Retail Customer Analytics Report:

Customer Analytics:

Customer analytics is the process of collecting, analyzing, and interpreting customer data to gain insights into their behavior, preferences, and trends. It involves using various analytical techniques and technologies to extract valuable information from customer-related data sources such as transactions, interactions, demographics, social media, and more.

The primary goals of customer analytics are:

1. **Understanding Customer Behavior:** By analyzing patterns in customer data, businesses can gain a deeper understanding of how customers interact with their products or services, their purchase habits, preferences, and the factors that influence their decisions.
2. **Improving Customer Experience:** Customer analytics helps businesses identify areas where they can enhance the customer experience. By understanding customer preferences and pain points, companies can tailor their products, services, and marketing efforts to better meet customer needs.
3. **Increasing Customer Retention:** Analyzing customer data can help businesses identify at-risk customers and take proactive measures to retain them. By identifying patterns associated with customer churn, companies can develop strategies to improve customer loyalty and reduce churn rates.
4. **Targeted Marketing and Personalization:** Customer analytics enables businesses to segment their customer base and target specific groups with personalized marketing campaigns. By delivering relevant and timely messages to customers, businesses can increase engagement and conversion rates.
5. **Optimizing Business Operations:** Customer analytics can also provide insights into the effectiveness of various business processes, such as sales, marketing, and customer service. By analysing data related to these processes, companies can identify inefficiencies and opportunities for improvement.

We have given a retail dataset to analyse. We have performed some initial data cleaning and data manipulations tasks and get some summary statistics at first. Next is we performed the RFM analysis on the given dataset.

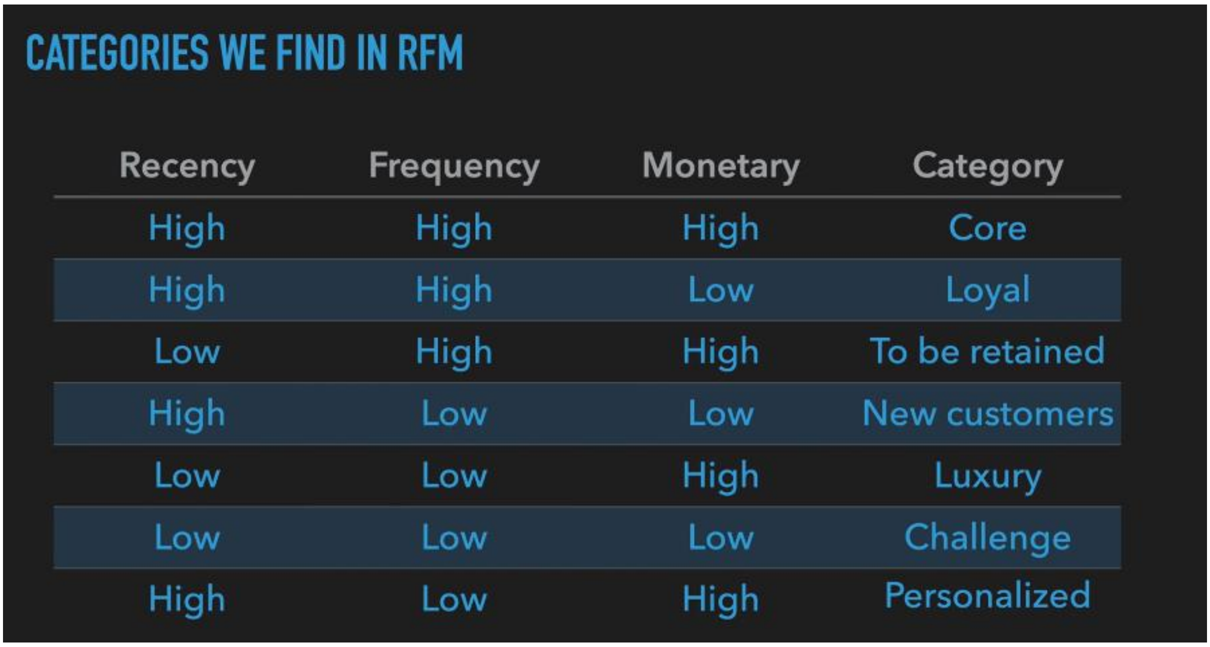
RFM Analysis:

RFM analysis is a data-driven customer segmentation technique used by businesses to categorize customers based on their purchasing behavior. The acronym "RFM" stands for Recency, Frequency, and Monetary Value, which are three key dimensions used to evaluate customer behavior:

1. **Recency (R):** Recency refers to how recently a customer has made a purchase. Customers who have made purchases more recently are often considered more valuable since they are likely to be more engaged with the business.
2. **Frequency (F):** Frequency measures how often a customer makes purchases within a given period. Customers who make frequent purchases are often more loyal and valuable to the business.
3. **Monetary Value (M):** Monetary value represents the total amount of money a customer has spent on purchases. Customers who have spent more money are typically more valuable to the business.

RFM can be done in 2 ways. One is statistical way where we segment each metric into 3 groups and create a multigroup. Second way is calculating RFM using Kmeans clustering where it clusters similar customers into same groups. RFM analysis helps us to find answers to questions like who is the best customer, which of the customers could contribute to the churn rate, which customers can be retained or are more valuable, which customers are likely to respond to engage with the marketing campaigns/ emails etc. 2 approaches to perform RFM is using ranking methods and using the K means method.

Rfm using ranking method:

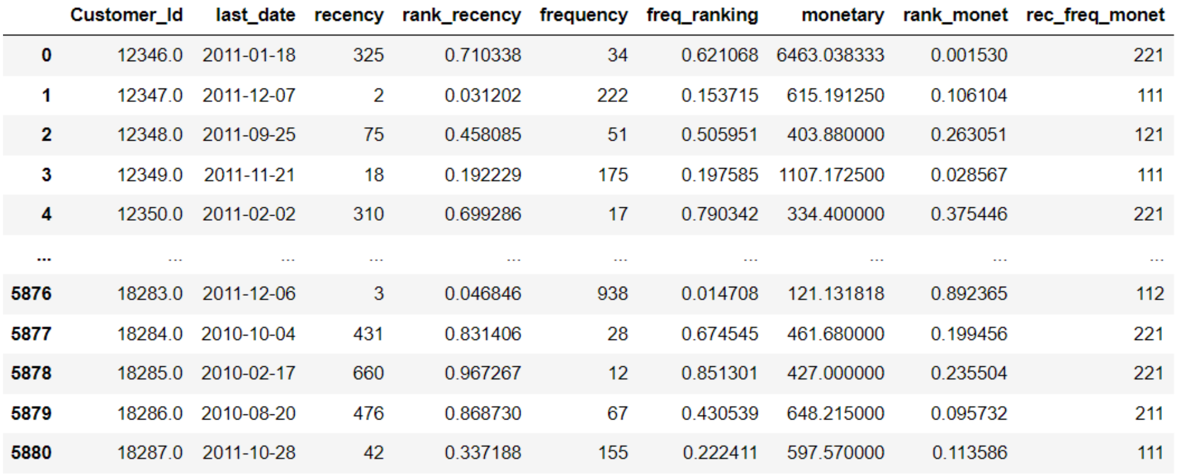


CATEGORIES WE FIND IN RFM

Recency	Frequency	Monetary	Category
High	High	High	Core
High	High	Low	Loyal
Low	High	High	To be retained
High	Low	Low	New customers
Low	Low	High	Luxury
Low	Low	Low	Challenge
High	Low	High	Personalized

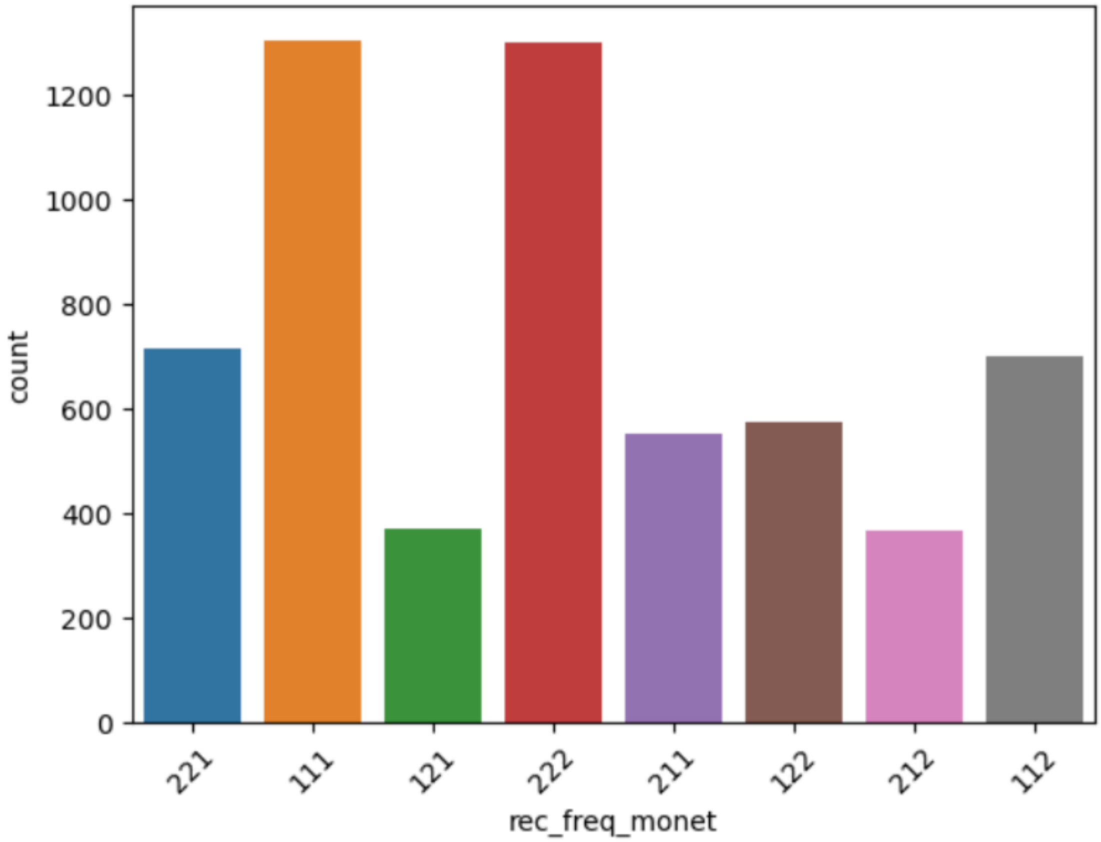
We need to segment customers from the retail dataset into groups based on their recent purchases, how often they buy, and how much they spend. First, we organized customer data by their latest purchase date to determine how recently they've bought something. Then, counted how frequently each customer shops.

Finally, calculated the average amount spent by each customer. These three aspects—recency, frequency, and monetary value—are combined to create a segmentation label for each customer, helping to identify who are the most valuable and engaged customers. This segmentation enables businesses to tailor their marketing efforts and services to meet the specific needs of different customer groups, ultimately enhancing customer satisfaction and loyalty. Created a function to identify the customers that belong to each category and their statistics.



	Customer_Id	last_date	recency	rank_recency	frequency	freq_ranking	monetary	rank_monet	rec_freq_monet
0	12346.0	2011-01-18	325	0.710338	34	0.621068	6463.038333	0.001530	221
1	12347.0	2011-12-07	2	0.031202	222	0.153715	615.191250	0.106104	111
2	12348.0	2011-09-25	75	0.458085	51	0.505951	403.880000	0.263051	121
3	12349.0	2011-11-21	18	0.192229	175	0.197585	1107.172500	0.028567	111
4	12350.0	2011-02-02	310	0.699286	17	0.790342	334.400000	0.375446	221
...	...	...	...	...	...	...	...	...	...
5876	18283.0	2011-12-06	3	0.046846	938	0.014708	121.131818	0.892365	112
5877	18284.0	2010-10-04	431	0.831406	28	0.674545	461.680000	0.199456	221
5878	18285.0	2010-02-17	660	0.967267	12	0.851301	427.000000	0.235504	221
5879	18286.0	2010-08-20	476	0.868730	67	0.430539	648.215000	0.095732	211
5880	18287.0	2011-10-28	42	0.337188	155	0.222411	597.570000	0.113586	111

```
Text(7, 0, '112')]
```

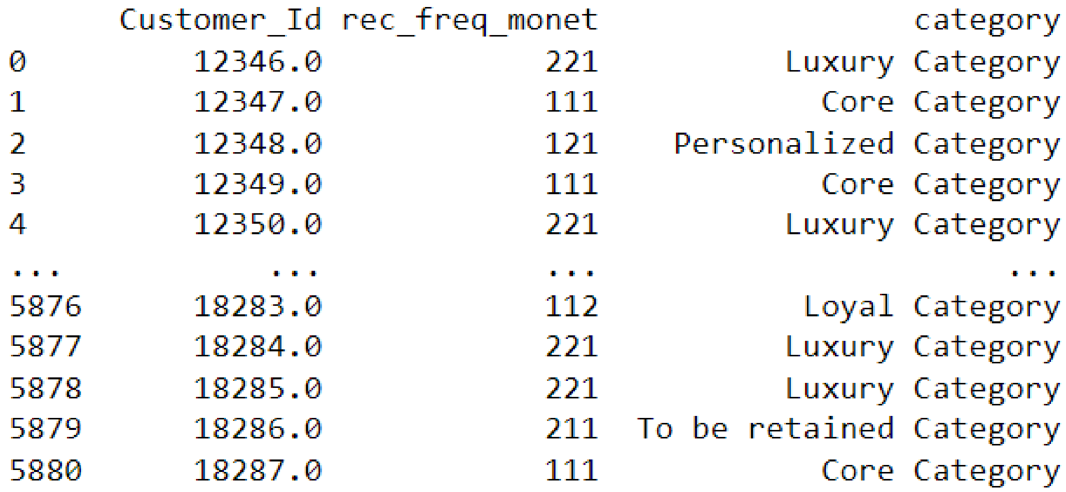


1. **111 (Core):** Customers with a high recency, high frequency, and high monetary value fall into the "Core" category. These are the most valuable customers who have recently made frequent purchases and have spent a significant amount of money. They are the backbone of the business and should be prioritized for personalized offers and exclusive benefits to maintain their loyalty.
2. **112 (Loyal):** Customers with high recency and frequency but low monetary value belong to the "Loyal" category. Although they shop frequently, they may not spend as much per transaction. Despite this, they are still loyal to the brand and regularly engage with it. They should be nurtured with targeted promotions and incentives to increase their spending.
3. **211(to be retained):** Customers with high frequency and monetary value but low recency are classified as "To be retained." These customers have made significant purchases in the past but haven't returned as often. They represent an opportunity for re-engagement, and efforts should be made to encourage repeat purchases and strengthen their loyalty.
4. **122(New Customers):** Customers with high recency, low frequency, and low monetary value belong to the " New customers " category. These are recent additions to the customer base who shop frequently but may not have spent much yet. They represent an opportunity for growth, and strategies should focus on converting them into higher-value customers through targeted marketing and promotions.
5. **221 (Luxury):** Customers with low recency, frequency, and high monetary value are categorized as "Luxury." These customers may not make frequent purchases or engage with your business recently, but when they do, they tend to spend significantly higher amounts compared to other customers. They likely have specific tastes or preferences and are willing to invest in premium products or services. Despite their infrequent interactions, they contribute substantially to your revenue and profit margins. Maintaining a positive relationship with these customers is crucial, as they are valuable assets to your business and may require personalized attention to ensure continued loyalty and satisfaction.
6. **222 (Challenge):** Customers with low recency, low frequency, and low monetary value are categorized as "Challenge." These customers present a significant opportunity for engagement and conversion. Despite their low spending and infrequent transactions, their consistent presence indicates some level of interest in your products or services. However, they may require targeted efforts to increase their purchase frequency and average spending.
7. **121 (Personalized):** Customers with high recency, high monetary value, but low frequency are categorized as "Personalized." These customers have made recent high-value purchases, indicating a strong interest and willingness to spend, but they do so infrequently. Despite their low frequency, their high spending makes them valuable to your business. They may have

specific preferences or needs that require personalized attention to encourage more frequent engagement and repeat purchases.

Identifying some category of customers:

Created a function that mapped the values of the 'recency frequency monetary' metric to their corresponding categories.



```
      Customer_Id rec_freq_monet                  category
0         12346.0            221          Luxury Category
1         12347.0            111            Core Category
2         12348.0            121    Personalized Category
3         12349.0            111            Core Category
4         12350.0            221          Luxury Category
...           ...            ...                      ...
5876      18283.0            112           Loyal Category
5877      18284.0            221          Luxury Category
5878      18285.0            221          Luxury Category
5879      18286.0            211  To be retained Category
5880      18287.0            111            Core Category
```

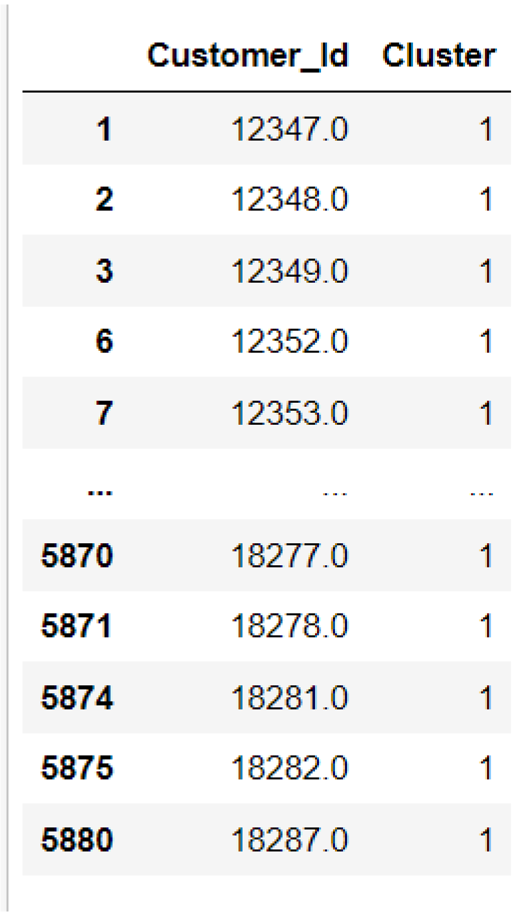
RFM using K means method:

To perform RFM (Recency, Frequency, Monetary) analysis using the k-means clustering method, we   follow these steps:

1. Prepare the RFM data: Calculate Recency, Frequency, and Monetary values for each customer.
2. Standardize the data: Normalize the RFM values to have a mean of 0 and a standard deviation of 1.
3. Choose the number of clusters (K): Determine the optimal number of clusters using techniques like the elbow method or silhouette score.
4. Apply k-means clustering: Use the standardized RFM data to perform k-means clustering.
5. Analyze the clusters: Examine the characteristics of each cluster to gain insights into customer segments.

Identifying the customers from each clusters:

These are the customers that belong to cluster 1



	Customer_Id	Cluster
1	12347.0	1
2	12348.0	1
3	12349.0	1
6	12352.0	1
7	12353.0	1
...	...	...
5870	18277.0	1
5871	18278.0	1
5874	18281.0	1
5875	18282.0	1
5880	18287.0	1

Customer Life time Value(CLV):

Customer lifetime value (CLV or CLTV) is a metric that indicates the total revenue a business can reasonably expect from a single customer account throughout the business relationship. The metric considers a customer's revenue value and compares that number to the company's predicted customer lifespan.

And this is a tool mainly used by marketers to identify the customers and if they subject to be very good customers for the brand, or occasional customers or customers who will not come all the time. And as you know that any brand and any retailer spend a lot of capital, a lot of investment on marketing and marketing campaigns So identifying if the customer.Is a valuable customer or not will better strategize the use of this marketing expenditure towards the customer that we believe will bring high value towards the future.

To calculate customer lifetime value (CLV), we utilize RFM features including recency, average monetary value, and frequency. We further split the RFM score into three columns corresponding to recency, frequency, and monetary value. By summing these grades, we obtain an overall RFM score. With a dataset spanning two and a half years, we split it into training and testing sets. CLV can be represented as either a continuous dollar value or categorized into high, medium, and low segments using techniques like K-means clustering.

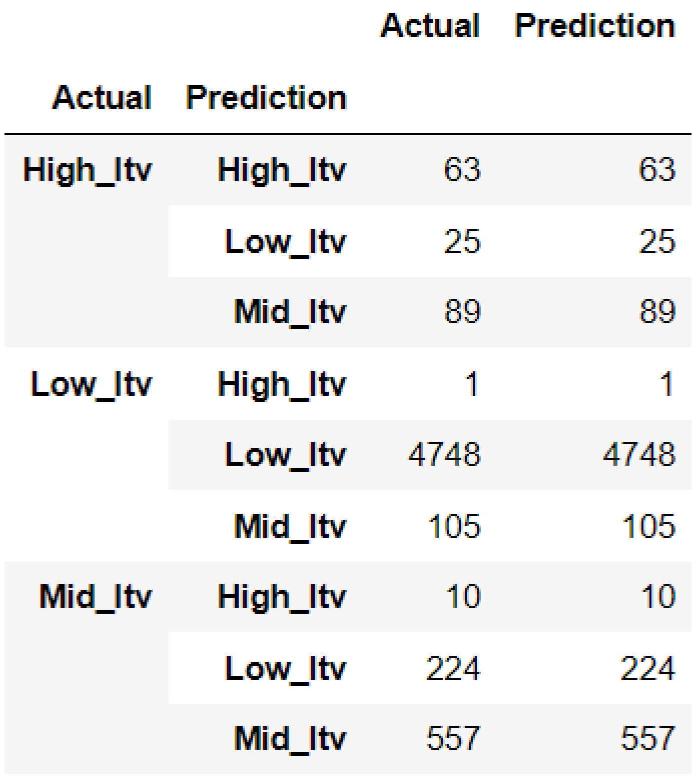
Subsequently, we train two models: a decision tree model and a multinomial logistic regression model due to the presence of three CLV categories. Logistic regression would suffice for binary categorization. These models enable us to predict future customer spending and tailor marketing efforts towards segments likely to yield high revenue.

To preprocess the RFM data, we first split the 'rec_freq_monet' column into three separate columns representing recency, frequency, and monetary value groups. Then, we mapped the values in each group according to a predefined value map. Subsequently, we calculated the overall RFM score by summing up the scores from each group. To calculate the Customer Lifetime Value (CLV), we aggregated the total revenue for each customer by summing their revenue transactions. This resulted in a dataset where each row represents a unique customer ID along with their corresponding CLV, which we denote as "ltv."

When we plotted the histogram of Ltv value we are bot able to visualise properly due to the presence of outliers which we conformed through a box plot. So we just removed the outliers by keeping only 99% of our data, removing those last few outliers.This is to remove the noise when we are training the model and to make the model more flexible and less biased to extreme values.

We proceeded to apply K-means clustering to classify Customer Lifetime Value (CLV) into three categories: low, mid, and high. After initializing the KMeans model with three clusters, we fitted it to the CLV data. The resulting clusters were then analyzed based on their mean CLV values. Cluster 0 was designated as representing low CLV customers, cluster 1 as mid CLV customers, and cluster 2 as high CLV customers. We then mapped these clusters to their corresponding CLV categories: low, mid, and high, for further analysis and targeting in marketing strategies.

We proceeded to model Customer Lifetime Value (CLV) using logistic regression, multinomial logistic regression, and decision tree classifier. Given the limited dataset size of 5800 observations, we opted for cross-validation to evaluate model performance robustly. For logistic regression and multinomial logistic regression, we utilized the RepeatedStratifiedKFold cross-validation method with three splits and three repeats. The resulting mean accuracy across iterations was found to be approximately 86%. We then pursued parameter tuning for the decision tree classifier using RandomizedSearchCV, aiming to enhance model performance. The best score obtained was 89%, indicating a slight improvement over the untuned model. Subsequently, we compared the predictions with the actual values, revealing an overall accuracy of 92%. These results suggest promising performance in predicting CLV, highlighting the potential effectiveness of machine learning algorithms in identifying high-value customers.



		Actual	Prediction
Actual	Prediction		
High_ltv	High_ltv	63	63
	Low_ltv	25	25
	Mid_ltv	89	89
Low_ltv	High_ltv	1	1
	Low_ltv	4748	4748
	Mid_ltv	105	105
Mid_ltv	High_ltv	10	10
	Low_ltv	224	224
	Mid_ltv	557	557

Market Basket Analysis:

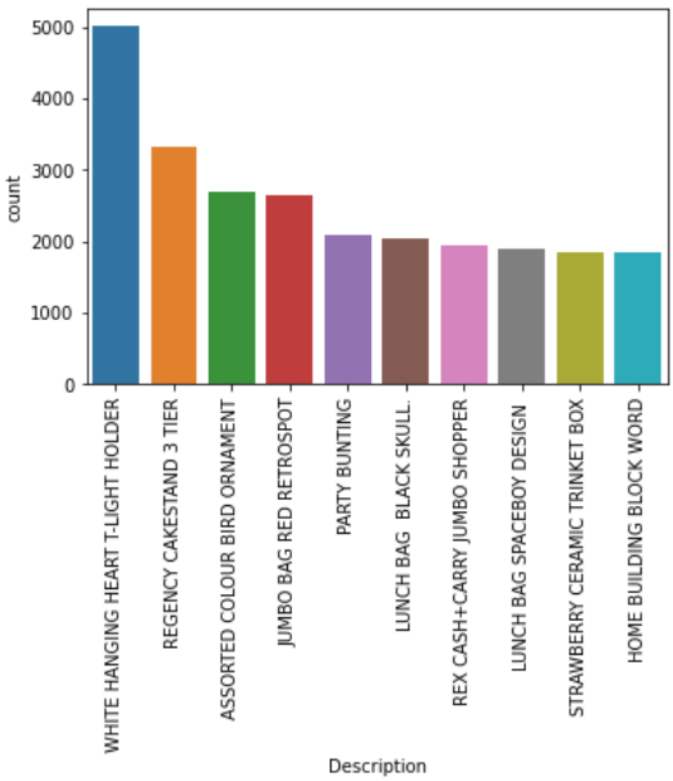
Market basket analysis is a data mining technique used to identify associations between products that are frequently purchased together by customers. It involves analysing transactional data to uncover patterns and relationships between items bought during the same shopping trip. The primary goal is to understand customer purchasing behaviour and to derive insights that can be used for various purposes such as product recommendations, cross-selling, and merchandising strategies..We conducted market basket analysis on the UK retailer dataset, generating rules for each product. These rules serve as recommendations; when a customer purchases a particular product, associated products or rules will be suggested to them. This facilitates personalized recommendations based on customer behaviour and preferences.

Market basket analysis is utilized to capture associations among items typically purchased together, with three key performance indicators (KPIs) to measure these associations. The first KPI is support, which calculates the frequency of item X and item Y being bought together over the total number of orders. Confidence measures how many times item X and item Y were bought together over the frequency of item X alone. Finally, lift evaluates the support of the association between item X and item Y relative to the individual supports of X and Y. These KPIs are assessed using the Apriori algorithm for smooth measurement.

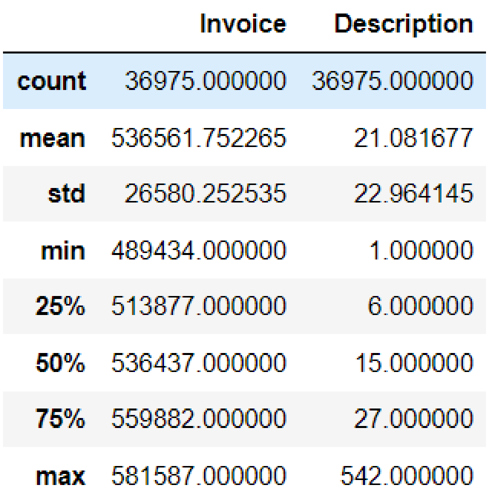
$$\begin{tikzcd} & Support = \frac{frq(X,Y)}{N} \\ Rule\colon X \Rightarrow Y \arrow[ur] \arrow[r] \arrow[dr] & Confidence = \frac{frq(X,Y)}{frq(X)} \\ & Lift = \frac{Support}{Supp(X) \times Supp(Y)} \end{tikzcd}$$

The market basket analysis process involves several key steps. Firstly, transactions are organized into baskets, where each basket represents a single order containing multiple items. Secondly, thresholds are set for rules to filter out insignificant occurrences. This ensures that only meaningful associations between items are captured, taking into account factors like confidence and support levels. Finally, the Apriori algorithm is applied to identify association rules, which are then used to make recommendations to customers based on their purchases. This allows for personalized recommendations to be made to customers based on their buying patterns and the relationships between different items in the market basket.

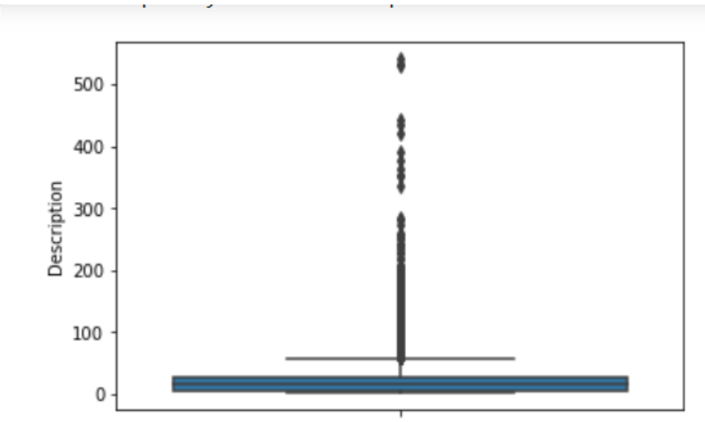
We will be using the cleaned retail dataset for the analysis. First visualize which items appear the most. WHITE HANGING HEART T-LIGHT HOLDER is the most popular item followed by ‘REGENCY CAKESTAND 3 TIER'



Next its good to know the to determine the average order size, which reveals important insights into basket composition. With an average of 21 items per basket and a 50th percentile of 15 items, it's evident that there are some significantly large orders. The disparity between the mean and median indicates the presence of outliers, likely contributing to the skew in distribution. Utilizing a box plot can visually represent this distribution, highlighting the range and spread of order sizes within the dataset.

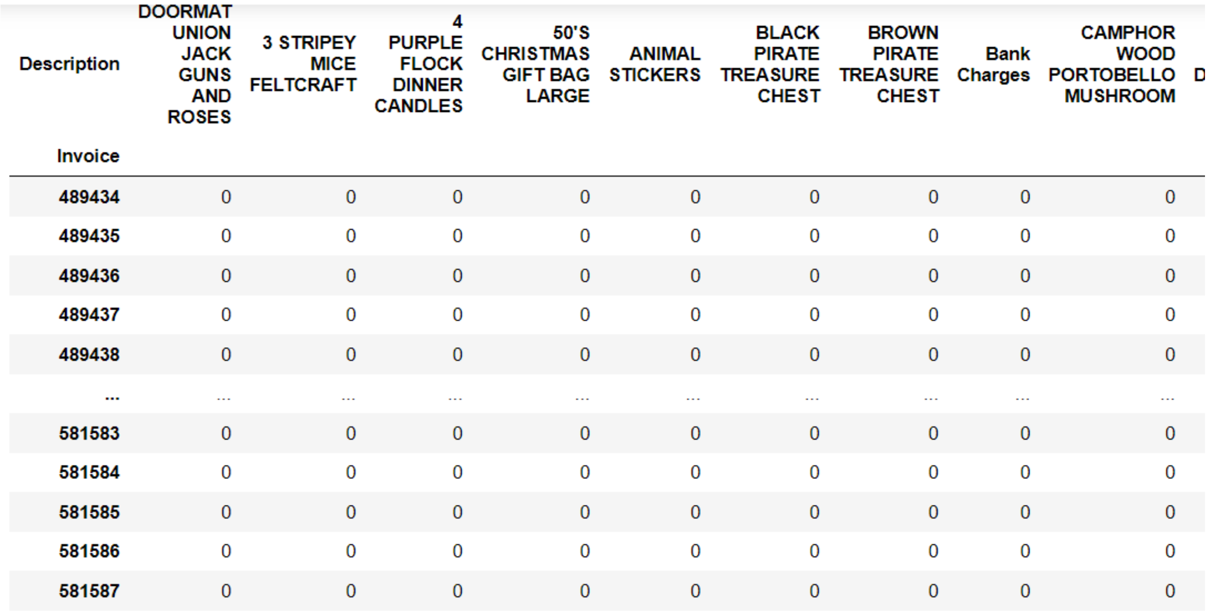


	Invoice	Description
count	36975.000000	36975.000000
mean	536561.752265	21.081677
std	26580.252535	22.964145
min	489434.000000	1.000000
25%	513877.000000	6.000000
50%	536437.000000	15.000000
75%	559882.000000	27.000000
max	581587.000000	542.000000



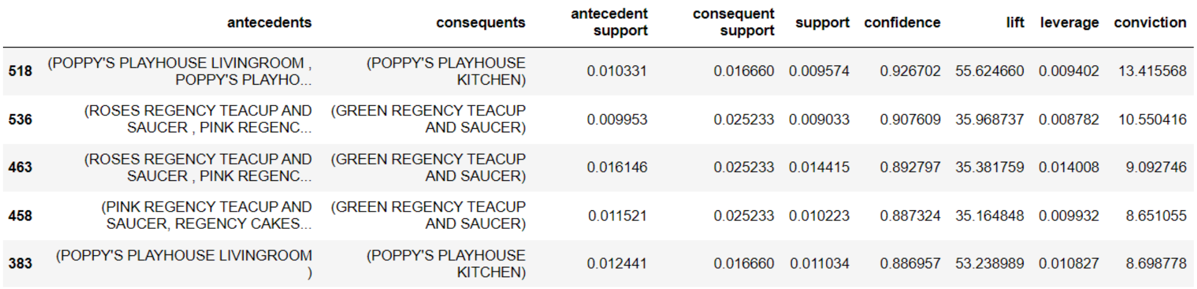
We can clearly see that we have a lot of outliers because the whole distribution is under 100.But we do have these orders that are above 300, 400 and even 500. Because this is a wholesale data, that means it's B2B business to business, that’s why the number of items as it's a wholesaler business is high. And that's why you have many items inside the order. Of course, in business to consumer, the amount will be different.

So now we need to prepare the data for the apriori algorithm or the market basket analysis. The data structure should be every invoice has item as 0 or 1. Items should be on the columns and the invoice should be the index. Just a count is required whether this item appear in this order or not. If item is bought in an order 1 else 0.



Description	DOORMAT UNION JACK GUNS AND ROSES	3 STRIPEY MICE FELTCRAFT	4 PURPLE FLOCK DINNER CANDLES	50'S CHRISTMAS GIFT BAG LARGE	ANIMAL STICKERS	BLACK PIRATE TREASURE CHEST	BROWN PIRATE TREASURE CHEST	Bank Charges	CAMPHOR WOOD PORTOBELLO MUSHROOM	D
Invoice										
489434	0	0	0	0	0	0	0	0	0	
489435	0	0	0	0	0	0	0	0	0	
489436	0	0	0	0	0	0	0	0	0	
489437	0	0	0	0	0	0	0	0	0	
489438	0	0	0	0	0	0	0	0	0	
...	...	...	...	...	...	...	...	...	...	
581583	0	0	0	0	0	0	0	0	0	
581584	0	0	0	0	0	0	0	0	0	
581585	0	0	0	0	0	0	0	0	0	
581586	0	0	0	0	0	0	0	0	0	
581587	0	0	0	0	0	0	0	0	0	

Now we will turn it to apriori form which is an antecedent and consequence. So the antecedent is the main item and the consequent is the item that is driven by the antecedent. he **apriori()** function from a library, likely **mlxtend**, is then used to generate frequent itemsets from the basket data. Frequent itemsets are sets of items that frequently occur together in transactions. In this step, **min_support** is specified as 0.009, which sets the minimum support threshold for an itemset to be considered frequent. The frequent itemsets obtained in the previous step are then used to generate association rules using the **association_rules()** function. Association rules represent relationships between items in the dataset. In this step, the **metric** parameter is set to 'lift', indicating that the lift metric will be used to evaluate the strength of association between items. The generated association rules are sorted based on confidence in descending order using the **sort_values()** function. Confidence is a measure of the reliability of the rule, indicating the likelihood that the consequent (items on the right-hand side of the rule) will be purchased given that the antecedent (items on the left-hand side of the rule) is purchased.



	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
518	(POPPY'S PLAYHOUSE LIVINGROOM , POPPY'S PLAYHO...	(POPPY'S PLAYHOUSE KITCHEN)	0.010331	0.016660	0.009574	0.926702	55.624660	0.009402	13.415568
536	(ROSES REGENCY TEACUP AND SAUCER , PINK REGENC...	(GREEN REGENCY TEACUP AND SAUCER)	0.009953	0.025233	0.009033	0.907609	35.968737	0.008782	10.550416
463	(ROSES REGENCY TEACUP AND SAUCER , PINK REGENC...	(GREEN REGENCY TEACUP AND SAUCER)	0.016146	0.025233	0.014415	0.892797	35.381759	0.014008	9.092746
458	(PINK REGENCY TEACUP AND SAUCER, REGENCY CAKES...	(GREEN REGENCY TEACUP AND SAUCER)	0.011521	0.025233	0.010223	0.887324	35.164848	0.009932	8.651055
383	(POPPY'S PLAYHOUSE LIVINGROOM)	(POPPY'S PLAYHOUSE KITCHEN)	0.012441	0.016660	0.011034	0.886957	53.238989	0.010827	8.698778

How to analyse this can be shown using an example.

Antecedents: Pink Regency Teacup and Saucers, Regency Cakes Consequents: Green Regency Teacup and Saucer.

Support:

- Antecedent support: 1.15%
- Consequent support: 2.52%
- Support for both: 1.02%

This indicates that 1.15% of transactions include Pink Regency Teacup and Saucers and Regency Cakes together, 2.52% include Green Regency Teacup and Saucer alone, and 1.02% include both combinations in the same transaction.

1. **Confidence**:
    - Confidence: 88%

The confidence of 88% suggests that when Pink Regency Teacup and Saucers and Regency Cakes are purchased together, there is an 88% likelihood that Green Regency Teacup and Saucer will also be purchased.

2. **Lift**:
    - Lift: 35.16

A lift of 35.16 indicates a strong association between the antecedents and consequents. It means that the likelihood of purchasing Green Regency Teacup and Saucer increases by 35.16 times when Pink Regency Teacup and Saucers and Regency Cakes are purchased together, compared to when they are purchased independently.

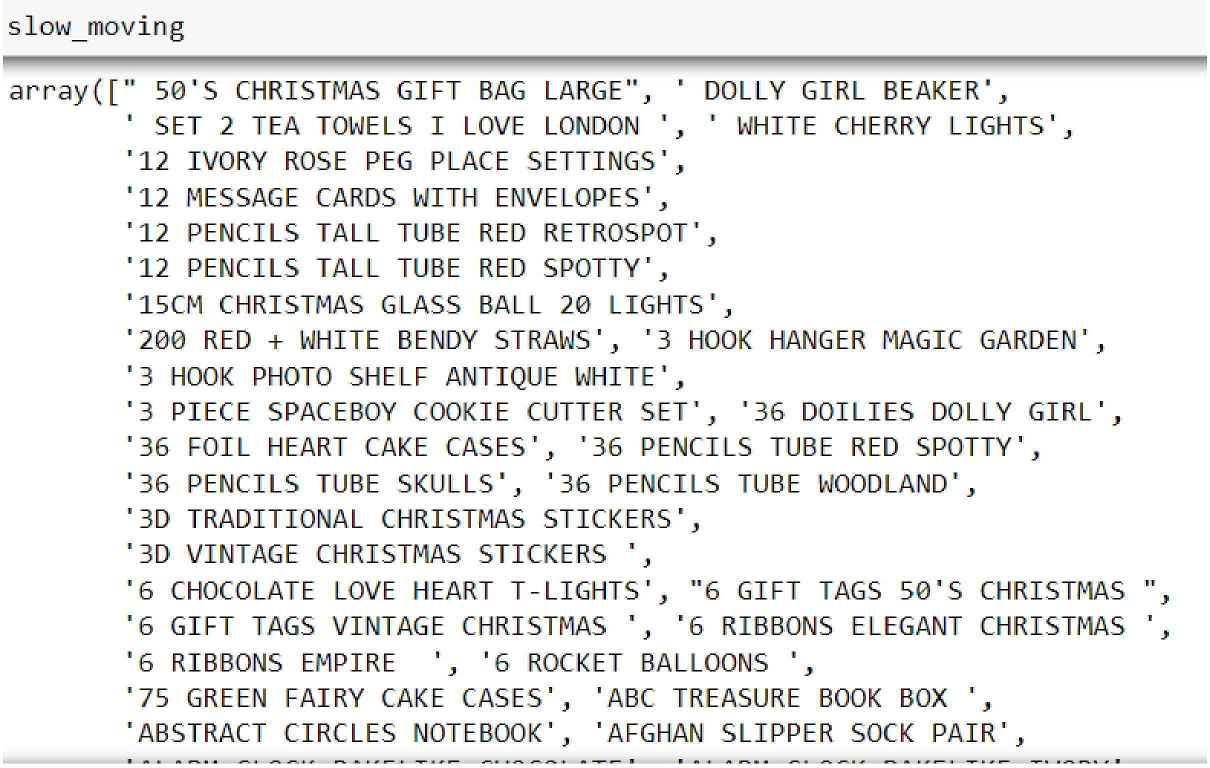
3. **Conviction**:
    - Conviction: 8.65

Conviction measures the degree of dependence between the antecedents and consequents. A conviction of 8.65 suggests that if a customer does not buy Pink Regency Teacup and Saucers and Regency Cakes together, they are 8.65 times more likely not to buy Green Regency Teacup and Saucer as well.

Based on these insights, the recommended actionable points are :

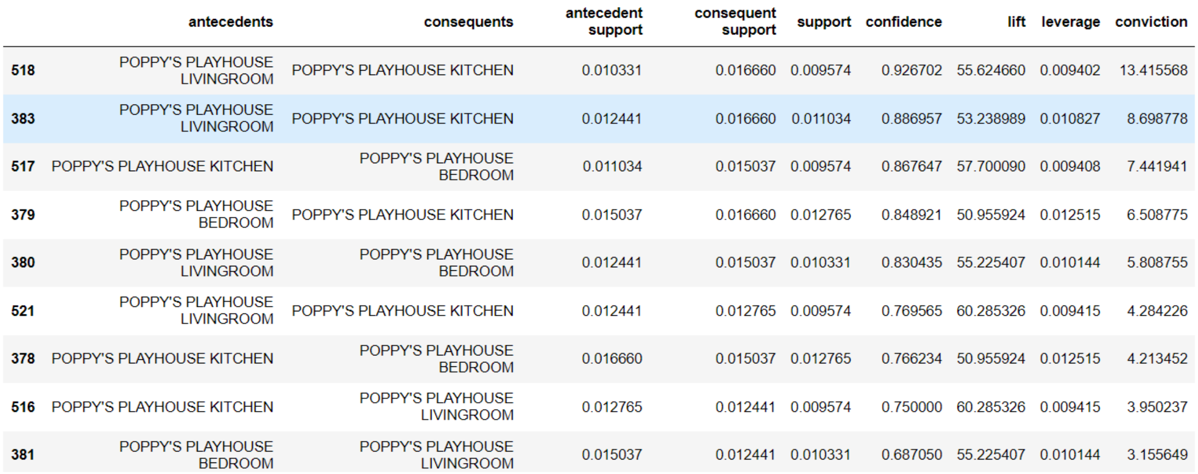
- Retailers can strategically place Pink Regency Teacup and Saucers and Regency Cakes near Green Regency Teacup and Saucer to encourage joint purchases, potentially increasing sales.
- Marketing campaigns or promotions can be targeted towards customers who purchase Pink Regency Teacup and Saucers and Regency Cakes to also consider buying Green Regency Teacup and Saucer, leveraging the high confidence level.
- Bundling or cross-selling strategies can be implemented to offer discounts or incentives for purchasing the combinations identified in the association rule, maximizing customer satisfaction and revenue.

Another thing we can do is to improve the sales of slow-moving items. First identify the slow-moving items. Then see where the slow-moving items are consequent and identify its antecedent. The recommended strategy should be to offer a special discount for the bundle of those 2 products together. By this we can push the sales of slow-moving items.



```
slow_moving
```

```
array([" 50'S CHRISTMAS GIFT BAG LARGE", ' DOLLY GIRL BEAKER',
       ' SET 2 TEA TOWELS I LOVE LONDON ', ' WHITE CHERRY LIGHTS',
       '12 IVORY ROSE PEG PLACE SETTINGS',
       '12 MESSAGE CARDS WITH ENVELOPES',
       '12 PENCILS TALL TUBE RED RETROSPOT',
       '12 PENCILS TALL TUBE RED SPOTTY',
       '15CM CHRISTMAS GLASS BALL 20 LIGHTS',
       '200 RED + WHITE BENDY STRAWS', '3 HOOK HANGER MAGIC GARDEN',
       '3 HOOK PHOTO SHELF ANTIQUE WHITE',
       '3 PIECE SPACEBOY COOKIE CUTTER SET', '36 DOILIES DOLLY GIRL',
       '36 FOIL HEART CAKE CASES', '36 PENCILS TUBE RED SPOTTY',
       '36 PENCILS TUBE SKULLS', '36 PENCILS TUBE WOODLAND',
       '3D TRADITIONAL CHRISTMAS STICKERS',
       '3D VINTAGE CHRISTMAS STICKERS ',
       '6 CHOCOLATE LOVE HEART T-LIGHTS', "6 GIFT TAGS 50'S CHRISTMAS ",
       '6 GIFT TAGS VINTAGE CHRISTMAS ', '6 RIBBONS ELEGANT CHRISTMAS ',
       '6 RIBBONS EMPIRE  ', '6 ROCKET BALLOONS ',
       '75 GREEN FAIRY CAKE CASES', 'ABC TREASURE BOOK BOX ',
       'ABSTRACT CIRCLES NOTEBOOK', 'AFGHAN SLIPPER SOCK PAIR',
       'ALARM CLOCK BAKELIKE CHOCOLATE', 'ALARM CLOCK BAKELIKE IVORY'
```



	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
518	POPPY'S PLAYHOUSE LIVINGROOM	POPPY'S PLAYHOUSE KITCHEN	0.010331	0.016660	0.009574	0.926702	55.624660	0.009402	13.415568
383	POPPY'S PLAYHOUSE LIVINGROOM	POPPY'S PLAYHOUSE KITCHEN	0.012441	0.016660	0.011034	0.886957	53.238989	0.010827	8.698778
517	POPPY'S PLAYHOUSE KITCHEN	POPPY'S PLAYHOUSE BEDROOM	0.011034	0.015037	0.009574	0.867647	57.700090	0.009408	7.441941
379	POPPY'S PLAYHOUSE BEDROOM	POPPY'S PLAYHOUSE KITCHEN	0.015037	0.016660	0.012765	0.848921	50.955924	0.012515	6.508775
380	POPPY'S PLAYHOUSE LIVINGROOM	POPPY'S PLAYHOUSE BEDROOM	0.012441	0.015037	0.010331	0.830435	55.225407	0.010144	5.808755
521	POPPY'S PLAYHOUSE LIVINGROOM	POPPY'S PLAYHOUSE KITCHEN	0.012441	0.012765	0.009574	0.769565	60.285326	0.009415	4.284226
378	POPPY'S PLAYHOUSE KITCHEN	POPPY'S PLAYHOUSE BEDROOM	0.016660	0.015037	0.012765	0.766234	50.955924	0.012515	4.213452
516	POPPY'S PLAYHOUSE KITCHEN	POPPY'S PLAYHOUSE LIVINGROOM	0.012765	0.012441	0.009574	0.750000	60.285326	0.009415	3.950237
381	POPPY'S PLAYHOUSE BEDROOM	POPPY'S PLAYHOUSE LIVINGROOM	0.015037	0.012441	0.010331	0.687050	55.225407	0.010144	3.155649

So here the slow-moving item POPPY’S PLAYHOUSE KITCHEN should be given a discount for those who buy this product along with POPPY’S PLAYHOUSE LIVING ROOM. Because they are buying those two items together, it will be appealing for them when they find both at a discounted price. And this is one of the main targets of the basket analysis, aside from recommendation to customers.