MegaMetrix Mart

A Project report submitted
In partial fulfilment of the requirements for the Degree of Bachelor of
Engineering/Technology

In
Information Technology/Computer Science and Engineering (IoT)
Semester – VII

By

Mr. Mukund Patel (12002080501042)

Academic Year 2022-23 (Odd)

Department of Information Technology
G H Patel College of Engineering & Technology
Bakrol Road, Vallabh Vidyanagar

CERTIFICATE

This is to certify that Seminar work embodied in this report entitled, **MegaMetrix Mart** was carried out by **Mr. Mukund Patel (12002080501042)**, at G H Patel College of Engineering & Technology for partial fulfilment of B.E. degree to be awarded by Charutar Vidya Mandal University (CVMU). This seminar work has been carried out under my supervision and is to the satisfaction of department.

Date: November 2023

Place: G.H. Patel College of Engineering and Technology

Faculty Guide   **Head of Department**

Prof. Mittal Darji   Dr. Nikhil Gondaliya

                                                                                              Professor & Head

Acknowledgement

The completion of any seminar work depends upon cooperation, co-ordination, and combined efforts of several sources of knowledge. I would like to express our/my deepest thanks to **Prof. Mittal Darji**, for his valuable inputs, guidance, encouragement, wholehearted cooperation, and constructive criticism throughout the duration of our project.

I hope that this seminar work report will provide all necessary information required to readers to fulfil their aspiration. Man's quest for knowledge never ends. Theory and practices are essential and complimentary to each other. I would like to express my sincere thanks to **Dr. Nikhil Gondaliya (Head of Department)** for wholehearted support.

Mr. Mukund Patel (12002080501042)

Table of content

List of Figures

List of Abbreviations

CC   Cloud Computing

IT   Information Technology

Abstract

The aim of this project was to develop data analytics solutions to help businesses grow using effective techniques. The project involved the creation of an ultimate data sheet for all marketing and growth purposes. It utilized Python code with various algorithms to create modules for clustering of customers based on their similarity, identifying high-spending customers, tracking repeat customers, making product recommendations, and compiling a list of customers' favourite products. All of this data was then combined into a single Excel sheet that can be used for multiple marketing strategies and can be used to make site changes based on the data.

CHAPTER – 1
Introduction To Project

1.1 PURPOSE OF THE PROJECT

The purpose of this project is to optimize business operations using advanced data analytics algorithms from a sales and marketing perspective. With the increasing amount of data generated by businesses, there is a growing need for tools and techniques to analyze this data and extract insights that can be used to improve business operations. Many companies struggle to make sense of this data and lack the expertise to apply advanced analytics algorithms effectively.

1.2 OBJECTIVE OF THE PROJECT

The objective of this project is to develop a data analytics solution that can optimize business operations from a sales and marketing perspective. To achieve this objective, the project has five specific goals. The first goal is to identify the key sales and marketing metrics that are critical to measuring the performance of a business. The second goal is to collect and analyze dat`a from various sources to generate insights into sales and marketing performance. The third goal is to develop advanced analytics algorithms that can be used to identify patterns and trends in the data. The fourth goal is to use the insights generated by the analytics algorithms to optimize business operations, such as improving marketing campaigns, streamlining sales processes, and enhancing customer engagement. Finally, the fifth goal is to evaluate the effectiveness of the data analytics solution in improving business operations and generating positive ROI for the company. By achieving these objectives, this project aims to help businesses improve their sales and marketing performance by leveraging the power of data analytics.

1.3 SCOPE OF THE PROJECT

The web development industry is a dynamic and rapidly growing sector that is constantly evolving. With the increasing importance of having a strong online presence, businesses are turning to web developers for a range of services, from website design to custom software development. As new technologies and tools emerge, web developers with expertise in areas such as data analytics and marketing are becoming increasingly valuable, making this an exciting and promising industry for those looking to enter it.

CHAPTER – 2
System Analysis

2.1 STUDY OF CURRENT SYSTEM

In the study of the current marketing and analysis domain, we have identified the components and processes of the existing system. We have examined the data analytics tools and techniques being used by businesses to analyse customer behaviour, preferences, and purchase patterns.

Some of the components of the current marketing and analysis domain include customer relationship management (CRM) software, web analytics tools, social media monitoring tools, statistical analysis software, data visualization tools, machine learning algorithms, email marketing platforms, and social media advertising platforms.

These components are used to collect and analyse customer data to understand their behaviour, preferences, and purchase patterns. By using these tools and techniques, businesses can identify trends and patterns in customer behaviour, which can help them make informed decisions about marketing strategies and product recommendations.

The data collected through these tools is analysed using various techniques such as clustering, segmentation, regression analysis, and machine learning algorithms. These techniques are used to identify customer segments, predict customer behaviour, and make recommendations for products and services.

2.2 PROBLEM AND WEAKNESSES OF CURRENT SYSTEM

The current marketing and analysis system faces several challenges that can limit its effectiveness. Some of the major issues include:

Lack of Integration: The current system often consists of multiple disconnected tools and platforms that do not integrate with each other. This makes it difficult to get a holistic view of customer behaviour and preferences. For example, customer data may be collected from multiple sources such as social media, email campaigns, and website analytics, but if these sources are not integrated, it can be challenging to get a complete understanding of customer behaviour and preferences.

Limited Insights: While the current system provides valuable insights into customer behaviour, preferences, and purchase patterns, it is often limited in terms of the depth and scope of analysis. Businesses may not have access to all the data they need, or they may not have the tools and expertise required to conduct more advanced analysis. This can lead to missed opportunities and suboptimal marketing strategies.

Incomplete Customer Profiles: The current system may not provide a complete view of customer behaviour and preferences. For example, businesses may not be able to track customer behaviour across all touchpoints, or they may not be able to capture data on offline interactions. This can lead to incomplete customer profiles and inaccurate insights into customer behaviour.

Lack of Human Expertise: Another challenge that businesses face is a lack of human expertise in data analytics. Analysing customer data requires specialized skills and knowledge, which

may not be available within the organization. This can lead to errors in data analysis, missed insights, and suboptimal marketing strategies.

2.3 PROPOSED SYSTEM

To address the challenges in the marketing and analysis domain, machine learning algorithms can be leveraged for data cleaning and integration, data security, and analysis in the proposed system. These algorithms can automate data cleaning and integration, detect and prevent security breaches, and analyse complex data sets to provide insights that may not be readily apparent to humans. Excel can also be used for data cleaning, integration, and analysis in the proposed system. Moreover, businesses can invest in training their employees in data analytics and machine learning techniques to improve their expertise in the proposed system. By implementing these solutions in the proposed system, businesses can improve their marketing and analysis capabilities and make informed decisions based on the insights gained from their data.

2.4 FEATURES OF PROPOSED SYSTEM

Data Cleaning and Integration: Machine learning algorithms can automate data cleaning and integration processes, reducing errors and inconsistencies in the data.

Data Security: Machine learning algorithms can detect and prevent security breaches, helping businesses protect sensitive data and maintain customer trust.

Analysis: Machine learning algorithms can analyse complex data sets to provide insights that may not be readily apparent to humans, such as clustering customers based on similarities in their behaviour or suggesting products or services that a customer is likely to be interested in based on their past behaviour.

the proposed system also includes a multipurpose Excel sheet that can be used for a variety of marketing and growth purposes. This Excel sheet is created using Python code that incorporates different algorithms for clustering customers based on their behaviour, identifying money-centric customers, tracking customer repetitions and creating recommendation lists based on the top products and customer favourites. The Excel sheet can be customized to suit specific business needs and can be used to create different marketing strategies based on the data available.

CHAPTER – 3
System Design

3.1 SYSTEM DESIGN & METHODOLOGY

The objective of the data analytics project was to help businesses grow by using effective techniques and algorithms. The project was divided into two main parts: creating an ultimate data sheet for all marketing and growth purposes, and building an advanced algorithm system to provide accurate product recommendations.

In the first part of the project, Python programming language was used to create different algorithms that could perform various tasks. The algorithms included clustering of customers based on their similarity, identifying money-centric customers who spend heavily on the site, repetition of customers to understand their visiting patterns, product recommendations based on collaborative filtering algorithm, and identification of favourite products for customers.

All of these algorithms were combined in an Excel sheet to create a powerful tool for businesses. The data gathered through these algorithms allowed businesses to create different marketing strategies and make changes to their site based on customer preferences.

In the second part of the project, an advanced algorithm system was built using data from around 25,000 products. The Python code developed was able to handle grammar and formatting errors, and provide the most appropriate and accurate product recommendations to customers. The system also had the feature of finding products based on the.

The methodology used in this project was highly effective in achieving the project objective of helping businesses grow. By using Python programming language to create different algorithms and combining them in an Excel sheet, businesses were able to gather data that could be used for various marketing purposes. The advanced algorithm system further enhanced the accuracy of product recommendations, leading to increased customer satisfaction and loyalty.

3.2 PYTHON LIBRARIES

Python has become a popular language for data science and machine learning, in large part due to the availability of many libraries that provide a wide range of functionalities. These libraries include tools for data manipulation, visualization, statistical analysis, machine learning algorithms, and deep learning frameworks. Among the many libraries available, some popular ones for data analysis and machine learning include scikit-learn, NumPy, pandas, Matplotlib, Seaborn, mlxtend, Surprise, and fuzzywuzzy.

Scikit-learn is a powerful machine learning library that includes a variety of tools for clustering, classification, regression, and more. NumPy is a fundamental library for scientific computing with Python that provides support for large, multi-dimensional arrays and matrices, as well as a wide range of mathematical functions to operate on these arrays. Pandas, on the other hand, is a library for data manipulation and analysis, and it provides a powerful data structure called a Data Frame that allows you to manipulate and analyse large datasets with ease.

Matplotlib is a popular plotting library that provides a wide range of tools for creating visualizations in Python, including line plots, scatter plots, histograms, and more. Seaborn is a

higher-level interface to Matplotlib that provides a range of built-in visualizations and tools for data exploration. Mlxtend is another library for machine learning and data analysis that includes a wide range of tools for association rule mining, feature selection, and more.

Surprise is a Python scikit for building and analysing recommender systems, and it provides a range of algorithms for collaborative filtering, including matrix factorization and k-NN. Finally, fuzzywuzzy is a library for fuzzy string matching in Python that can be used to match similar strings, correct spelling errors, and more. With these and many other libraries available, Python continues to be a powerful tool for data science and machine learning.

3.3 ALGORITHMS

3.3.1 K-MEANS CLUSTERING ALGORITHM

K-Means Clustering is an unsupervised learning algorithm used to solve clustering problems in machine learning or data science. This algorithm groups an unlabelled dataset into different clusters, where the value of K defines the number of pre-defined clusters that need to be created. The algorithm is iterative and divides the dataset into k different clusters in a way that each dataset belongs to only one group with similar properties. K-Means allows us to cluster data into different groups, making it a convenient way to discover the categories of groups in an unlabelled dataset on its own, without the need for any training.

K-Means is a centroid-based algorithm, where each cluster is associated with a centroid. The primary aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters. The algorithm takes the unlabelled dataset as input, divides it into k-number of clusters, and repeats the process until it finds the best clusters. The value of k should be predetermined in this algorithm.

The k-means clustering algorithm mainly performs two tasks: it determines the best value for K centre points or centroids by an iterative process, and assigns each data point to its closest k-centre. The data points that are near to the particular k-centre create a cluster. Hence each cluster has datapoints with some commonalities and is away from other clusters.

Python implementation of K-Means clustering involves using the scikit-learn library that provides many clustering algorithms, including K-Means. The implementation requires importing the necessary libraries, loading the dataset, and defining the number of clusters to create. After that, the K-Means algorithm is trained on the dataset, and the clusters are visualized using Matplotlib.

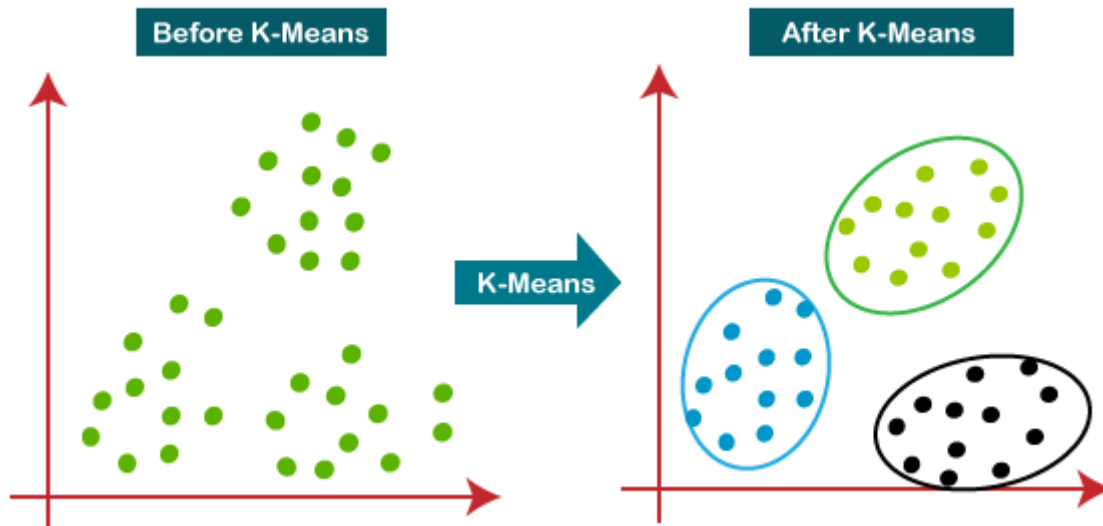The below diagram explains the working of the K-means Clustering Algorithm:

Fig 3.1 K-means Clustering

The K-Means algorithm works in the following steps: first, the number of clusters K is chosen. Then, K random points or centroids are selected. Each data point is then assigned to its closest centroid, forming K clusters. The variance is calculated, and new centroids are placed for each cluster. This process is repeated until no reassignments occur. The resulting model is then ready.

3.3.2 COLLABORATIVE FILTERING

Collaborative filtering is a machine learning algorithm used to make personalized recommendations to users based on their past behaviours and preferences. It works by finding similarities between users or items and using that information to make predictions about which items a user may like or dislike.

Collaborative filtering is a machine learning algorithm used to make personalized recommendations to users based on their past behaviors and preferences. It works by finding similarities between users or items and using that information to make predictions about which items a user may like or dislike.

Here is a simple diagram of how collaborative filtering works:

COLLABORATIVE FILTERING

LIKED BY ALICE AND BOB

SIMILAR USERS

LIKED BY ALICE, RECOMMENDED TO BOB
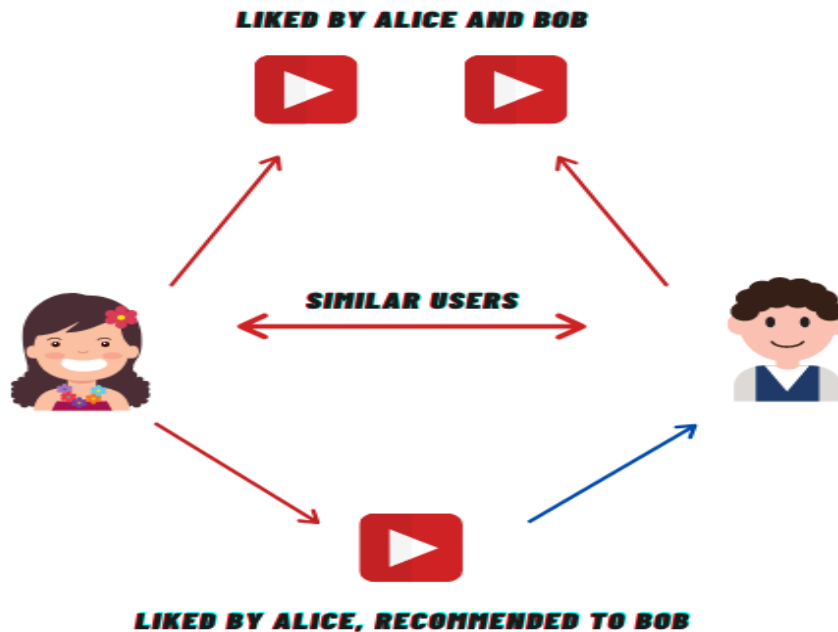
Fig 3.2 Collaborative Filtering

Collaborative filtering is a popular technique used in recommender systems to generate personalized recommendations for users. Here is a simple diagram that outlines the steps involved in collaborative filtering:

The first step in collaborative filtering is data collection, which involves collecting user-item interactions such as purchases, ratings, likes, and so on. Once the data is collected, the next step is to calculate the similarity between users or items based on their interactions with the items. There are different methods to calculate similarity, such as cosine similarity or Pearson correlation.

User-user collaborative filtering involves finding similar users to the target user and recommending items that the similar users have liked but the target user has not yet interacted with. This approach is based on the assumption that users who have similar preferences for items are likely to have similar preferences for other items as well.

Item-item collaborative filtering, on the other hand, involves finding similar items to the target item and recommending those items to users who have interacted with the target item. This approach is based on the assumption that users who have interacted with a particular item are likely to be interested in similar items.

After finding similar users or items, the next step is to generate a predicted rating for the target user and target item based on the ratings of similar users or items. Finally, a list of recommended items for the target user is generated based on the predicted ratings.

3.3.3 FUZZY LOGIC

The 'Fuzzy' word means the things that are not clear or are vague. Sometimes, we cannot decide in real life that the given problem or statement is either true or false. At that time, this concept provides many values between the true and false and gives the flexibility to find the best solution to that problem.

Example of Fuzzy Logic as comparing to Boolean Logic
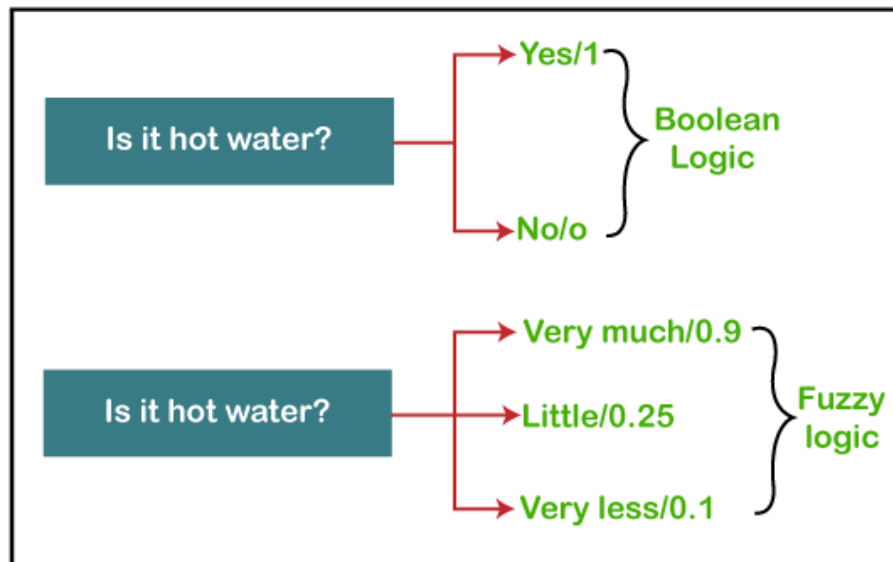
Fig 3.3 Fuzzy Logic

Fuzzy logic contains the multiple logical values and these values are the truth values of a variable or problem between 0 and 1. This concept was introduced by Lofti Zadeh in 1965 based on the Fuzzy Set Theory. This concept provides the possibilities which are not given by computers, but similar to the range of possibilities generated by humans.

In the Boolean system, only two possibilities (0 and 1) exist, where 1 denotes the absolute truth value and 0 denotes the absolute false value. But in the fuzzy system, there are multiple possibilities present between the 0 and 1, which are partially false and true.

The Fuzzy logic can be implemented in systems such as micro-controllers, workstation-based or large network-based systems for achieving the definite output. It can also be implemented in both hardware or software.

To implement fuzzy logic in an advanced search bar, it is important to preprocess the search terms before applying the algorithm. This involves converting all letters to lowercase, removing special characters, and removing stop words. This helps to ensure that the search terms are normalized and can be accurately compared to other strings.

Once the search terms have been pre-processed, a fuzzy matching algorithm can be applied to calculate the similarity between the search terms and other strings. There are many different fuzzy matching algorithms available, each with their own strengths and weaknesses. The Levenshtein distance algorithm and the Jaro-Winkler distance algorithm are two popular options.

After applying the fuzzy matching algorithm, the search results can be ranked based on their similarity to the search terms. The most similar results will be ranked higher than less similar results. This ranking can be further refined by adjusting the weighting of different search parameters, such as fuzzy matching thresholds and search result rankings.

To ensure that the advanced search bar is effective and user-friendly, it is important to test and refine the system. This involves gathering feedback from users and analysing search patterns to identify any areas for improvement. The search bar can be further optimized by adjusting the weighting of different search parameters, such as fuzzy matching thresholds and search result rankings.

CHAPTER – 4
Use Case

4.1 THE ADVANCED ANALYTICS MASTERSHEET

The Advanced Analytics Master Sheet is to provide businesses with the ultimate marketing and analytics tool that can drive growth and success. This Excel file includes six major columns, each with several minor columns that provide a comprehensive view of customer behaviour and preferences. This report aims to provide an overview of the Excel file and explain how businesses can use it to their advantage. By leveraging the power of this tool, businesses can gain valuable insights into their customer base, identify trends and patterns, and make data-driven decisions that can lead to growth and success.

4.1.1 CUSTOMERS DETAILS

	customer_id	email	mobile	whatsapp
4	6599	110HLOUNGE@GMAIL.COM	(508) 904-8250	(508) 904-8250
14	6616	TUSHAR9722@GMAIL.COM	(718) 340-9282	(718) 340-9282
23	6637	360DELINYC@GMAIL.COM	(443) 515-7879	(443) 515-7879
25	6641	MUNAWAR_2042@YAHOO.COM	(347) 613-8864	(347) 613-8864
26	6643	162SMOKESHOP@GMAIL.COM	(647) 247-1473	(647) 247-1473
38	6665	ABDOYAHYA143@GMAIL.COM	(347) 285-4075	(347) 285-4075
43	6680	-	(908) 230-3513	(908) 230-3513
44	6681	nutansls@gmail.com	(908) 230-3513	-
51	6696	SUBJANALI@YAHOO.COM	(917) 463-9338	(917) 463-9338
54	6701	JUNAIDBINMAHFOOZ@GMAIL.COM	(347) 925-4492	(347) 925-4492
60	6718	JAGRUTI444@YAHOO.COM	(603) 880-7655	(603) 880-7655
68	6729	ROCKYPATELRP1234@GMAIL.COM	(347) 476-4700	(347) 476-4700
71	6740	ANDYP95USA@GMAIL.COM	(617) 331-8488	(617) 331-8488
72	6742	zandani199@icloud.com	(347) 264-1849	(347) 264-1849
73	6743	SMPATEL8510@GMAIL.COM	(860) 709-0264	(860) 709-0264
92	6787	AMPMEXPRESS123@GMAIL.COM	(651) 731-4148	(651) 731-4148
97	6794	-	(347) 656-4689	-
104	6823	ANKITDEV2130@GMAIL.COM	(347) 476-4700	(347) 476-4700
114	6837	MOHAMMADAQEEL75@GMAIL.COM	(551) 655-2014	(551) 655-2014
121	6851	PATELPD89@GMAIL.COM	(774) 357-9162	(774) 357-9162
147	6896	PATELSANDY423@GMAIL.COM	(609) 592-6585	(609) 592-6585
151	6904	ASHWIN1095@YAHOO.COM	(603) 858-2248	(603) 858-2248
158	6919	GARVITPATEL.3@GMAIL.COM	(732) 986-5096	(732) 986-5096
173	6946	BILLYSLIQUORS@GMAIL.COM	(646) 578-0016	(646) 578-0016
184	6968	BORDERLINEBEVERAGES@GMAIL.COM	(443) 469-4023	(443) 469-4023
185	6969	JAYPATEL15588@GMAIL.COM	(508) 665-9414	(508) 665-9414
187	6975	-	(917) 600-8079	-
194	6988	NIKUL1000@YAHOO.COM	(603) 769-9875	(603) 769-9875

Sheet1 | Sheet2 | money-centric view | repetition | no_repetition | Cluster 0

Ready   383 of 2685 records found   Accessibility: Investigate

Fig 4.1 Customers Details

The details of our customers can be a valuable resource for personalized marketing and communication. By using the customer's name in email communication, for example, we can create a more personalized and engaging message. Additionally, we can segment our email lists and SMS campaigns based on the email and mobile fields. This allows us to target specific groups of customers with tailored promotions and messages.

One way to use this information effectively is by sending promotions to customers who have not made a purchase in a long time. By creating a segment of customers who have not made a purchase in a certain time period, we can send targeted promotions or discounts to encourage them to make a purchase. This approach can be especially effective when combined with other sorting columns in our customer database, allowing us to easily copy the email or phone number of the targeted customers and send them marketing materials directly.

4.1.2 CLUSTERS

Fig 4.2 Clusters

One way to use the cluster information is to identify customers who are most likely to respond to a specific promotion. By filtering customers based on their spending behaviour, such as high repeat rate and high spend per day, we can create promotions that are targeted towards those customers. This can lead to increased customer engagement and higher sales.

Another way to use the cluster information is to identify trends in customer behavior. For example, if we notice that a particular cluster is showing a decline in repeat rate, we can investigate the cause and take corrective action. This could involve changes to our product or service offerings, pricing strategies, or marketing campaigns.

15

By using cluster analysis to identify new customer segments, we can gain insights into customer behaviour that were previously unknown. For example, we may discover that a particular cluster has a high affinity for a specific product category or prefers a certain type of shopping experience. This information can be used to tailor our offerings to better meet the needs of our customers.

In summary, cluster analysis is a valuable tool for gaining insights into customer behaviour and identifying new customer segments. By filtering customers based on their spending behaviour and analysing cluster characteristics, we can create targeted promotions, identify trends in customer behaviour, and tailor our offerings to better meet the needs of our customers.

Fig 4.3 Individual Clusters

4.1.3 MONEY-CENTRIC VIEW

Fig 4.4 Money-Centric view

In customer relationship management, the money-centric view is a popular approach for identifying high-value customers. By analyzing customer data related to their spending habits, we can identify customers who are likely to be high-value and create targeted campaigns to retain them.

One way to use the money-centric view is to sort customers by their total spent and identify the top 10% of customers. These high-value customers can be targeted with personalized campaigns to encourage them to continue doing business with the company. This approach can be particularly effective for businesses that rely on repeat purchases and customer loyalty.

Another way to use the money-centric view is to analyze the spend per day column. Customers who have a high spend per day are likely to be frequent buyers and are more likely to make a repeat purchase soon. These customers can be targeted with relevant promotions or offers to encourage them to make another purchase.

Finally, we can also use the total days buy column to identify loyal customers. Customers who have made a purchase on a large number of days are likely to be loyal to the company and are valuable assets. These customers can be targeted with loyalty programs, such as rewards or exclusive offers, to further increase their loyalty and encourage them to continue doing business with the company.
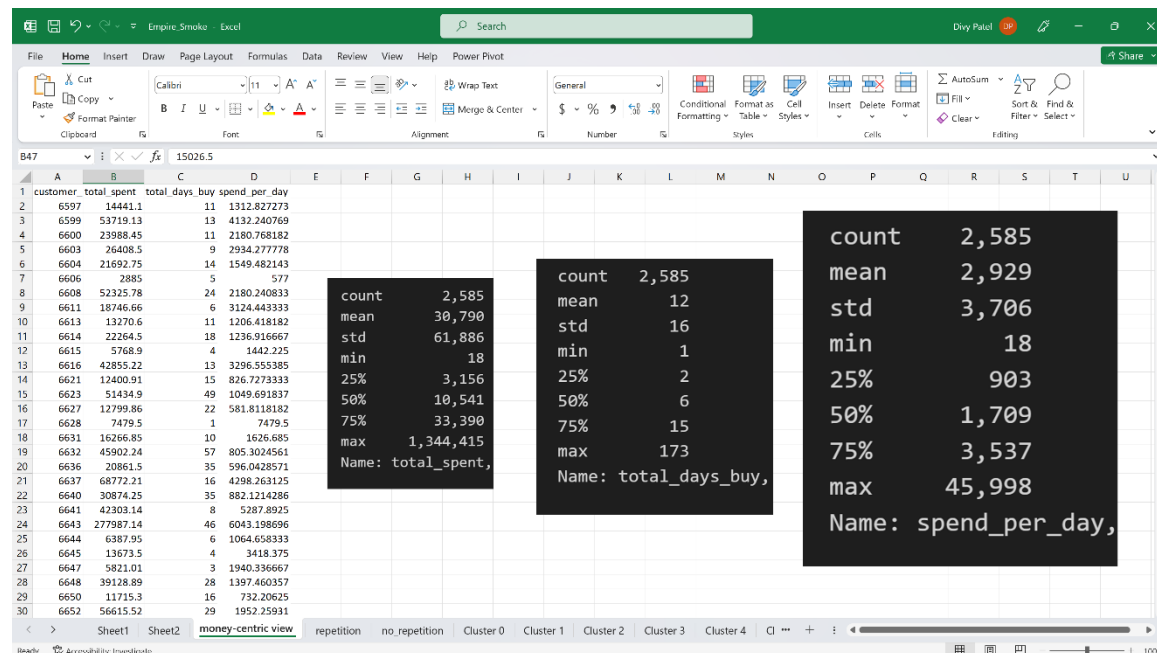

Fig 4.5 Money-Centric Statistics

4.1.4 REPETITIVE CUSTOMERS PATTERNS

Cluster no.	money-centric view			repetition		
cluster	total_spent	total_days_buy	spend_per_day	repeat_day	count	repeat_rate
Cluster 4	53719.13	13	4132.240769	9	19	61.29032258
Cluster 1	42855.22	13	3296.555385	17	11	61.11111111
Cluster 6	68772.21	16	4298.263125	4	15	62.5
Cluster 3	42303.14	8	5287.8925	16	12	66.66666667
Cluster 6	277987.14	46	6043.198696	4	36	52.17391304
Cluster 4	76610.4	20	3830.52	8	25	64.1025641
Cluster 2	45373.06	10	4537.306	26	6	50
Cluster 0	81996.68	22	3727.121818	12	13	54.16666667
Cluster 1	80311.6	23	3491.808696	9	18	52.94117647
Cluster 1	109003.44	16	6812.715	11	16	64
Cluster 6	148845.65	14	10631.83214	11	17	62.96296296
Cluster 1	95053.46	30	3168.448667	6	33	66
Cluster 0	62046	6	10341	37	3	50
Cluster 4	362105.39	111	3262.210721	1	134	69.43005181
Cluster 6	62448.23	13	4803.71	12	17	70.83333333
Cluster 3	30903.31	9	3433.701111	38	5	55.55555556
Cluster 1	175641.57	43	4084.687674	6	7	14.28571429
Cluster 0	52604.35	9	5844.927778	20	4	40
Cluster 1	59856.91	10	5985.691	15	14	66.66666667
Cluster 3	69437.1	7	9919.585714	14	9	75
Cluster 4	163549.66	28	5841.059286	5	33	57.89473684
Cluster 0	45500.91	6	7583.485	14	4	57.14285714
Cluster 1	52782.32	10	5278.232	15	12	63.15789474
Cluster 4	79371.16	21	3779.579048	7	28	68.29268293
Cluster 1	304520.09	79	3854.684684	2	59	52.67857143
Cluster 5	108370.35	10	10837.035	19	9	60
Cluster 6	82011.87	23	3565.733478	9	17	54.83870968
Cluster 1	190994.06	21	9094.955238	7	31	68.88888889

Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cl ...

Fig 4.6 Repetitive Customers

Repetitive customer patterns are an important aspect of customer analysis that can help businesses identify customers who are likely to make repeat purchases and those who are highly engaged with their brand. By analyzing customer purchase data, we can gain insights into customer behavior and use that information to tailor marketing strategies accordingly.

One way to identify customers who are likely to make repeat purchases is by using repetition information. Customers with a high repeat rate are likely to make repeat purchases in the future, and we can target them with relevant promotions. We can use the repeat day column to identify customers who are likely to make a repeat purchase soon. Customers with a short repeat day are likely to make a repeat purchase soon, and we can target them with promotions that encourage repeat purchases.

Another way to identify highly engaged customers is by analyzing the count column. Customers who have made a large number of purchases are likely to be highly engaged with our brand, and we can target them with relevant promotions that reward their loyalty. By identifying and targeting these customers, we can improve customer retention and increase sales.
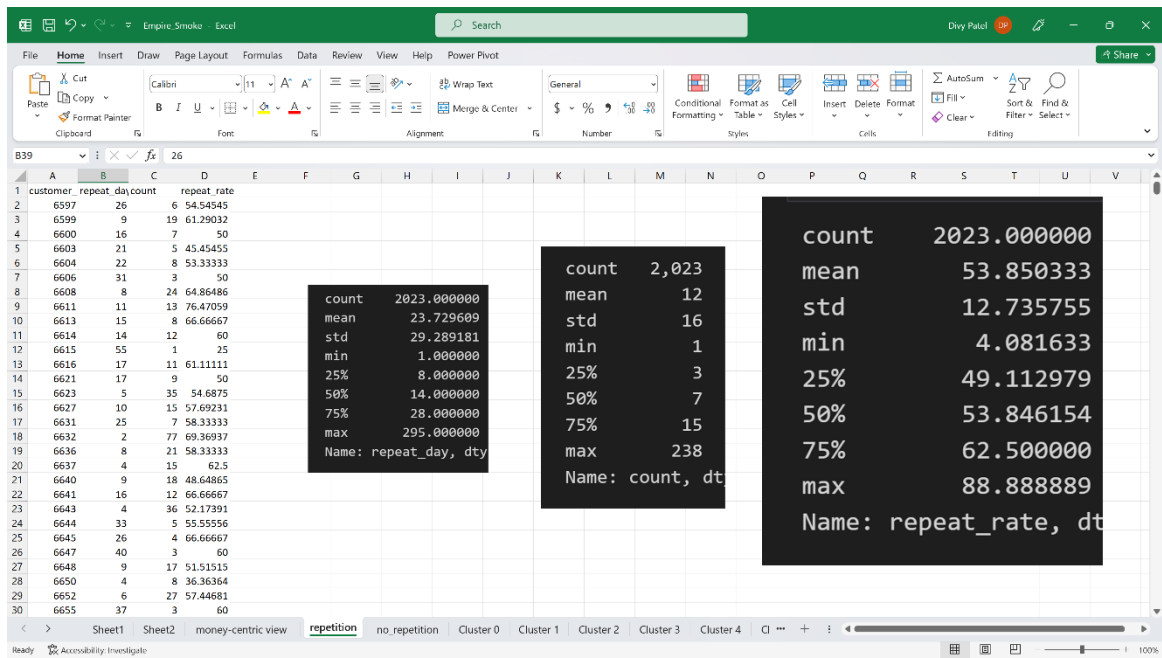
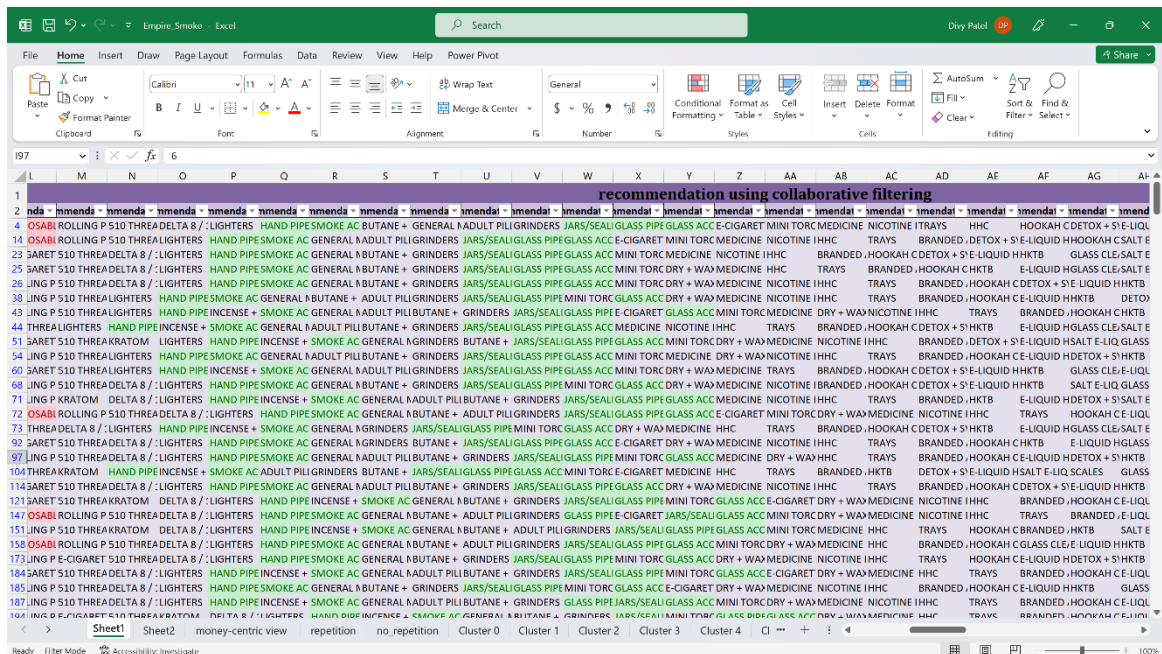Fig 4.7 Repetitive Customers Statistics

4.1.5 RECOMMENDATION

Fig 4.8 Recommendation

Recommendation systems are an essential component of modern e-commerce platforms. These systems provide personalized recommendations to customers based on their past purchase history and browsing behavior. The recommendations can take various forms, such as product suggestions, complementary products, and related products.

One way to use recommendation information is to personalize the product recommendations for each customer. By analyzing a customer's past purchase history and browsing behavior, the system can identify the products that are most likely to interest them. For example, if a customer has previously purchased products that are similar to a particular recommendation, we can highlight that recommendation to them. This approach can help to increase customer engagement and ultimately lead to more sales.

Another way to use recommendation information is to identify new products to promote to customers. By analyzing the recommendations that are frequently made to customers, the system can identify new products that may be of interest to them. For example, if we notice that a particular product is frequently recommended to customers, we can promote that product more aggressively. This approach can help to increase the visibility of new products and attract more customers to the platform.

In addition, recommendation information can be used to identify potential product bundles. By analyzing the recommendations that are frequently made together, the system can identify product bundles that may be attractive to customers. For example, if we notice that certain products are frequently recommended together, we can create a bundle of those products and promote it to customers. This approach can help to increase the value of the purchase for customers and lead to more sales for the platform.
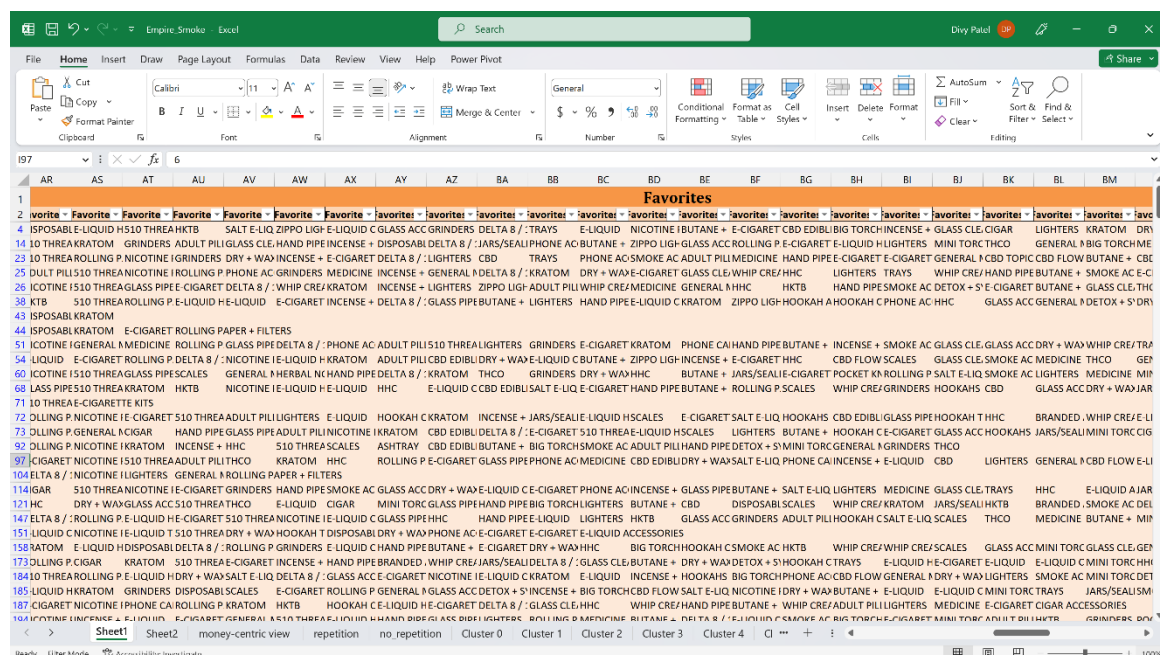
4.1.6 FAVORITE

Fig 4.9 Favorite

Personalizing product recommendations is a powerful way to increase customer engagement and loyalty. By using favorite information, we can gain valuable insights into each customer's preferences and tailor our recommendations accordingly. For instance, if a customer has previously purchased a product that they have marked as a favorite, we can highlight related

products to them. This can help us to showcase products that are most relevant to the customer and increase the chances of them making another purchase.

Moreover, favorite information can be analyzed to identify trends in customer preferences. By analyzing the frequency and types of products that are marked as favorites, we can gain insights into which products are most popular among our customers. This can help us to identify potential gaps in our product offerings and take action to promote products that are in high demand.

In addition, analyzing favorite information can help us to identify potential product opportunities. By tracking which products are frequently marked as favorites, we can gain insights into which products are resonating with customers and potentially develop similar products to meet their needs and preferences.

Conclusion

The Advanced Analytics Master Sheet is a significant step forward in using data analysis to improve sales and marketing efforts. By combining various data points and algorithms, businesses can learn more about how customers behave and what they like. This approach helps make smarter decisions and develop marketing plans that engage customers and boost loyalty, leading to business growth. The project's success in creating complex algorithms like K-Means clustering, collaborative filtering, and fuzzy logic shows its ability to help businesses group customers, give personalized suggestions, and refine marketing strategies.

Putting customer details, cluster analysis, money-focused viewpoints, repetitive customer patterns, suggestions, and favorites all in one Excel sheet makes the project practical and easy to use. This tool makes it simple to access and analyze data, so businesses can make informed choices and run targeted marketing campaigns that connect with different customer groups. The project also stresses the importance of getting better by listening to customer feedback and watching how campaigns perform. This shows its flexibility and dedication to adapting to changing customer needs.

In summary, the Advanced Analytics Master Sheet shows how data analysis can change sales and marketing for the better. With advanced algorithms and comprehensive data skills, this project not only helps businesses understand customers but also improve marketing strategies for successful growth.

References

[1] W3school: https://www.w3schools.com/ MySQL/DEFAULT.asp

[2] Code With Harry: https://www.youtube.com/@CodeWithHarry

[3] AWS: https://aws.amazon.com/

[4] Developer Mozilla: https://developer.mozilla.org/en-US/

[5] Shashank Kalanithi: https://www.youtube.com/@ShashankData