Distribution of Goods and Services – a case study [7] This paper named titled as "Data-Driven Machine Learning System for Optimization of Processes Supporting the Distribution of Goods and Services: A Case Study, "is a case study of the application of a data-driven machine learning system to improve processes supporting the distribution of goods and services is presented. The study focuses on the application of machine learning algorithms to assess data and forecast how effective procedures will be, enabling real-time optimization and enhanced performance. In addition to discussing the difficulties encountered in the system's development and deployment, the authors give the system's implementation results.

Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities [8] Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) are the models used in this study to process and evaluate text data from news stories, and a fully connected neural network is used to forecast stock values. The authors test their proposed model against a number of well-established machine learning techniques using a huge dataset of financial news stories and stock prices from a stock market index. In terms of prediction accuracy and stability, they discover that the deep learning model performs better than the conventional approaches. The findings imply that including news articles in stock price forecasting can be advantageous and offer a fresh viewpoint for financial research. The study offers insightful information on the application of deep learning.

The papers [7] & [8] discuss different ways to understand and improve the supply chain cycle.

Hub location problems: A review of models, classification, solution techniques, and applications [9] The paper "Hub Location Problems: A Review of Models, Classification, Solution Techniques, and Applications" provides a comprehensive overview of hub location problems in the field of operations research and logistics. The authors start by defining hub location problems, which involve the selection of facilities (hubs) to serve as intermediaries between suppliers and customers. They then present a classification of hub location problems into different categories, including p-median, p-center, and capacitated hub location problems. The authors also review the various solution techniques used to solve hub location problems, including mathematical programming, heuristics, and metaheuristics. They discuss the strengths and weaknesses of each approach and provide examples of real-world applications where these techniques have been used to solve hub location problems in transportation, telecommunications, and other industries. Finally, the author summarizes the current state of research in the field of hub location problems and identifies potential avenues for future research. They conclude that hub location problems are a rich and challenging area of study, with significant real-world applications and a need for continued development of new and improved solution techniques.

Using clustering analysis in a capacitated location-routing problem [10] In this paper the main goal is to determine the best locations for facilities (such as warehouses or distribution centers) to serve customers, as well as the most efficient routes for delivery vehicles to take between these facilities and customers. To preprocess the customer data and cluster them, the authors suggest utilizing clustering analysis, a machine learning technique. Following that, the capacitated LRP solution is guided by the clustering in an effort to increase the solution's overall effectiveness and quality. The results are contrasted with those attained using conventional techniques as the authors assess their methodology using a real-world dataset. In terms of both the quality of the solution and the computational time needed to get it, they discover that their method produces better solutions. The authors conclude by demonstrating the ability of clustering analysis to enhance the resolution of capacitated location-routing problems and arguing that this strategy has promising applications in a variety of industries, including transportation, and distribution.

A unified warehouse location-routing methodology [11] In order to optimize the locations of warehouses and the routes used by delivery vehicles in a logistics system, a new method is presented in this study. The authors suggest a unified approach that incorporates routing and warehouse site considerations into a single optimization problem. The authors begin by going over prior studies in the area of warehouse location-routing issues and pointing out the shortcomings of existing approaches. They next go over their suggested approach, which unifies routing and location decisions for warehouses into a single mathematical model. The model takes into consideration variables including client demand, delivery truck capacity, and warehousing capacity. The authors assess their strategy using actual datasets and contrast the outcomes with those attained through the use of conventional techniques. They discover that their strategy produces solutions that are better in terms of both the solution's quality and the amount of time needed to compute them. The authors present a fresh and cohesive method for resolving warehouse location-routing issues, and they show how it has the potential to boost the efficacy and efficiency of logistics systems. Their approach has potential implications in several industries, including distribution, supply chain management, and transportation.

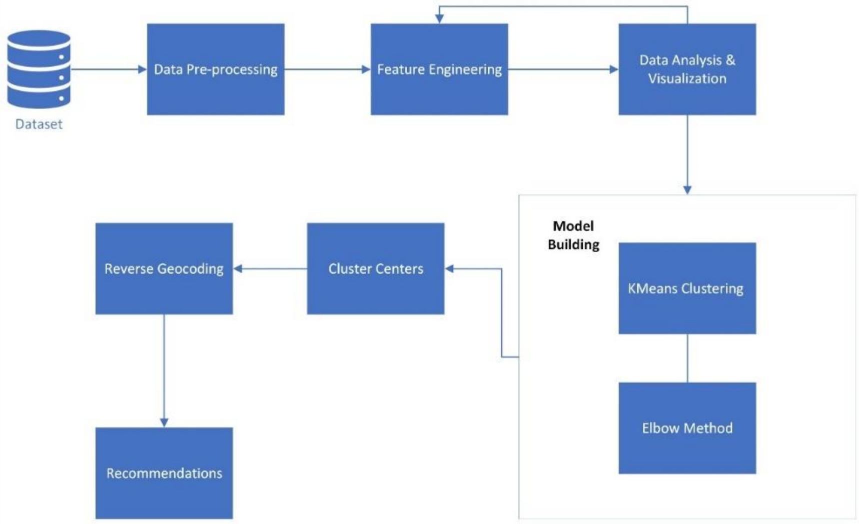
The papers [9] - [11] presented various ways to solve the location routing problems to improve the supply chain cycle.

In our work, we combine these individual ideas and develop a recommender system to improve the supply chain cycle using the given locations for each product in the database.

This paper focuses on developing a recommendation system based on the product orders, utilizing the coordinates on the map to generate clusters for suggesting a location for the selected product.

III. **Proposed Methodology**

Figure 3.1 System Architecture Diagram



Design Overview

R studio an IDE for R programming language, served as the primary tool in this study. As per figure 3.1, the location and product information of all orders was extracted to form the given dataset for further processing and analysis. The extracted information was pre-processed, and a new cleaner dataset was created. Data Analysis was performed on the features to explore and make sure if the dataset is ready or not. Different maps were created to visualize the neighborhood as per the product location groups. The maps also helped visualize the variance in locations for each product in the database.

After this basic analysis, the features were further filtered and analyzed to prepare a cleaner dataset for clustering. This filtered dataset was now ready to be clustered.

A model was built based on the content-based filtering approach by applying K-Means clustering. The desired product for which the recommendation is to be generated is selected here and clusters are formed based on the Euclidean distance between the several latitudes and longitudes from which this product orders were made. The Euclidean distance calculates the length of the line segment between two points in Euclidean space. In this case, the coordinates are the latitudes and longitudes of each location. It is obtained by the following equation:

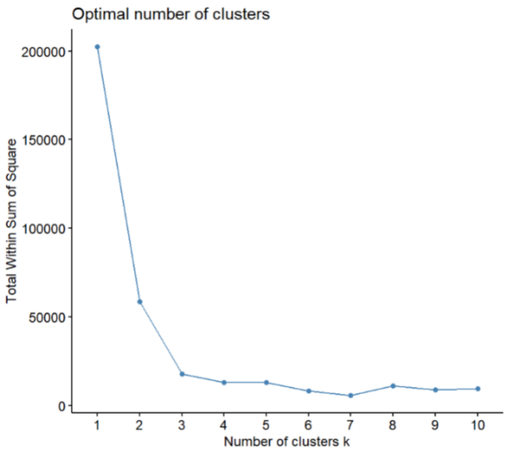
$$D = \sqrt{} [(X_2 - X_1)^2 - (Y_2 - Y_1)^2]$$

Where:
D = Euclidean Distance
(X_1, Y_1) = latitude and longitude of the first point
(X_2, Y_2) = latitude and longitude of the second point

The optimal number of clusters is selected by using the elbow method. The elbow method is a graphical representation of sum of square distance between the points in a cluster and the centroid of the cluster or the Within-Cluster Sum of Square. The clusters are selected by plotting the graph of Within-Cluster Sum of Square on the y-axis for different values of K (number of clusters) on the x-axis. The k value is selected when the reduction in Within-Cluster Sum of Square value becomes insignificant which can be seen an as elbow in the graph.

For example, in figure 3.2 below the selected k value from the graph is 3.

Figure 3.2 Elbow Graph



After selecting the number of clusters needed, the model is processed, and we are presented with the cluster centers in the form of latitudes and longitudes that represent the recommended locations for this product's warehouse.

By reverse geocoding, using the cluster center latitudes and longitudes the exact city and country were extracted for recommending the best locations for the products warehouse.

A. Dataset Description and Exploratory Data Analysis

The selected dataset is about the supply chain information by the company DataCo Global.

The dataset consists of several columns with information on the products, shipping, customer details and sales.

The dataset is a mix of structured and unstructured data. It consists of 180519 rows and 53 columns. There are 118 unique products in this dataset.

This dataset contains a lot of information on many orders for several products. Each order in the dataset has information on the product - name, description, image; the customer ordering the product – customer name, email, address, price, department; the order fulfillment details like – mode of payment, product location, product market, quantity ordered, sales, profit, order status, order date, mode of payment; information on delivery – delivery status, shipping days, sales per customer, late delivery risk etc.

Exploring various features of the dataset:

The figures 3.3 and 3.4 below are representations of the 10 most and least expensive products in the dataset. The "SOLE E35 Elliptical" is the most expensive product and the "Clicgear 8.0 Shoe Brush" is the least expensive product in the dataset.

Fig.3.3 Most expensive Products

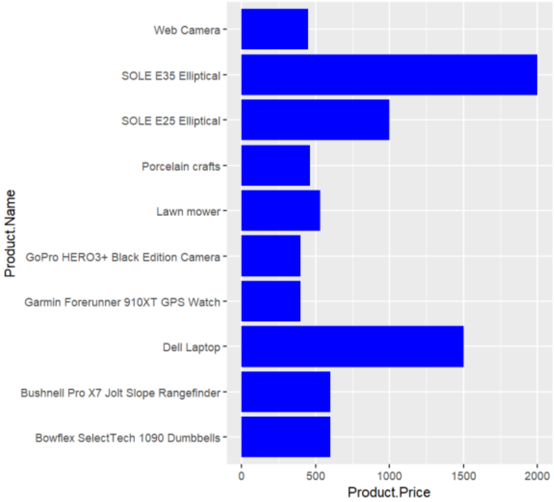


Fig.3.4 Least Expensive Products

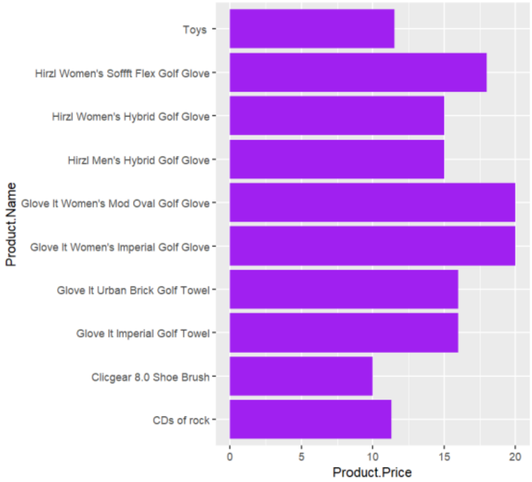
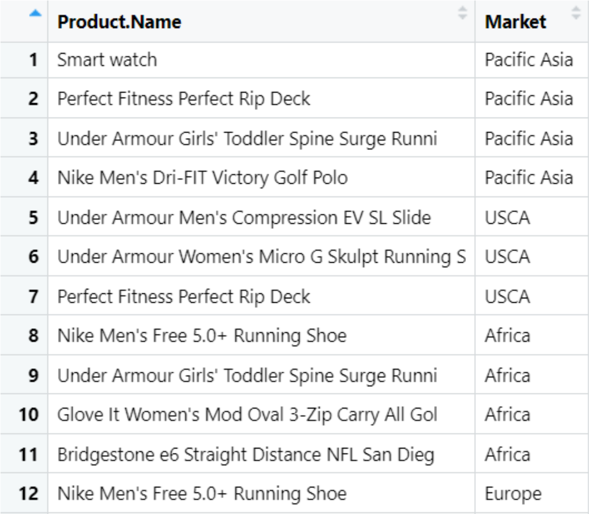


Table 3.1 below is a snippet of the product names vs market table. It shows the product names and the market they are sold and shipped from.

Table 3.1 Displaying distinct Products and Markets



	Product.Name	Market
1	Smart watch	Pacific Asia
2	Perfect Fitness Perfect Rip Deck	Pacific Asia
3	Under Armour Girls' Toddler Spine Surge Runni	Pacific Asia
4	Nike Men's Dri-FIT Victory Golf Polo	Pacific Asia
5	Under Armour Men's Compression EV SL Slide	USCA
6	Under Armour Women's Micro G Skulpt Running S	USCA
7	Perfect Fitness Perfect Rip Deck	USCA
8	Nike Men's Free 5.0+ Running Shoe	Africa
9	Under Armour Girls' Toddler Spine Surge Runni	Africa
10	Glove It Women's Mod Oval 3-Zip Carry All Gol	Africa
11	Bridgestone e6 Straight Distance NFL San Dieg	Africa
12	Nike Men's Free 5.0+ Running Shoe	Europe

Fig.3.5 Days for shipping real vs scheduled

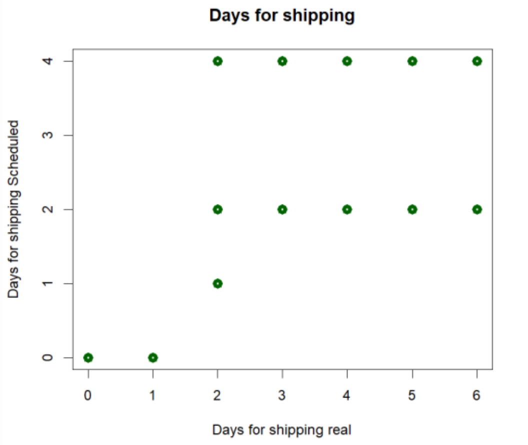
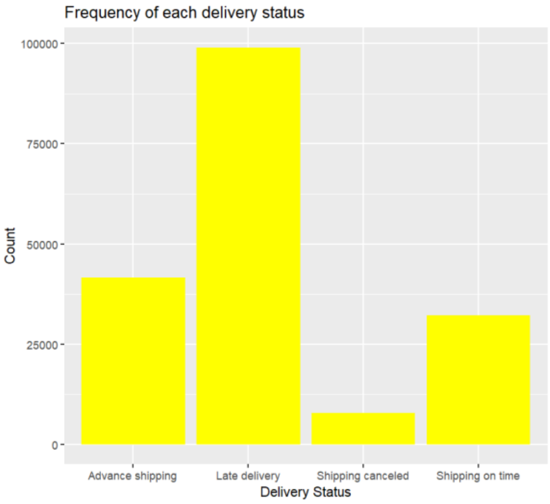


Figure 3.5 above compares the scheduled delivery dates to the real delivery dates. From the graph it can be said that there were several shipping delays

Fig.3.6 Count of each Delivery Status



Looking at figure 3.6, it can be said that most of the orders are delivered late. The number of orders shipped in advance and on time combined is less than the orders that are shipped late. This shows a reason for late delivery.

The columns that would be used in this project are divided into two categories.

1. Location Details – The columns with information about the location and other details about where location details about the order.

   The columns in this category are Latitude, Longitude, Customer City and Market.

2. Product Details – The columns which give information about the product and the order made.

   The columns in this category are Benefit per Order, Product Name, Product Image.

These columns were further filtered upon more exploration and analysis of the dataset.

B. Data Pre-processing

After exploring the dataset, in this step all the unnecessary columns were removed. The column names were changed to make them easier to understand and work on.

The remaining columns or the selected features were - Benefit_per_order, Customer_City, Latitude, Longitude, Market, Product_Image, Product_Name.

Then the rows with all null values were removed and the rows that had any empty cells were omitted from the final dataset.

The pre-processed dataset to be used for analysis now had 180519 rows and 7 columns remaining.

C. Data Analysis

The selected columns were further analyzed for understanding and visualization:

Figures 3.7 and 3.8 below represent the 10 most and least ordered products with their order counts in the dataset. As per the figures, the "Perfect Fitness Perfect Rip Deck" was ordered close to 2500 times and is the most ordered product. The "SOLE E25 Elliptical" and "Bowflex Tech 1090 Dumbbells" were ordered only 10 times and are the least ordered products in the dataset.

Fig 3.7: 10 most ordered products

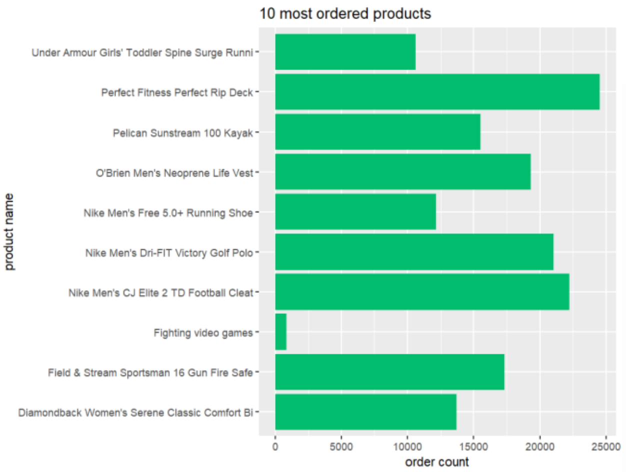
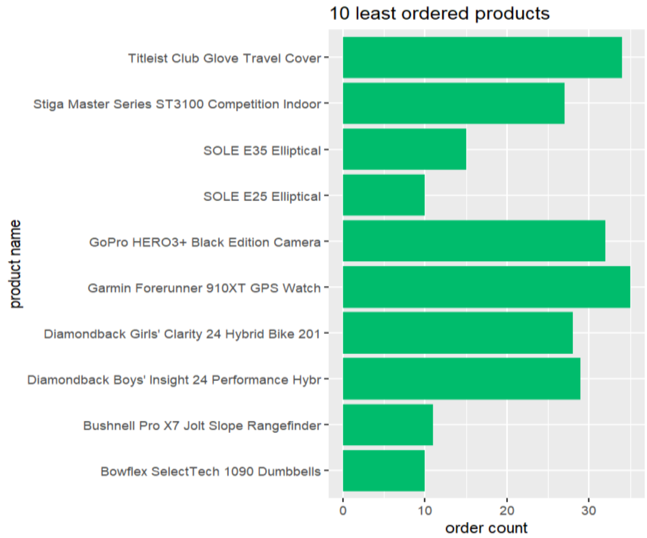


Fig 3.8: 10 least ordered products



Figures 3.9 and 3.10 show the orders with the most and least benefit. As per the figure 3.9, the product order with "Field & Stream Sportsman 16 Gun Fire Safe" has the most benefit per order. In figure 3.10, a lot of the orders have a negative benefit or losses.

Fig 3.9 Highest Benefits

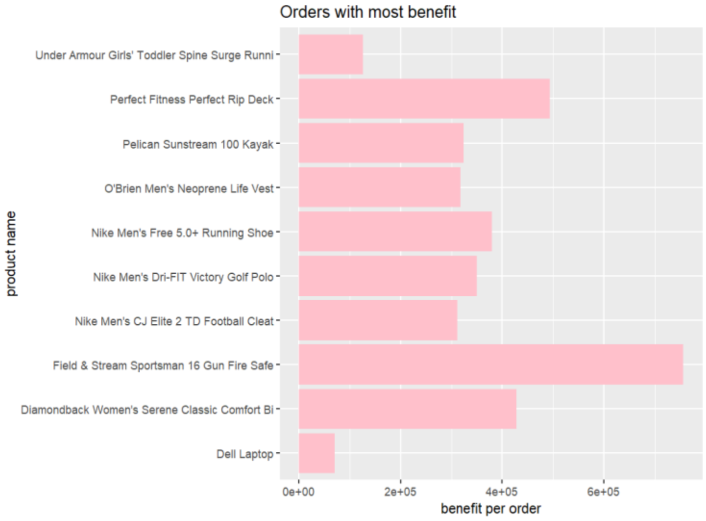


Fig 3.10: Least Benefits

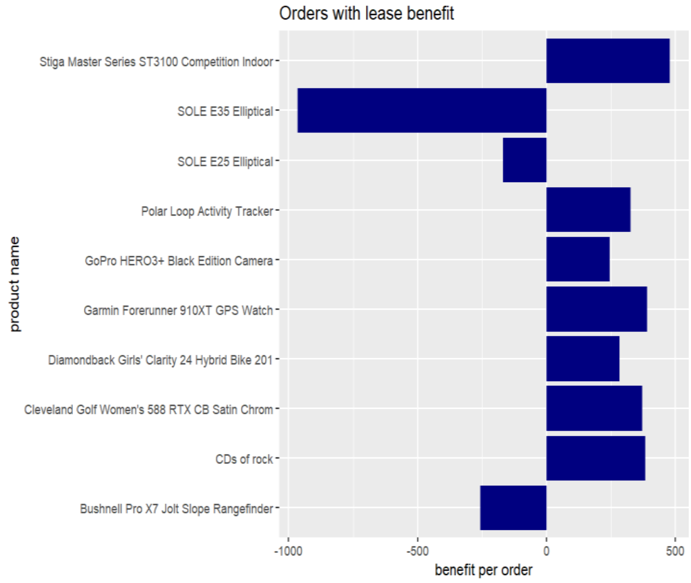
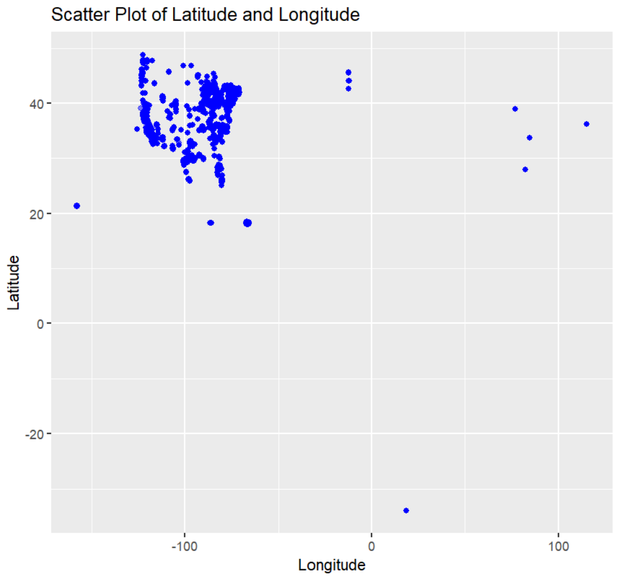


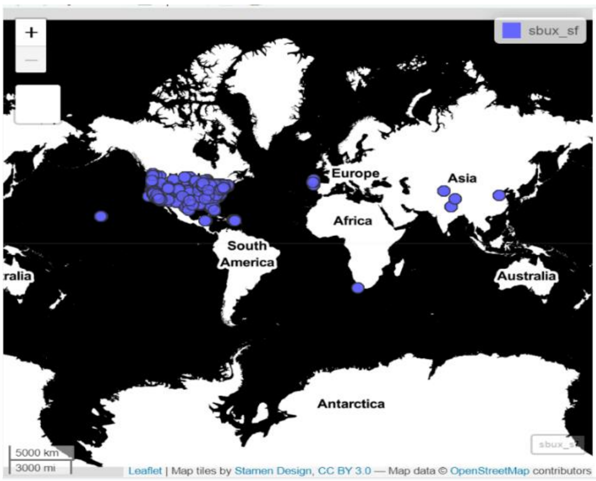
Figure 3.11: Scatter plot of Latitude and Longitude



The figure 3.11 above is a scatter plot of latitudes and longitudes of the orders made. It was created to visualize the variance in the latitudes and longitudes provided.

Figure 3.12 is a map representation of where the orders come from. It was built based on the latitudes and longitudes of the orders already provided in the database. The plot is interactive and hovering over each point shows the order and product details.

Fig 3.12: Map of the product orders and their locations



Leaflet | Map tiles by Stamen Design, CC BY 3.0 — Map data © OpenStreetMap contributors

D. Model Building

To build the model, the features - latitude, longitude and the product names were selected. Every product was assigned an identification number to use in the model. A legend was created to map the product to its ID. A product for which the recommendations were needed was selected and all the orders for that product were filtered.

Using the elbow method, the optimal number of clusters were selected for the product order locations. A cluster neighborhood using the KMeans clustering algorithm was generated based on the selected number of clusters. The cluster centers pointed to the latitudes and longitudes of recommended locations for the warehouses.

The city and country names were extracted using reverse geocoding from the cluster centers. A data frame was created with all the warehouse location recommendations for the selected product.

Statistical Method:

Content based filtering was used here by comparing products, orders, and closer locations to obtain the solution.

The best locations were derived using the unsupervised clustering algorithm KMeans Clustering by sorting the location co-ordinates where the same or similar orders come in most.

The location co-ordinates and product information from the orders were used as inputs and a neighborhood was generated by forming clusters. The number of most optimal clusters(K) was selected as per the chosen product using the elbow method. Using the Euclidean distance of all data points centers were calculated to make the recommendations as per the chosen product.

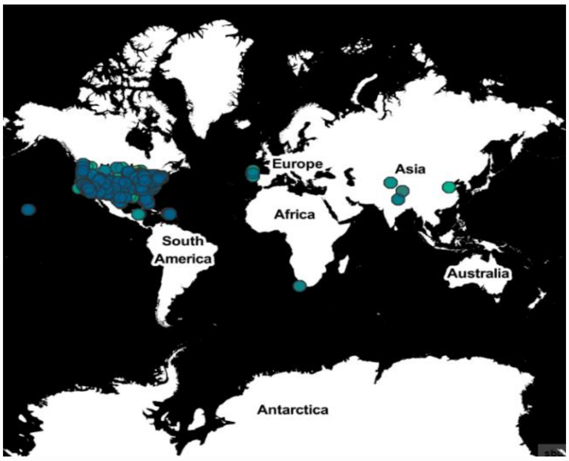
IV. **Results**

The model produced different cluster plots for different products that give us the location recommendations for every product.

The sections below are the steps performed to produce the results and successfully develop the recommendation system that generates suggestions for the selected product. Two examples have also been shown to better explain the work.

Fig 4.1 Map for the order locations grouped by product

Maps



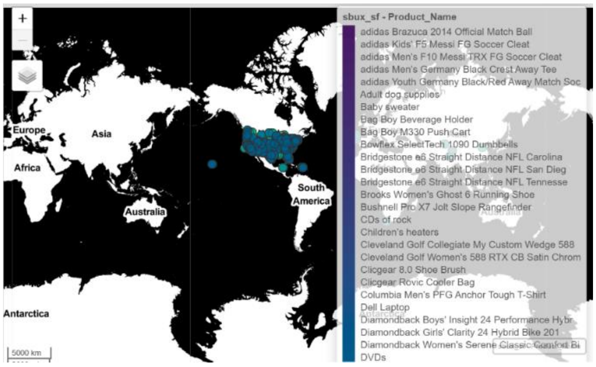
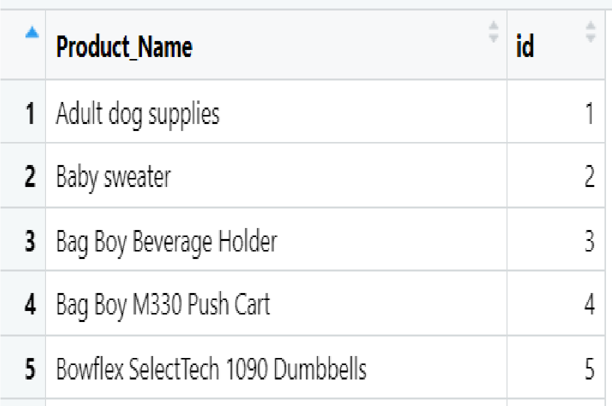

+
Europe
Asia
Africa
Australia
South America
Antarctica
5000 km
sbux_sf - Product_Name
adidas Brazuca 2014 Official Match Ball
adidas Kids' F5 Messi FG Soccer Cleat
adidas Men's F10 Messi TRX FG Soccer Cleat
adidas Men's Germany Black Crest Away Tee
adidas Youth Germany Black/Red Away Match Soc
Adult dog supplies
Baby sweater
Bag Boy Beverage Holder
Bag Boy M330 Push Cart
Bowflex SelectTech 1090 Dumbbells
Bridgestone e6 Straight Distance NFL Carolina
Bridgestone e6 Straight Distance NFL San Dieg
Bridgestone e6 Straight Distance NFL Tennesse
Brooks Women's Ghost 6 Running Shoe
Bushnell Pro X7 Jolt Slope Rangefinder
CDs of rock
Children's heaters
Cleveland Golf Collegiate My Custom Wedge 588
Cleveland Golf Women's 588 RTX CB Satin Chrom
Clicgear 8.0 Shoe Brush
Clicgear Rovic Cooler Bag
Columbia Men's PFG Anchor Tough T-Shirt
Dell Laptop
Diamondback Boys' Insight 24 Performance Hybr
Diamondback Girls' Clarity 24 Hybrid Bike 201
Diamondback Women's Serene Classic Comfort Bi
DVDs

Fig 4.2 Map with Legend for the order locations grouped by product

The Maps in figures 4.1 and 4.2 above show the distribution of various products and their order details and point them out based on their location.

The tables and figures below show the steps used to achieve the recommendations:

Table 4.1 Legend Table that maps products to an ID



▲	Product_Name	id
1	Adult dog supplies	1
2	Baby sweater	2
3	Bag Boy Beverage Holder	3
4	Bag Boy M330 Push Cart	4
5	Bowflex SelectTech 1090 Dumbbells	5

A table with all the distinct products was generated and an ID was assigned to each product as in the table above.

Next a product and its ID were selected and a subset of the dataset with rows containing the selected product only was created.

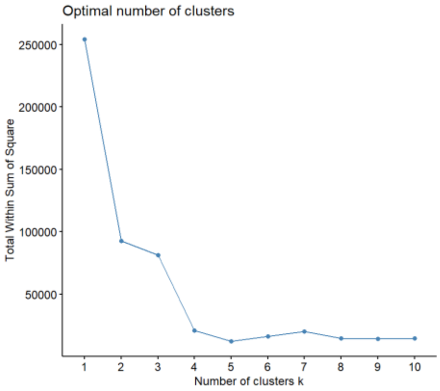
Using the elbow method as in figures 4.3 and 4.5 the optimal number of clusters required were selected.

The clusters were generated based on the selected numbers from the above step as shown in figures 4.4 and 4.6. These clusters each calculate a center which is a pair of latitude and longitude. This latitude and longitude can be interpreted as the recommended location to store the selected product in.

To understand the output, the cluster centers are reverse geocoded to return the county, state, and country that the latitude and longitudes of the outputs were pointing to. This is represented in tables 4.2 and 4.3 below.

Example 1: Recommendation for the product Adult Dog Diapers:

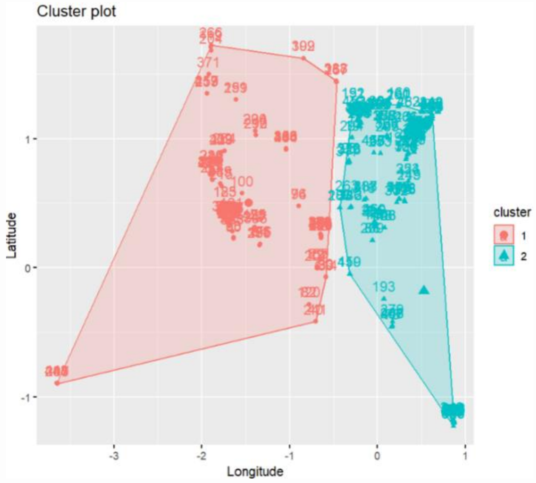
Fig. 4.3 Elbow Curve for Product – Adult Dog Diapers

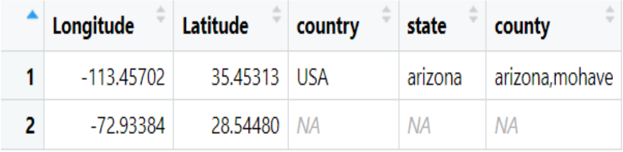


The number of clusters to be generated here was selected to be 2.

Fig 4.4 Clusters for Adult Dog Diapers

Table 4.2 Generated recommendations for warehouse locations for Adult Dog Diapers



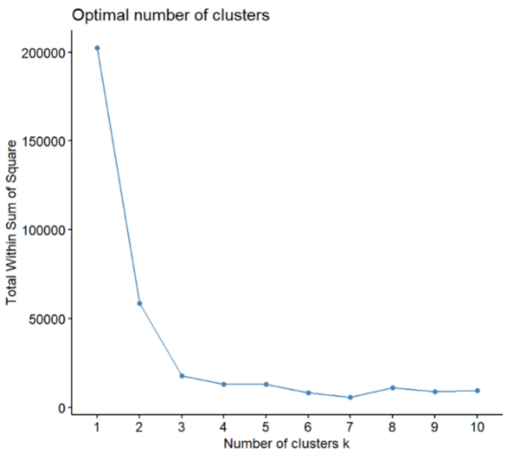


	Longitude	Latitude	country	state	county
1	-113.45702	35.45313	USA	arizona	arizona,mohave
2	-72.93384	28.54480	NA	NA	NA

The product “**Adult Dog Diapers”** is recommended to be stored in Mohave County, Arizona, USA.

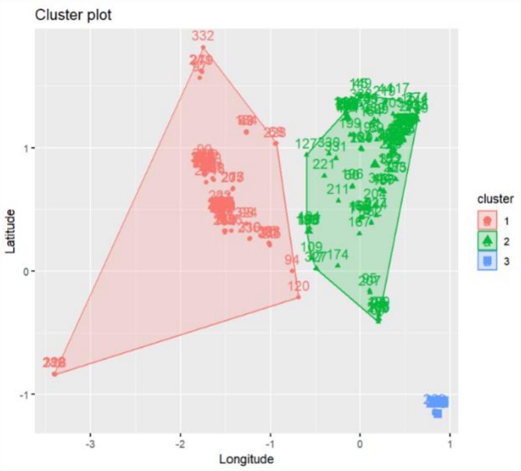
Example 2: Recommendation for the product First Aid Kit:

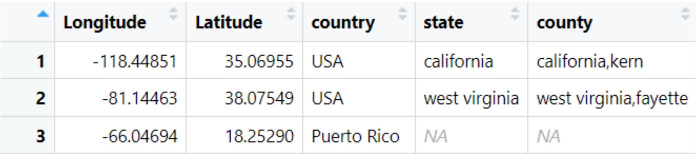
Fig 4.5 Elbow Curve for Product – First Aid Kit



The number of clusters to be generated here was selected to be 3.

Fig 4.6 Clusters for First Aid Kit





	Longitude	Latitude	country	state	county
1	-118.44851	35.06955	USA	california	california,kern
2	-81.14463	38.07549	USA	west virginia	west virginia,fayette
3	-66.04694	18.25290	Puerto Rico	NA	NA

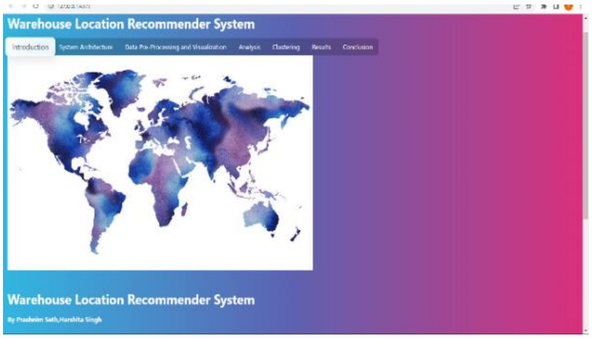
Table 4.3 Generated recommendations for warehouse locations for First Aid Kit

The product “**First Aid Kit**” is recommended to be stored in Kern County, California, USA and Fayette County, West Virginia, USA.

Some rows might have NA in country, state, and county due to the latitude and longitude pointing to a location without a name or some place in the ocean.

Images of shiny app

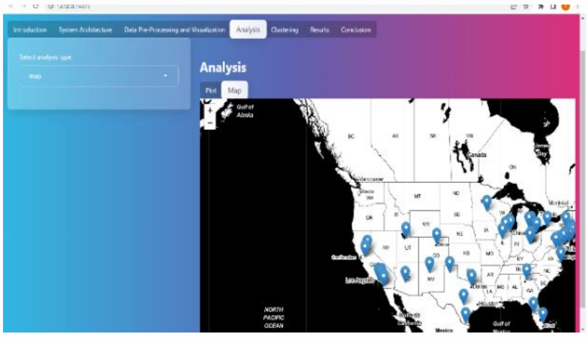
A shiny app was created to better present the idea of this project by showing a demonstration of the recommendation system with the help of a user interface. The images below are screenshots of the created shiny app.



Warehouse Location Recommender System

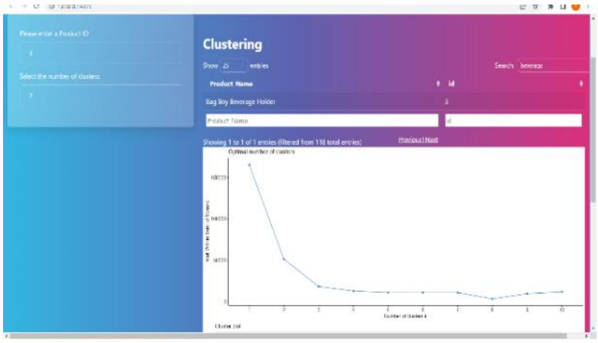
Warehouse Location Recommender System

By Pradeim Seth,Harshita Singh



Select analysis type: map

Analysis



Please enter a Product ID: 1

Select the number of clusters: 3

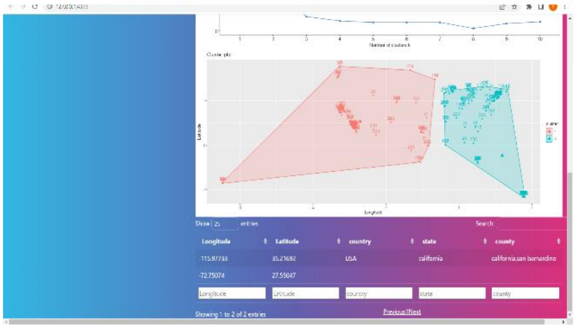
Clustering

Show 25 entries   Search: beverage

Product Name	id
Bag Boy Beverage Holder	3

Showing 1 to 1 of 1 entries (filtered from 118 total entries)

Cluster plot



Show 25 entries Search:

Longitude	Latitude	country	state	county
-115.97733	35.21692	USA	california	california,san bernardino
-72.75074	27.55047			

Showing 1 to 2 of 2 entries Previous 1 Next

V. Conclusion

In this study, a recommendation system was developed that utilizes K-Means clustering, to propose optimal warehouse locations for the selected product based on the product and order locations.

This research builds upon the foundation laid by previous studies conducted in the domain of supply chain and recommender systems. Studies such as Kumar et al. [2], and Trattner and Elsweiler [6] have contributed to valuable insights into recommendation system methodologies in specific domains. Additionally, studies such as Tarapata et al. [7], and Seyedan and Mafakheri [8] have provided further context for this study by exploring data-driven approaches and predictive analytics for supply chain optimization.

The findings of this study revealed that majority of the orders in the dataset come from the country USA, with "Perfect Fitness Perfect Rip Deck" emerging as the most in-demand product. Furthermore, the study identified specific geographical locations where is it most beneficial to store each product in the dataset, such as Mohave County, Arizona, USA for the product "Adult Dog Diapers" and Kern County, California, USA along with Fayette County, West Virginia, USA for the product "First Aid Kit".

In future iterations, this recommendation system could be enhanced by incorporating the feature "benefits per order" to assign weight to each point and then find the cluster centers. Additionally, the automation of the elbow method could streamline the determination of picking the optimal number of clusters, eliminating the need for manual intervention.

Disclosure Statement

No potential conflict of interest was reported by the authors.

VI. **References**

[1]https://www.nvidia.com/en-us/glossary/data-science/recommendation-system/

[2] Kumar, Manoj & Yadav, Dharmendra & Singh, Ankur & Kr, Vijay. (2015). A Movie Recommender System: MOVREC. International Journal of Computer Applications. 124. 7-11. 10.5120/ijca2015904111. https://www.researchgate.net/publication/283042228_A_Movie_Recommender_System_MOVREC

[3] Binu Thomas and Amruth K John 2021 IOP Conf. Ser.: Mater. Sci. Eng. **1085** 012011 https://iopscience.iop.org/article/10.1088/1757-899X/1085/1/012011/pdf

[4] Reema Sikka, Amita Dhankhar and Chaavi Rana. Article: A Survey Paper on E-Learning Recommender System. International Journal of Computer Applications 47(9):27-30, June 2012. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d2ea43f25c1dc9ab7b6a240edcbf8bb0368cbf0c

[5] Philip, Simon & Shola, Peter & Abari, Ovye. (2014). Application of Content-Based Approach in Research Paper Recommendation System for a Digital Library. International Journal of Advanced Computer Science and Applications. 5. 10.14569/IJACSA.2014.051006. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=71afcb79cd9acad12a557ff35f2a0b2fa60ac5f3#page=49

[6] Trattner, Christoph & Elsweiler, David. (2017). Food Recommender Systems: Important Contributions, Challenges and Future Research Directions. https://arxiv.org/abs/1711.02760

[7] Tarapata, Zbigniew & Nowicki, Tadeusz & Antkiewicz, Ryszard & Dudzinski, Jaroslaw & Konrad, Janik. (2020). Data-Driven Machine Learning System for Optimization of Processes Supporting the Distribution of Goods and Services – a case study. Procedia Manufacturing. 44. 60-67.10.1016/j.promfg.2020.02.205. https://www.researchgate.net/publication/341137356_Data-Driven_Machine_Learning_System_for_Optimization_of_Processes_Supporting_the_Distribution_of_Goods_and_Services_-_a_case_study

[8] Seyedan, M., Mafakheri, F. Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *J Big Data* **7**, 53 (2020). https://journalofbigdata.springeropen.com/articles/10.1186/s40537-020-00329-2

[9] Zanjirani Farahani, Reza & Hekmatfar, Masoud & Boloori, Alireza & Nikbakhsh, Ehsan. (2013). Hub location problems: A review of models, classification, solution techniques, and applications. Computers & Industrial Engineering. 64. 1096–1109. 10.1016/j.cie.2013.01.012. https://www.researchgate.net/publication/257177917_Hub_location_problems_A_review_of_models_classification_solution_techniques_and_applications

[10] Barreto, Sérgio & Ferreira, Carlos & Paixão, José & Santos, Beatriz. (2007). Using clustering analysis in a capacitated location-routing problem. European Journal of Operational Research.179.968-977.10.1016/j.ejor.2005.06.074.
https://www.researchgate.net/publication/221951620_Using_clustering_analysis_in_a_capacitated_location-routing_problem

[11] Perl, J. & Daskin, Mark. (1984). A unified warehouse location-routing methodology. Journal of Business Logistics. 5. 92-111. https://www.researchgate.net/publication/284555507_A_unified_warehouse_location-routing_methodology

[12]https://rpubs.com/FelipeMonroy/619723

[13]https://r-spatial.github.io/mapview/reference/mapviewOptions.html

[14] https://data.mendeley.com/datasets/ 8gx2fvg2k6/5