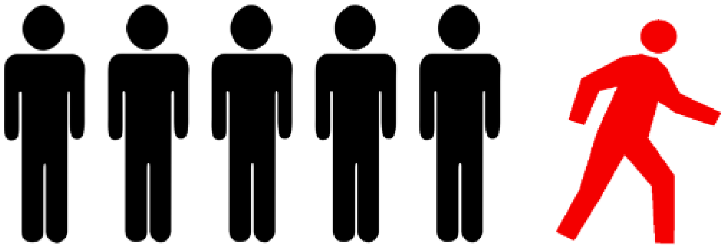
Churn Prediction for FoodCorp



1. Executive Summary:

FoodCorp, a prominent player in the fast-moving consumer goods (FMCG) sector, faces a critical imperative: effectively managing customer churn to sustain its market position. ConsultingCorp's comprehensive analysis offers a strategic roadmap to address this challenge, encompassing churn definition refinement, pioneering churn prediction methodologies, and insightful marketing strategies to fortify FoodCorp's market resilience.ConsultingCorp proposes a nuanced churn definition anchored on a 51-day inactivity threshold, meticulously crafted to discern churn patterns within FoodCorp's customer base. Leveraging transactional data, the analysis reveals that 75.01% of customers exhibit a median lapse between visits falling below this threshold. This refined definition crystallizes actionable intelligence, identifying 11.83% of active customers ripe for targeted interventions, translating to an average of 12.16 daily churn predictions with perfect classification accuracy. Such precision empowers FoodCorp to proactively engage with at-risk customers, nurturing enduring loyalty amidst market dynamics.

The technical endeavour culminates in the development of a sophisticated churn prediction system, leveraging extensive feature engineering predominantly executed in SQL. Anchored in precision-centric methodologies, the predictive model attains a commendable precision score of 93%, signifying its efficacy in identifying churn instances with unprecedented accuracy. Augmented by rigorous preprocessing and robust evaluation protocols, this system furnishes FoodCorp with a resilient framework for strategic decision-making in customer retention endeavours.Central to ConsultingCorp's analysis is the elucidation of nuanced marketing insights, differentiating churned from non-churned segments with granularity. Through immersive exploratory data analysis and compelling visualizations, the insights transcend observation, facilitating personalized marketing campaigns and engagement initiatives. Moreover, the creation of detailed customer personas guides the development of tailored communication strategies and bespoke loyalty programs, fostering enduring customer relationships amidst dynamic market dynamics.

In conclusion, ConsultingCorp's analysis serves as a guiding beacon in FoodCorp's pursuit of churn management excellence. Armed with actionable insights and advanced technical capabilities, FoodCorp is poised to navigate the complexities of the FMCG landscape with precision, fostering enduring customer loyalty and driving sustainable growth.

2.Current levels of churn:

The analysis of current churn levels within FoodCorp's customer base reveals crucial insights into retention dynamics and strategy effectiveness. Churn, defined as a significant deviation from regular purchasing behaviour, presents a formidable challenge for FoodCorp. While conventional churn definitions often relate to customers switching between companies, nuances in the fast-moving consumer goods (FMCG) sector emphasize disruptions in purchasing patterns rather than outright defection. ConsultingCorp proposes a formal churn definition, with a 51-day inactivity period denoted as "β" days, to categorize churn as an absence exceeding this threshold. This selection is strategically justified, balancing the need to identify meaningful churn events with practical resource allocation considerations. The recommendation is substantiated by compelling evidence from transactional data analysis, indicating that approximately 75.01% of customers exhibit fewer median days between visits than the chosen threshold. With this definition, FoodCorp can anticipate targeting approximately 11.83% of active customers for interventions, translating to an average of 12.16 predictions per day with a perfect classifier. This approach enhances the effectiveness of retention strategies, enabling FoodCorp to proactively engage with at-risk customers and cultivate enduring loyalty, positioning the company for sustained growth and competitiveness in the FMCG market landscape.

Furthermore, the churn definition of a 51-day inactivity period, as recommended by ConsultingCorp, will be utilized alongside the results provided by ConsultingCorp for further analysis and decision-making within FoodCorp. This definition serves as a critical parameter in understanding and addressing churn dynamics within the customer base, guiding strategic interventions and resource allocation efforts.

3. Technical Insights into the Churn Prediction System:

3.1 Features:

In the meticulous construction of a comprehensive dataset for customer churn analysis, a strategic and systematic approach was adopted. The process began with a thorough exploration of a diverse set of features, meticulously selected based on their relevance to churn prediction and their potential to provide meaningful insights into customer purchasing behaviour. The dataset was filtered with only active customers and after getting the active customers data, started with the feature engineering. Temporal features were engineered using a carefully devised windowing strategy, enabling the capture of trends and patterns over time.This endeavour commenced with the compilation of summary tables, amalgamating transactional data from various sources to form a unified dataset encompassing all pertinent customer information. Over 20 features were thoughtfully crafted, spanning both temporal and non-temporal dimensions of customer behaviour. Temporal aspects of customer behaviour were designed to quantify customer engagement frequency, while non-temporal attributes offered insights into product category expenditures, features are engineered for a preceding window from the reference date, enriching the dataset with nuanced insights.

To optimize data representation and address multicollinearity concerns, sophisticated techniques including correlation analysis were employed and selected relevant features while discarding the irrelevant features .To reduce high correlation between spend and quantity, the aggregation of spend and quantity features into average spend per quantity for preceding windows, ensuring the retention of essential information while reducing dataset dimensions.Furthermore, employing a diverse array of feature importance methodologies, including VarianceThreshold ,GenericUnivariateSelect etc, facilitated the identification of key predictors crucial for predictive modeling. Ultimately, the adoption of the Cross-Validated Permutation Importance Analysis with Random Forest Classifier emerged as the preferred approach due to its superior predictive performance with score 83.29. This rigorous selection process yielded a succinct list of the top 10 features, encapsulating the most influential factors in churn prediction, refer below figure 1.

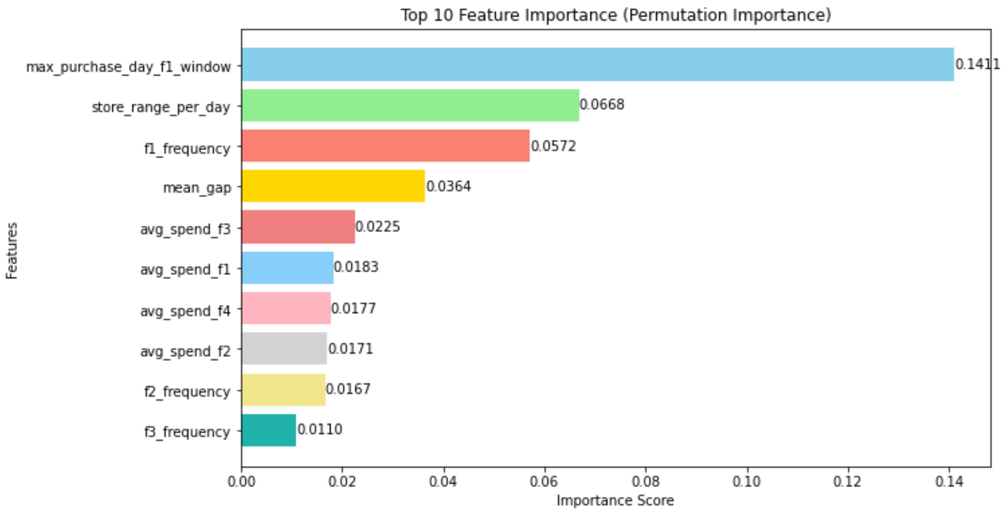


Figure.1: Top 10 Features with Importance Score

Central to this effort was the implementation of a well-defined temporal windowing strategy, essential for accurately delineating input and output features within specified timeframes. With careful consideration of parameters such as output window size and tumbling window size, each window effectively encapsulated pertinent customer behaviour, ensuring precise temporal alignment of features. In essence, this methodical and strategic approach to feature engineering and selection lays a robust foundation for customer churn analysis. By combining meticulous data curation, sophisticated feature engineering techniques, and a comprehensive temporal windowing strategy, this process enables the extraction of actionable insights to inform strategic decision-making and enhance customer-centric initiatives.

3.1.1 Temporal Feature Engineering Strategy:

The temporal feature engineering strategy leverages a structured approach to delineate temporal windows and derive informative features from customer transaction data. This strategy involved defining distinct temporal windows and deriving insightful features from transactional data. Specifically, the `output_window_size` and `tumbling_window_size` parameters were set to 51 days to capture relevant customer behaviours leading up to the churn event. In this strategy, the reference day, denoted as `ref_day`, was calculated based on the maximum purchase date (`now`) minus the output window size (51 days). This reference day served as the anchor point for analyzing customer behaviors within the defined temporal windows.

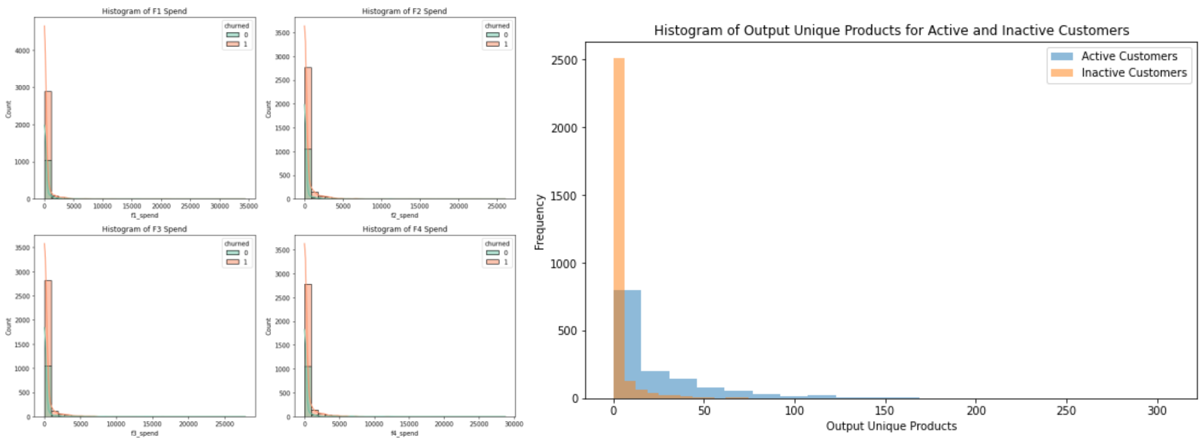


Figure.2:Features Distribution for churner and non- churner as well active and inactive customers

The SQL query executed to extract pertinent customer data utilized these parameters to filter transactions effectively. Within the SQL query, various temporal features were computed, including mean and standard deviation of purchase intervals (`mean_gap` and `std_gap`), range of stores visited per day (`store_range_per_day`), and frequency of transactions within different temporal windows.

For instance, the input features (`f1_spend`, `f2_spend`, `f3_spend`, `f4_spend`) were calculated based on the customer's spending behavior within tumbling windows preceding the reference day. On the other hand, the output feature (`output_feature`) represented the cumulative purchase value within the output window, reflecting customer activities leading up to churn. Additionally, metrics such as total stores visited, total unique products purchased, and purchase frequency within each temporal window were computed to capture diverse aspects of customer behaviour over time, refer figure.2. This comprehensive windowing strategy ensured that temporal features accurately captured customer behaviours preceding churn, providing valuable insights for predictive modeling and targeted retention efforts. By leveraging temporal patterns embedded within transactional data, businesses can proactively identify and mitigate churn risks, fostering long-term customer loyalty and profitability,refer figure.3.

```
Observation Window:
|---------------------------------|----------------------------------|-------------|
|         557               |             608 (Now)          |  609 (Now + 1 day)

Output Window (51 days):
|---------------------------------|----------------------------------|-------------|
|         557               |        557 (Reference Date)     |  608 (Now)

Tumbling Window (51 days):
|---------------------------------|----------------------------------|-------------|

|         506               |          557 (Reference Date)   |  608 (Now)
```

Figure.3: Windowing Strategy used for temporal feature construction

- Observation Window: This is the period during which customer behaviour is observed and analyzed. In this diagram, it is represented by the left section labeled "Observation Window." The observation window encompasses the entire duration of the analysis, during which data is collected and processed.

- Output Window: The output window represents the timeframe leading up to the churn event, where the customer's behaviour is examined to predict churn. In the diagram, it is depicted by the right section labeled "Output Window." The duration of the output window is determined by the variable `output_window_size`, which in this case is set to 51. This means that the output window spans 51 time units leading up to the reference day.

- Tumbling Window: The tumbling window refers to a fixed-size window that moves forward in time, capturing customer transactions and behaviour within each window. It is represented in the lower part of the diagram labeled "Tumbling Window." The size of the tumbling window is determined by the variable `tumbling_window_size`, which is also set to 51 in this example. The tumbling window moves forward in time, capturing customer data within each window as it progresses.

- Reference Day: The reference day, denoted by the variable `reference_day`, is calculated based on the current time (`now`) minus the output window size. In this example, `now` is set to 608, and the output window size is 51. Therefore, the reference day is 608 - 51 = 557. The reference day serves as the anchor point for analyzing customer behaviors within the defined temporal windows.

This diagram illustrates how the observation window, output window, and tumbling window are used in the churn prediction system. The observation window collects data over the entire analysis period, while the output window focuses on the timeframe leading up to the churn event for predicting churn. The tumbling window moves forward in time, capturing customer data within fixed-size windows for analysis.

3.2 Prediction Approach:

The chosen prediction approach for customer churn analysis yielded promising results, as evidenced by the numerical values obtained during model evaluation. With a baseline accuracy of 75.35%, the predictive model achieved a notable precision of 93%, indicating its ability to accurately identify true churn instances among the predicted churn observations. Despite a slightly lower overall accuracy of 68.5% compared to the baseline, the model's precision score underscores its proficiency in identifying actionable churn cases with minimal false positives.

The decision to select this particular model was justified by its superior precision performance, which holds paramount importance in customer churn analysis. A high precision score implies a low false positive rate, meaning that the model effectively identifies customers who are likely to churn while minimizing the risk of incorrectly flagging loyal customers as churners. This precision-focused approach aligns with business objectives, as it enables targeted intervention strategies to retain at-risk customers and optimize resource allocation.

Moreover, the selected model underwent rigorous preprocessing and tuning steps to optimize its performance. Feature engineering efforts focused on capturing nuanced customer behaviour through a diverse set of features, including temporal and non-temporal attributes. Standardization of features using MinMaxScaler within a pipeline framework ensured uniformity in feature scales, facilitating model convergence and enhancing predictive accuracy.

Model selection and tuning were conducted methodically, leveraging GridSearchCV to identify the optimal hyperparameters for each algorithm. The chosen model, LinearSVC, demonstrated superior precision performance, making it well-suited for the task of customer churn prediction. Its ability to handle high-dimensional data and effectively separate churners from non-churners contributed to its selection as the preferred predictive algorithm.

In conclusion, the adoption of a precision-focused prediction approach, coupled with meticulous preprocessing and tuning steps, culminated in the development of a robust model for customer churn analysis. The model's emphasis on precision ensures accurate identification of churn instances, empowering businesses to implement targeted retention strategies and optimize customer relationship management efforts effectively. The utilization of essential libraries like scikit-learn and the encapsulation of models within Pipeline objects streamlined preprocessing, while systematic hyperparameter tuning using GridSearchCV ensured optimal model performance. This comprehensive approach facilitated the development of a predictive model capable of delivering actionable insights to drive strategic decision-making in customer retention endeavours.

3.3 Evaluation Method:

The evaluation method employed for assessing the performance of the predictive model in customer churn analysis embraced a multifaceted approach, encompassing precision, accuracy, and baseline comparison. Precision, denoting the ratio of true positive predictions to all positive predictions, emerged as a pivotal metric, crucial for gauging the model's aptitude in correctly identifying churn instances while mitigating false positives. This metric assumed paramount importance given its direct relevance to targeted intervention strategies and judicious resource allocation in customer retention endeavours. Additionally, accuracy, representing the proportion of correctly classified instances among the total number of instances, offered a holistic assessment of the model's overall performance. The baseline accuracy, functioning as a benchmark for comparison, provided valuable insights into the model's efficacy with a simplistic predictive strategy. Collectively, these evaluation metrics facilitated a comprehensive understanding of the model's predictive prowess and its alignment with the strategic imperatives of FoodCorp.

Following meticulous setup of frameworks for diverse machine learning models and extensive experimentation and comparison, the optimal model for customer churn prediction surfaced through rigorous evaluation. The process entailed delineating pipelines for Support Vector Classifier (SVC) and Random Forest Classifier (RFC), specifying hyperparameter grids, and employing GridSearchCV to discern the most effective combination of parameters. Despite considering alternatives such as XGBoost and AdaBoost, the Linear SVC model with tailored hyperparameters consistently outshone others, yielding the highest precision score on the test data. With an impressive precision score of approximately 93%, the selected model showcased exceptional accuracy in classifying positive cases (churn), underlining its efficacy in addressing FoodCorp's churn management concerns. This decision reflects a judicious balance between precision, accuracy, and computational efficiency, solidifying the Linear SVC model's pivotal role as the linchpin of FoodCorp's customer churn prediction framework.

Furthermore, the evaluation process incorporated numerical values obtained through rigorous model testing and validation, furnishing tangible evidence of the model's performance. Quantitative metrics provided substantive insights into the model's effectiveness in delineating churn instances from non-churners. Leveraging holdout testing with distinct training, validating, and testing sets, alongside Cross-Validation Score, facilitated robust evaluation. Additionally, GridSearchCV enabled systematic exploration of hyperparameter spaces, ensuring optimal model configuration. This comprehensive evaluation methodology engendered a nuanced understanding of the model's predictive capabilities and its potential ramifications on strategic decision-making in customer retention initiatives.

3.4 Summary:

The customer churn analysis aimed to develop a robust predictive model to identify and retain at-risk customers, thereby optimizing customer relationship management efforts. The project encompassed several key phases, including data preprocessing, feature engineering, model selection, and evaluation.In the data preprocessing phase, meticulous attention was dedicated to cleansing and transforming raw transactional data into a structured format suitable for analysis. Essential steps included data cleaning, handling missing values, and standardizing feature scales using MinMaxScaler within a pipeline framework.

Feature engineering played a crucial role in capturing nuanced customer behavior, with over 20 features meticulously crafted to encapsulate both temporal and non-temporal facets of customer purchasing behaviour. This process involved the creation of various temporal features to capture customer engagement frequency and non-temporal attributes to provide insights. The model selection process involved the exploration of diverse machine learning algorithms, including LinearSVC and RandomForestClassifier, encapsulated within Pipeline objects to streamline preprocessing steps. Hyperparameters for each model were systematically tuned using GridSearchCV to identify the optimal configuration, with a precision-focused approach guiding the selection of the best-performing model.Evaluation of the predictive model's performance was conducted using precision, accuracy, and baseline comparison metrics. The model demonstrated a notable precision score of 93%, indicating its proficiency in accurately identifying churn instances with minimal false positives. Despite a slightly lower overall accuracy of 68.5% compared to the baseline accuracy of 75.35%, the model's precision underscored its effectiveness in targeted intervention strategies for customer retention.

Overall, the project's success stemmed from a combination of meticulous data preprocessing, comprehensive feature engineering, methodical model selection and tuning, and thorough evaluation. The resulting predictive model provides actionable insights to inform strategic decision-making in customer retention efforts, empowering businesses to optimize resource allocation and enhance customer relationship management practices.

4.Insight Report:

4.1 Insights for Marketing:

Understanding the distinction between churn and non-churn customers is crucial for crafting effective marketing strategies that drive business growth and foster customer loyalty. The analysis conducted has revealed significant insights into customer behaviour, delineating key differences between these two segments. By segmenting customers based on their spending patterns, distinct groups of active and inactive customers were identified. Notably, a majority of customers (67.2%) fall into the inactive category, underscoring the importance of re-engagement strategies to revitalize dormant relationships. Furthermore, the rigorous churn determination process employed, which factors in the timing of customers' last purchases relative to observation windows, has enabled accurate classification of churned and non-churned customers. This insight serves as a foundational pillar for devising targeted retention efforts and personalized marketing campaigns,refer figure.4.

Visualizations such as histograms, scatter plots, and box plots provide compelling insights into the purchasing behaviour of active and inactive customers, as well as churned and non-churned segments, facilitating a deeper understanding of customer dynamics and informing targeted marketing strategies(refer below digram). Subsequently, attention is focused on active customers, as they represent a critical segment for further examination. This meticulous methodology ensures the accurate identification of churned customers, providing marketers with essential insights to develop effective churn prediction strategies and enhance customer retention initiatives.

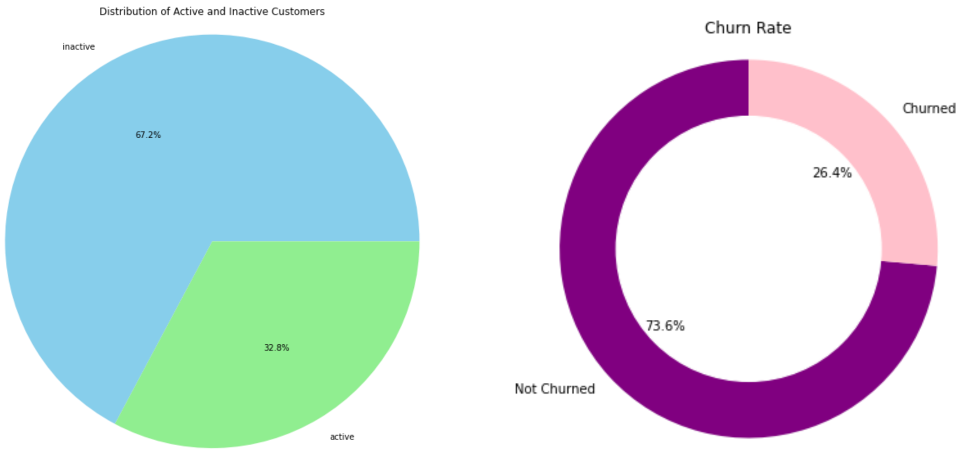


Figure.4: Active and Inactive customers and Churn rate

The analysis also sheds light on the nuanced dynamics of customer behaviour, including mean gap-related information and the number of unique products purchased by churned and non-churned customers. Leveraging exploratory data analysis techniques, such as visualization and correlation analysis, actionable insights into the drivers of churn and retention were uncovered. Armed with these insights, marketing teams can tailor communication strategies, loyalty programs, and promotional offers to cater to the unique needs and preferences of both churn and non-churn customers which are basically the active customers. Additionally, the pen portraits of churning and non-churned customers provide a deeper understanding of their respective purchase frequencies and mean gap patterns, offering valuable guidance for targeted engagement initiatives. Equipped with a profound understanding of customer spending patterns and the likelihood of churn, businesses are empowered to craft targeted marketing strategies tailored to meet the unique needs and preferences of their clientele. Delving into intricate details such as purchase frequency, product preferences, and engagement levels, companies can orchestrate personalized communication initiatives aimed at cultivating stronger bonds with their customer base. Introducing bespoke loyalty programs and incentives not only acknowledges the loyalty of existing customers but also incentivizes continued patronage, fostering a sense of belonging and affinity towards the brand.

Moreover, deploying strategic promotional offers, meticulously designed to resonate with individual customer segments, serves as a potent tool to captivate their interest and nurture enduring brand loyalty. Whether engaging with active customers or those steadfastly loyal and immune to churn, a targeted approach to marketing ensures that every interaction reinforces the brand's value proposition and strengthens its position in the marketplace. Implementing proactive retention initiatives based on churn predictions can significantly reduce churn rates and enhance customer lifetime value. By offering tailored incentives, exceptional customer service, and proactive outreach, businesses can nurture long-term relationships and drive sustainable growth. Leveraging predictive analytics and customer segmentation, businesses can proactively address churn risk, maximize customer lifetime value, and drive long-term profitability.

4.1 Technical Report:

The insights detailed in the marketing summary are supported by a rigorous technical analysis conducted on customer data. This section provides a detailed overview of the methodologies employed, data preprocessing steps, and analytical techniques used to derive actionable insights.

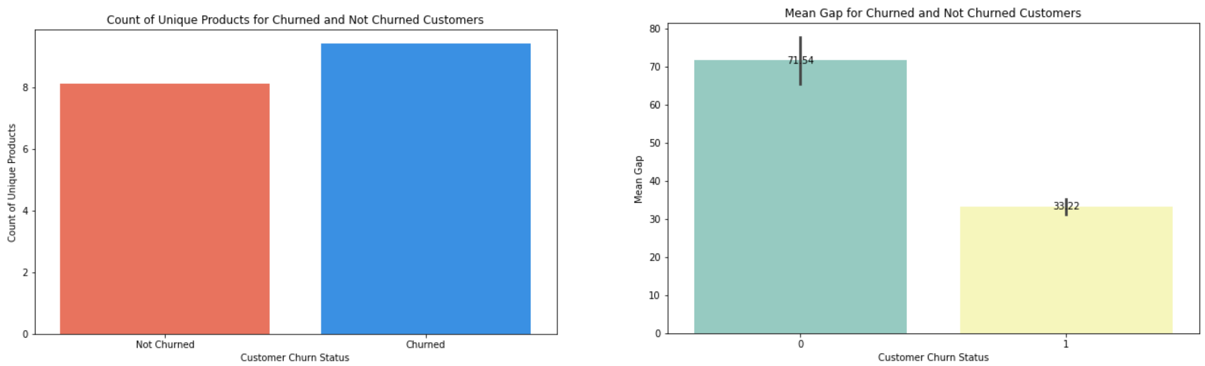


Figure.5: Insights for Churned and not Churned customers (0,1)

1. Data Preprocessing: The analysis begins with data preprocessing to ensure the quality and integrity of the dataset. This includes handling missing values, removing duplicates, and encoding categorical variables. Additionally, feature engineering techniques are applied to extract relevant features that capture customer behaviour and spending patterns. Notably, the creation of the "active_customers" feature, which categorizes customers based on their spending activity, serves as a pivotal step in segmenting the customer base for further analysis.

2. Churn Determination Process: A meticulous churn determination process is employed to accurately classify customers as churned or non-churned. This process hinges on the timing of customers' last purchases relative to predefined observation windows. Specifically, the difference between the maximum purchase day within the initial tumbling window and the minimum purchase day within the output window is calculated. If this difference exceeds a predetermined threshold, typically set at 51 days, the customer is categorized as churned; otherwise, they are labeled as non-churned. This method ensures the accurate identification of churned customers, forming the foundation for subsequent churn prediction and retention efforts.

3. Exploratory Data Analysis (EDA): Exploratory data analysis techniques, including visualization and correlation analysis, are employed to uncover insights into customer behaviour and preferences. Visualizations such as histograms, scatter plots, and box plots provide valuable insights into the distribution of key variables, including purchase frequency, spending patterns, and number of unique products between purchases. These visual aids facilitate a deeper understanding of customer dynamics and inform targeted marketing strategies aimed at reducing churn and enhancing customer retention.

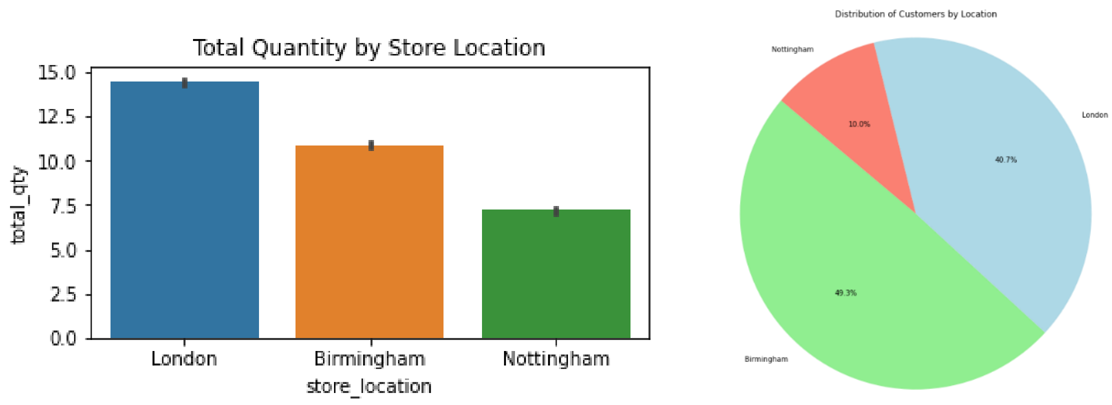


Figure.6: Exploratory Data Analysis (by Location)

4. Predictive Modeling:

Predictive modeling techniques play a pivotal role in forecasting churn rates and identifying customers at risk of churn. Through the utilization of various machine learning algorithms such as logistic regression, decision trees, and random forests, trained on comprehensive historical customer data, future churn patterns can be accurately predicted. By analyzing an array of features ,these models demonstrate efficacy in effectively classifying customers as churned or non-churned. This classification capability enables the implementation of proactive retention initiatives and the development of targeted marketing campaigns aimed at mitigating churn risk and fostering enduring customer loyalty. Additionally, alongside logistic regression, decision trees, and random forests, other models like Linear Support Vector Classifier (SVC) were explored and evaluated to ensure the selection of the most suitable model for FoodCorp's churn prediction needs.

5. Pen Portraits:

Pen portraits are constructed to provide a concise summary of the characteristics and behaviours of churned and non-churned customers who are filtered out first as active customers. These portraits highlight key metrics such as purchase frequency, average spend per quantity, and engagement levels, offering valuable insights for personalized engagement initiatives and targeted communication strategies.

By combining these analytical techniques, businesses can gain a comprehensive understanding of customer churn dynamics and develop data-driven strategies to mitigate churn risk, enhance customer loyalty, and drive sustainable growth.

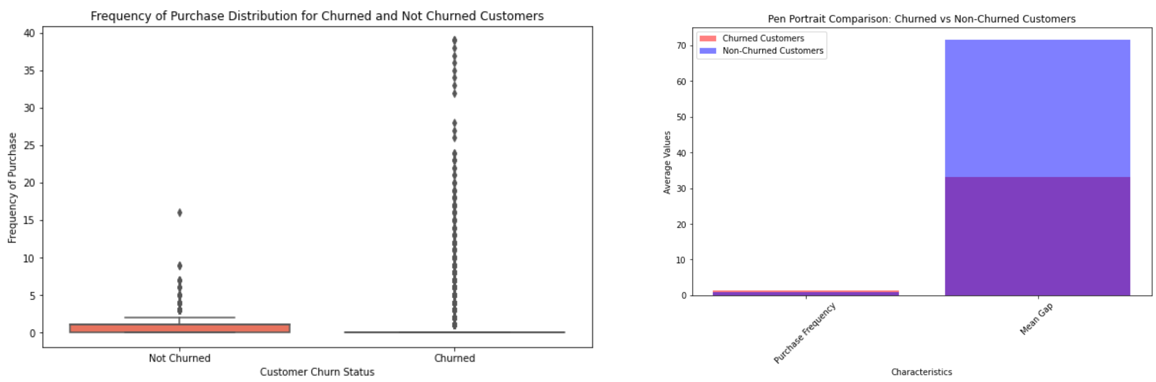


Figure.7: Churners Vs Non Churners