Mahindra First Choice Services

Presented by :
Vishnu Sreenivasan
Quid Morbiwala
Karan Pande
Hemant Gupta

Business Problem

Geolocation Based Analysis:

- Extract insights regarding the ownership pattern of cars based on location features of the customers.
- Find out how the different type of services and service time vary with location.
- Analyse the revenue generated at different locations throughout the country

Market Segmentation:

- Divide the customer base into different segments based on the activeness, number of visits to garage, service type and revenue generated.
- This will help to better understand the customer behavioural patterns and have targeted marketing campaigns and rewards based on the segments

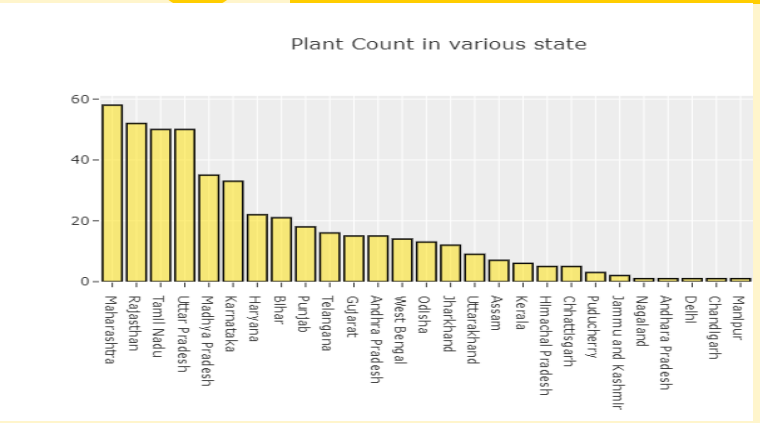
Customer Lifetime Value Prediction:

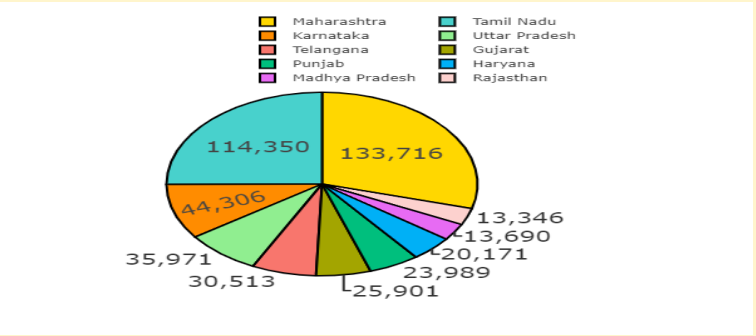
- Predict the value given by the customer over his lifetime based on the average spend over a year and number of visits.
- Analysing CLV would help Mahindra First Choice focus on the customers who are going to generate good value to the company.

Exploratory Data Analysis

Lets Start with the first steps of Visualizations

Geolocation Analysis



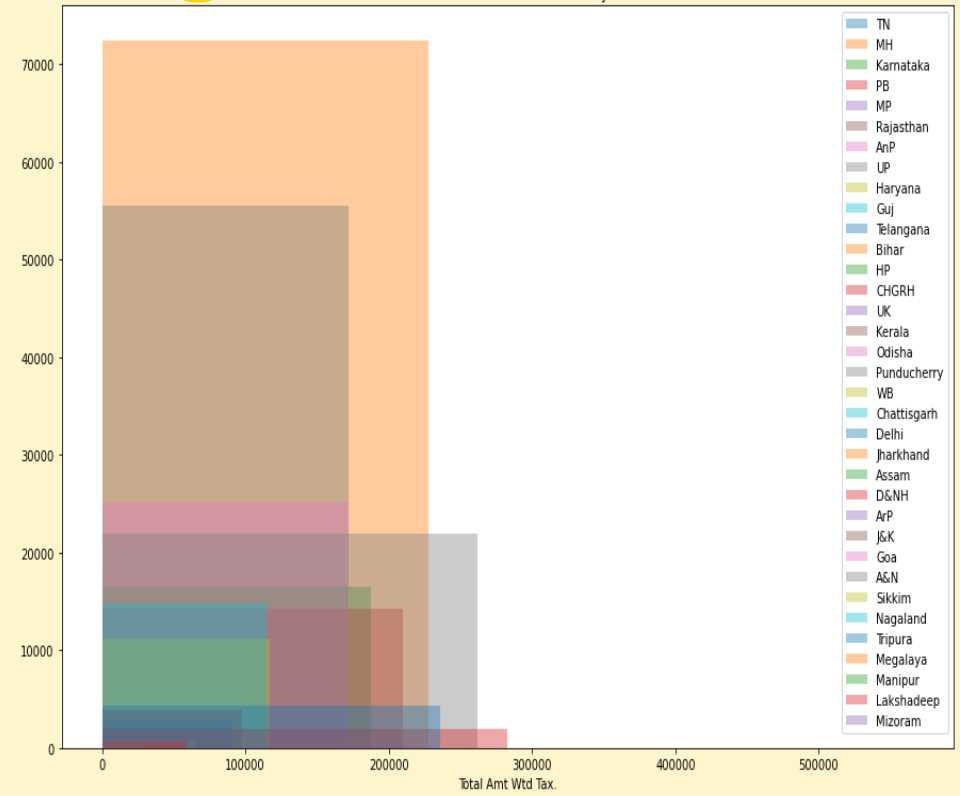


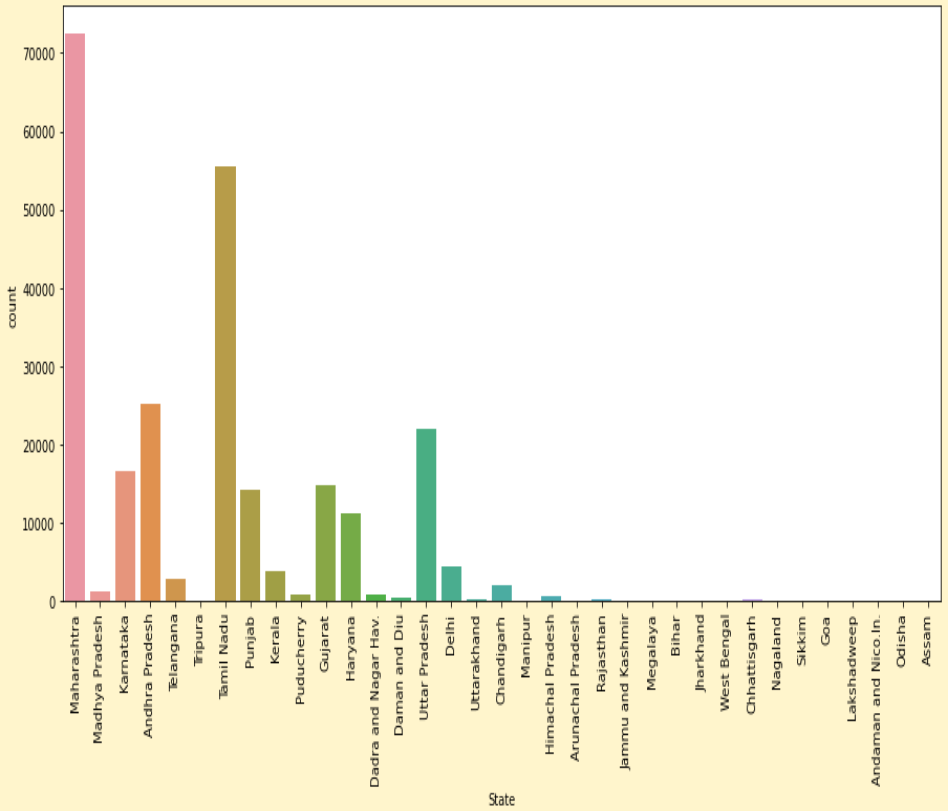
Plant count vs car count

- Maharashtra, Rajasthan, Tamil Nadu, UP,MP tops in number of plants
- Maha,TN, Karnataka, UP, Telangana tops in car count
- Plants vs car count is very less in states like Karnataka, Telangana
- Possible reasons could be car owners in Karnataka ,Telangana are more as compared to other states
- On the other hand Rajasthan is 2nd in terms of count but the car count is very less.
- Possible reasons could be customers shift to other service provider or car owners are less in this state as compared to other states

Which state has most revenue?

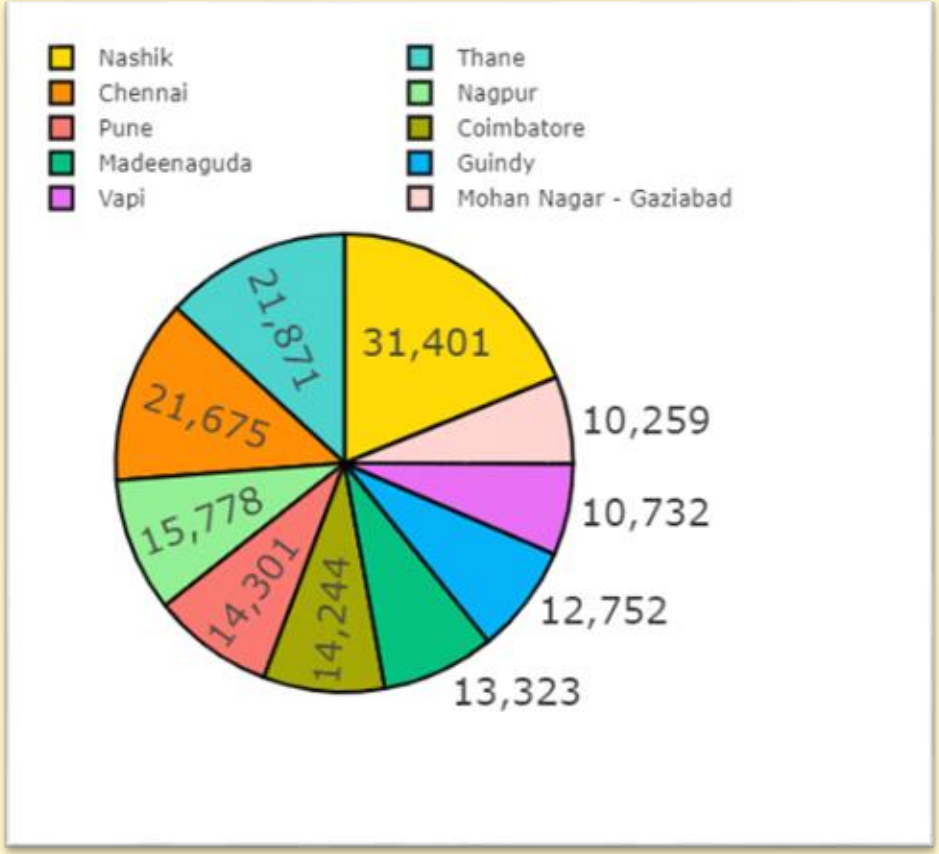
Geo locaton based Revenue Analysis





- Maharashtra has highest customer base followed by Tamil Nadu.
- Top 3 revenue generating states are MH TN AP

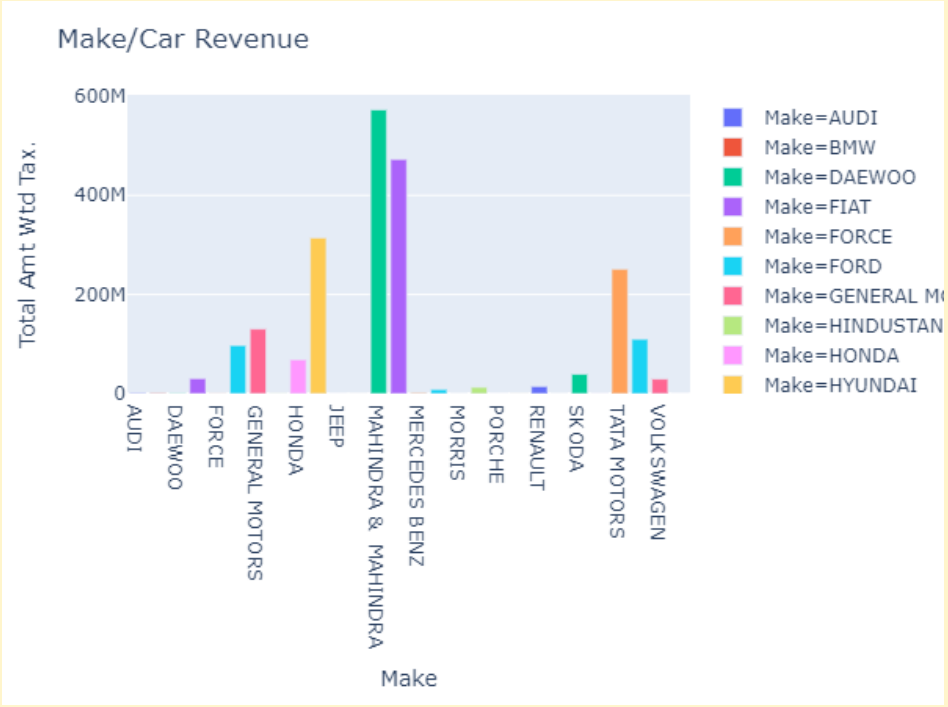
Which city has most revenue

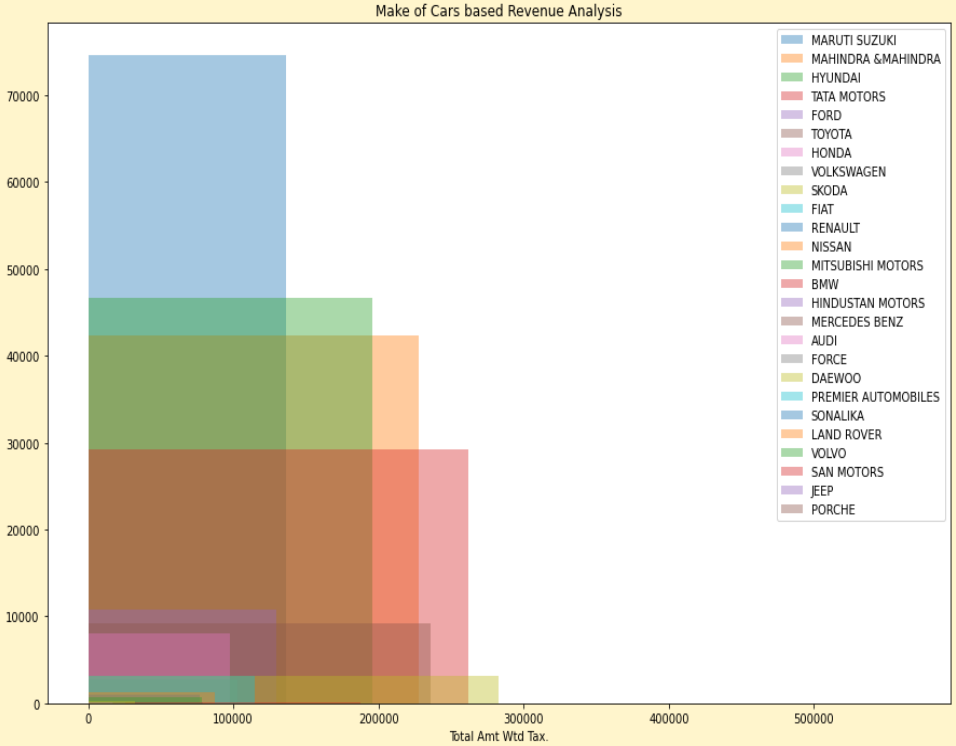


Top 5 cities given above top the list, this is obvious as they are top major metropolitan cities in India

- **Nashik(31,401)**
- **Thane(21,871)**
- **Chennai(21,675)**
- **Nagpur(15,778)**
- **Pune(14,301)**

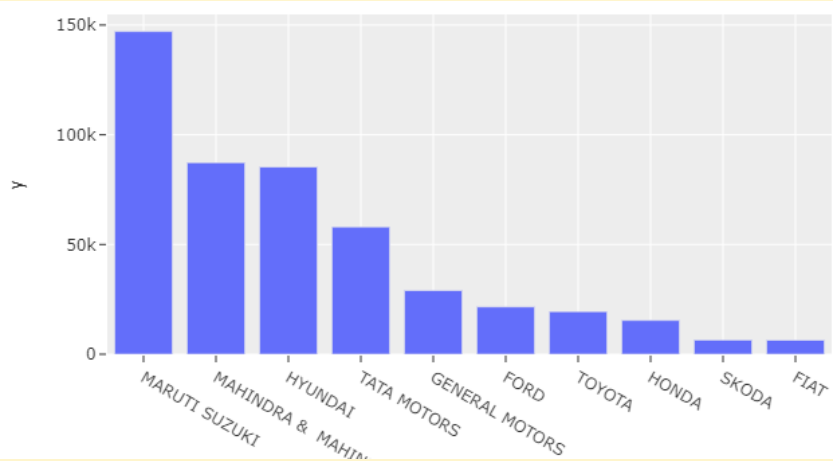
Which make/car is more popular?

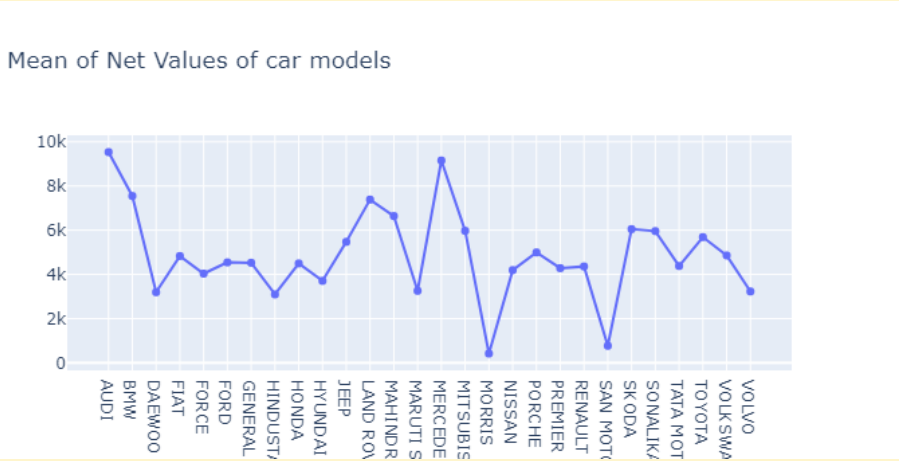




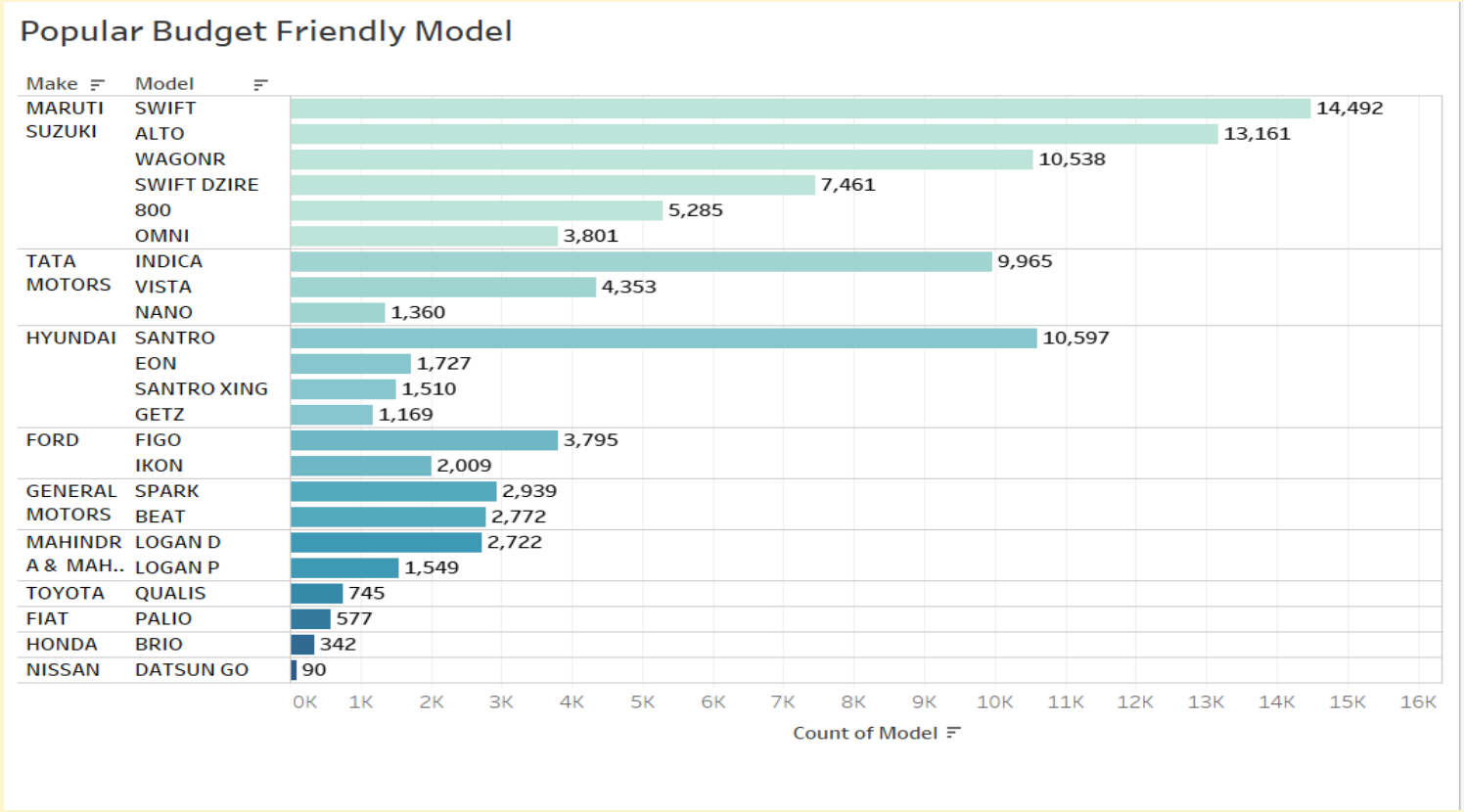
- M&M contributed the most towards revenue ,followed by Maruti Suzuki, Hyundai, Tata Motors
- Meanwhile, Maruti Suzuki tops on the basis of count but the revenue produced is less as compared to Mahindra and Mahindra
- Make of Cars based Revenue Analysis

Which make/car is more popular?

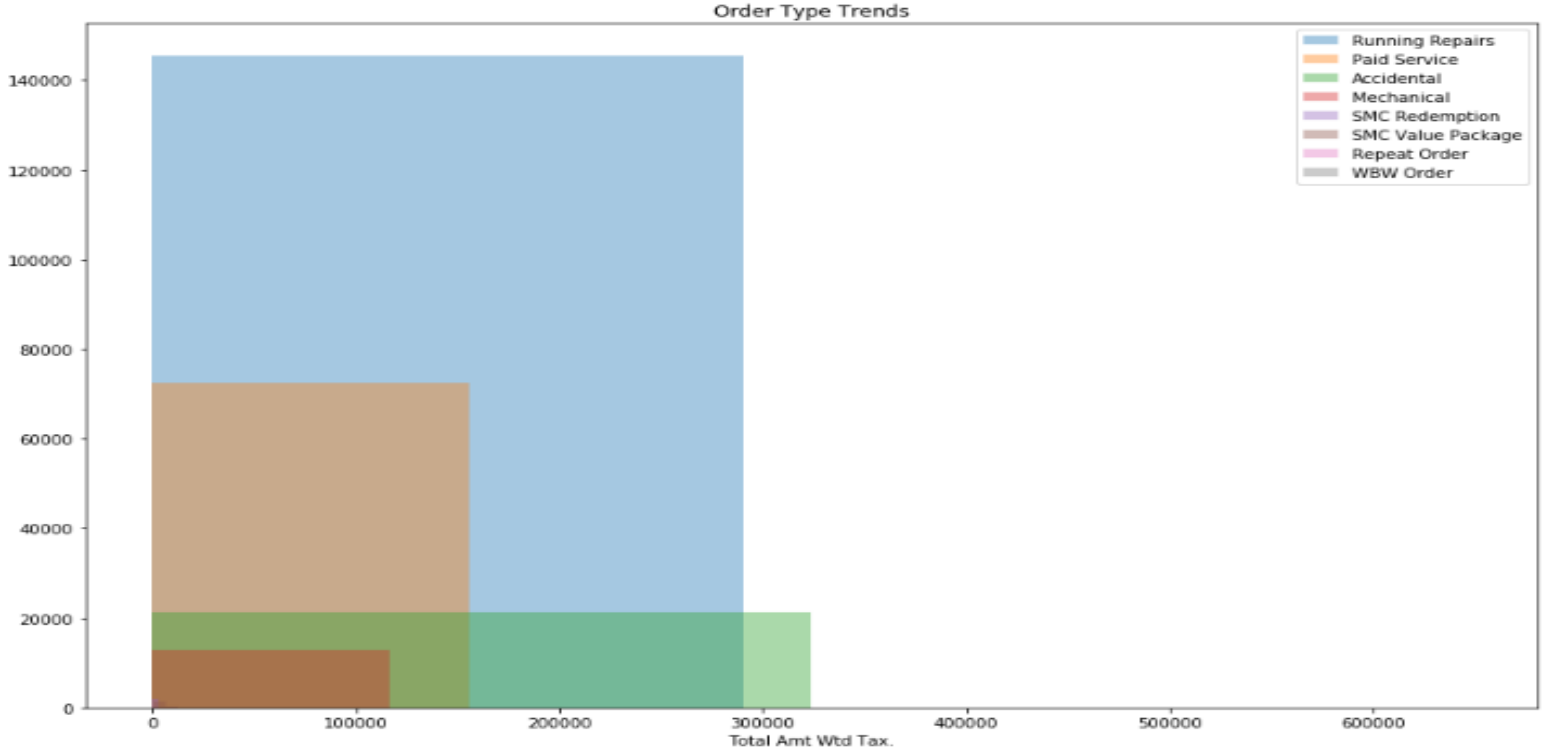




Popular Budget Friendly Car Models



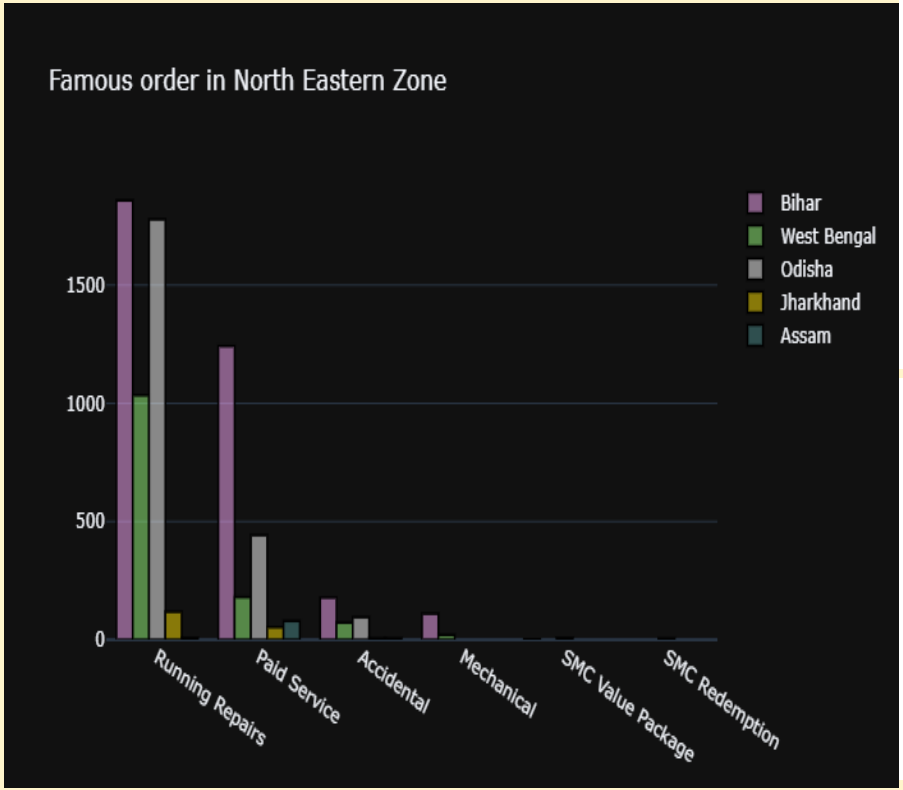
What are the most popular services?

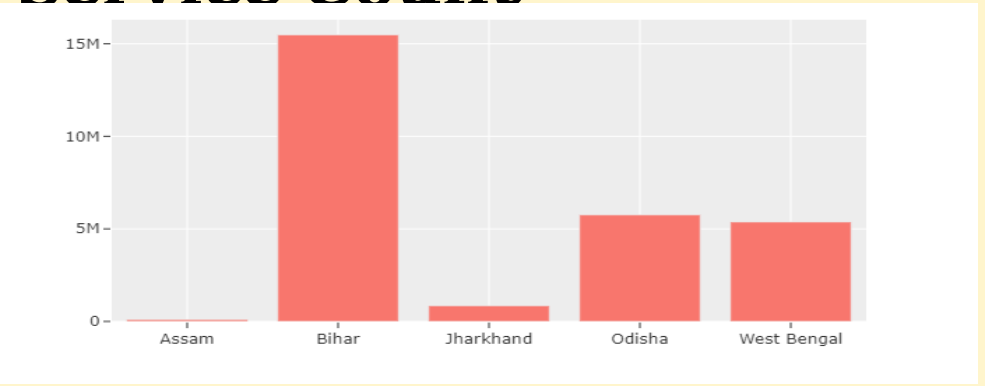


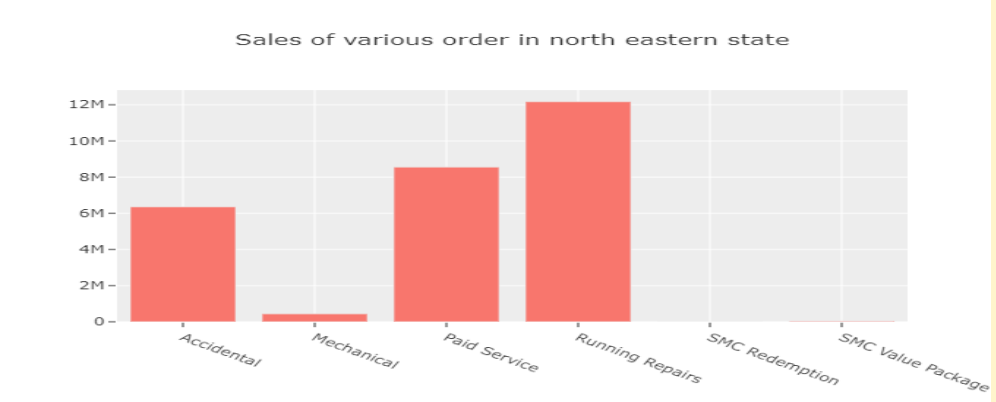
- Frequency of Running services customers is highest.
- Accidental service order generates highest revenue.

Zonal Analysis

Northern Eastern Zone Service Count

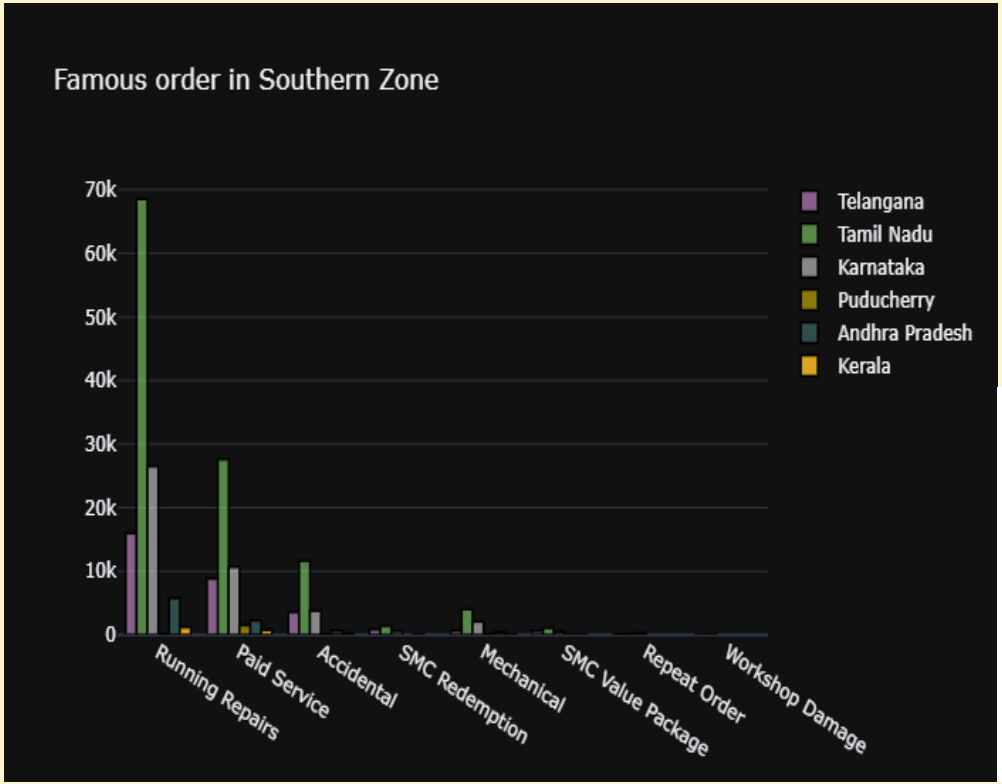


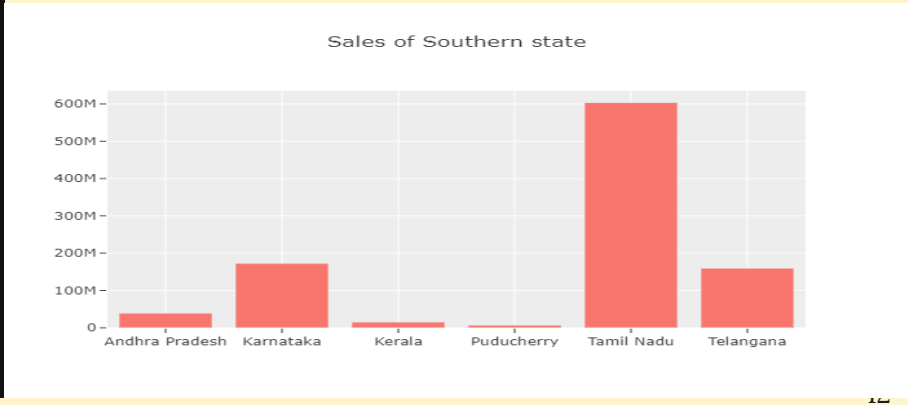




Zonal Analysis

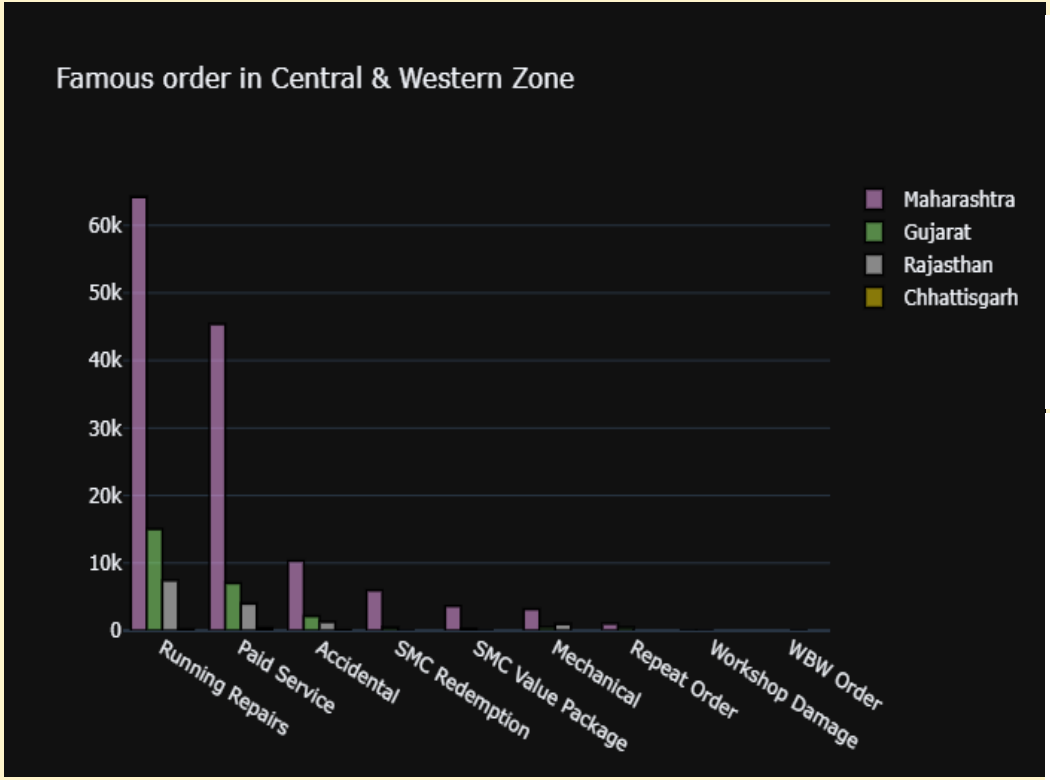
Southern Zone Service Count

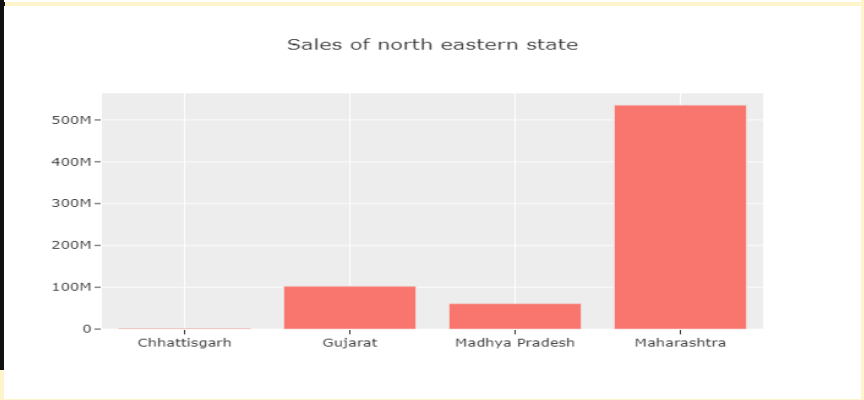




Zonal Analysis

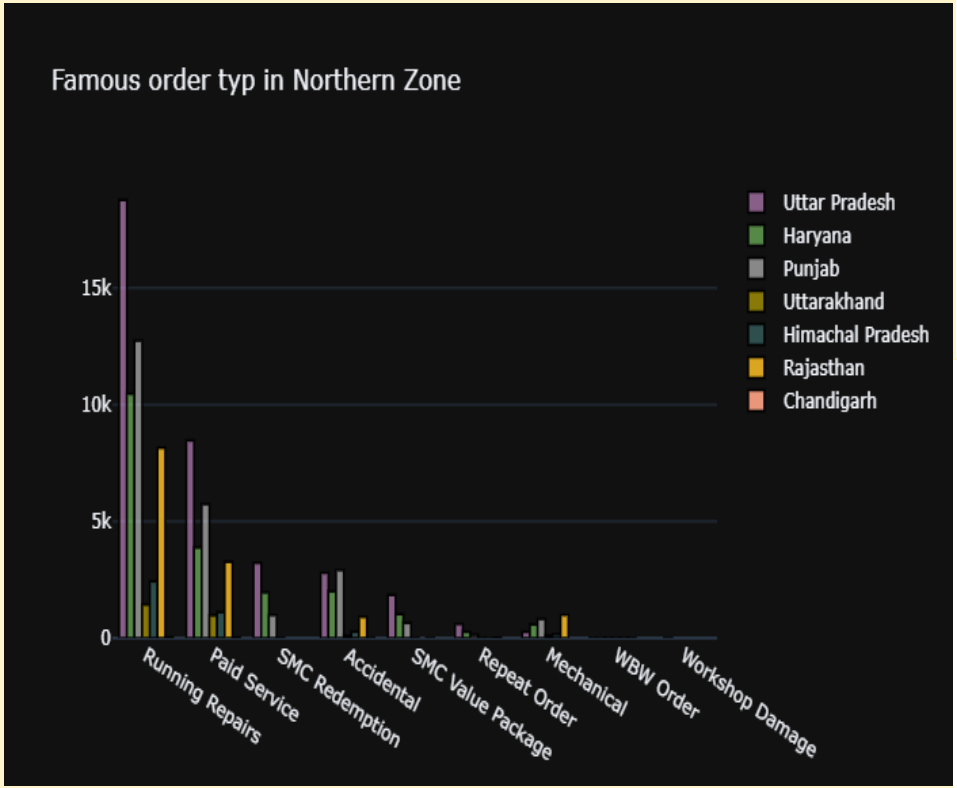
Central & Western Zone Service Count

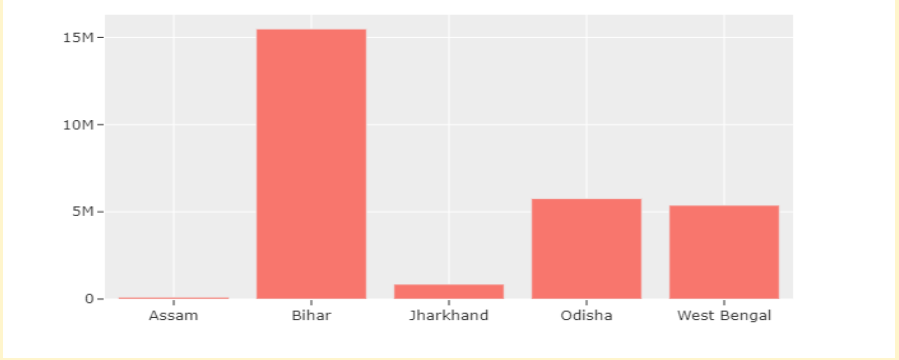


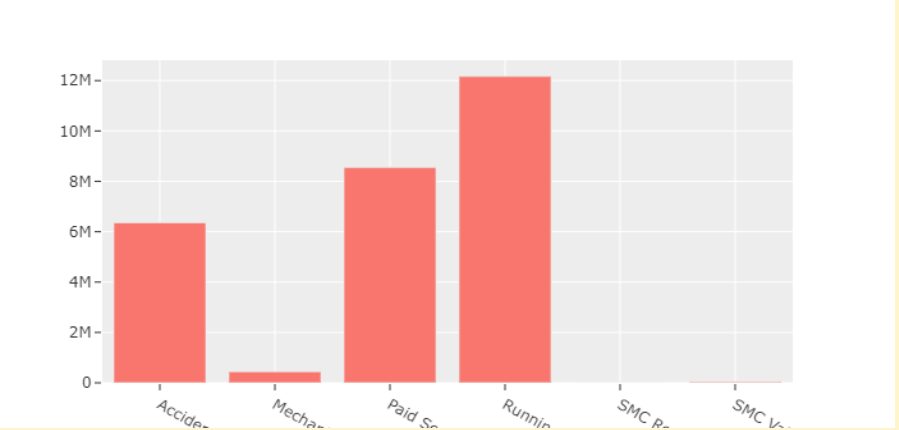


Zonal Analysis

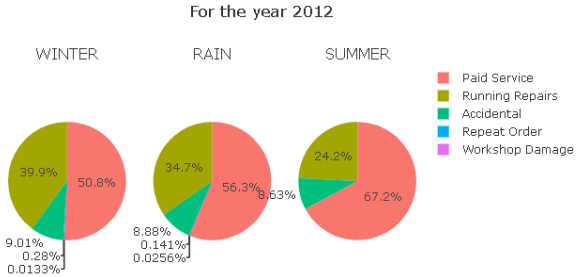
Northern Zone Service Count

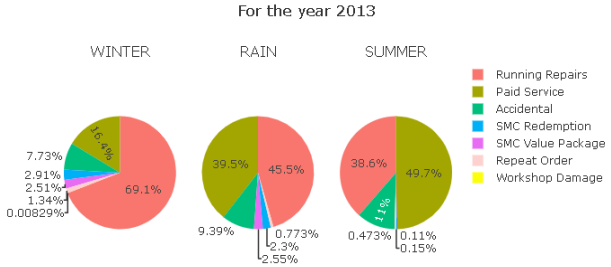


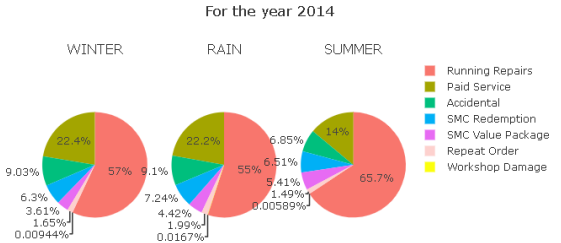


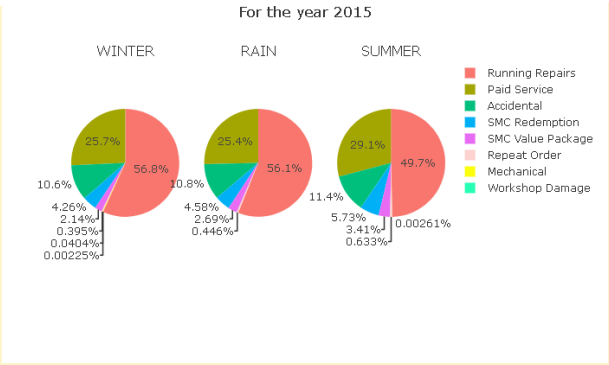


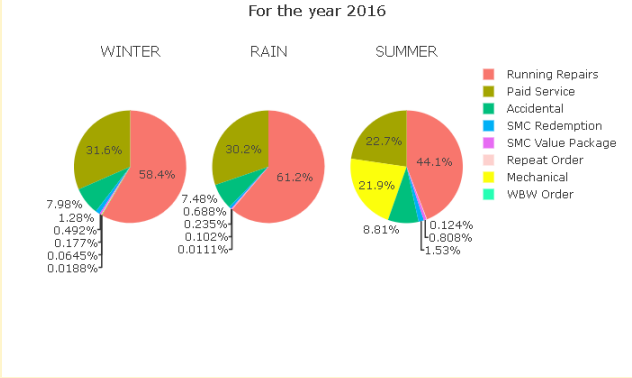
Seasonal Analysis of order types





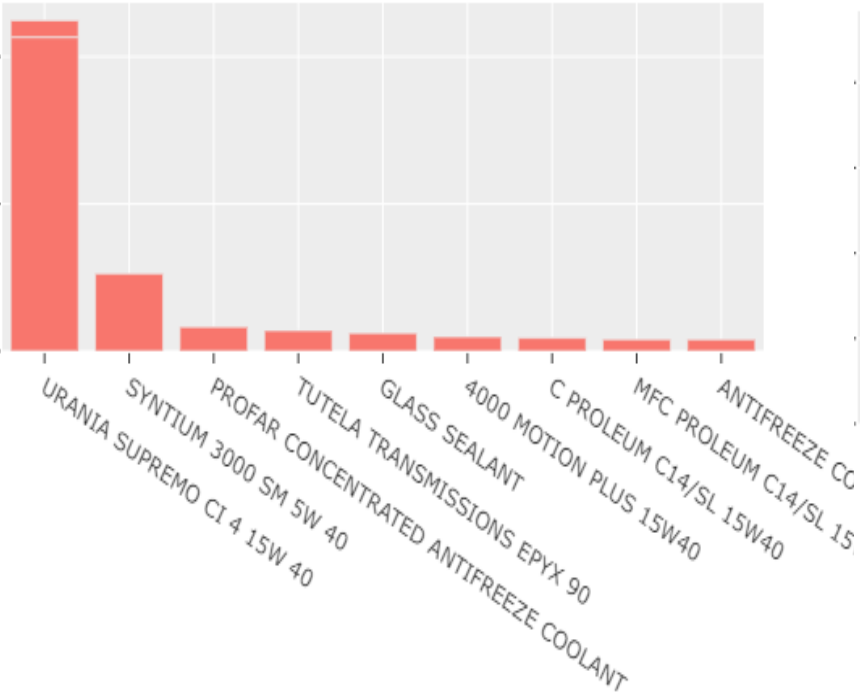




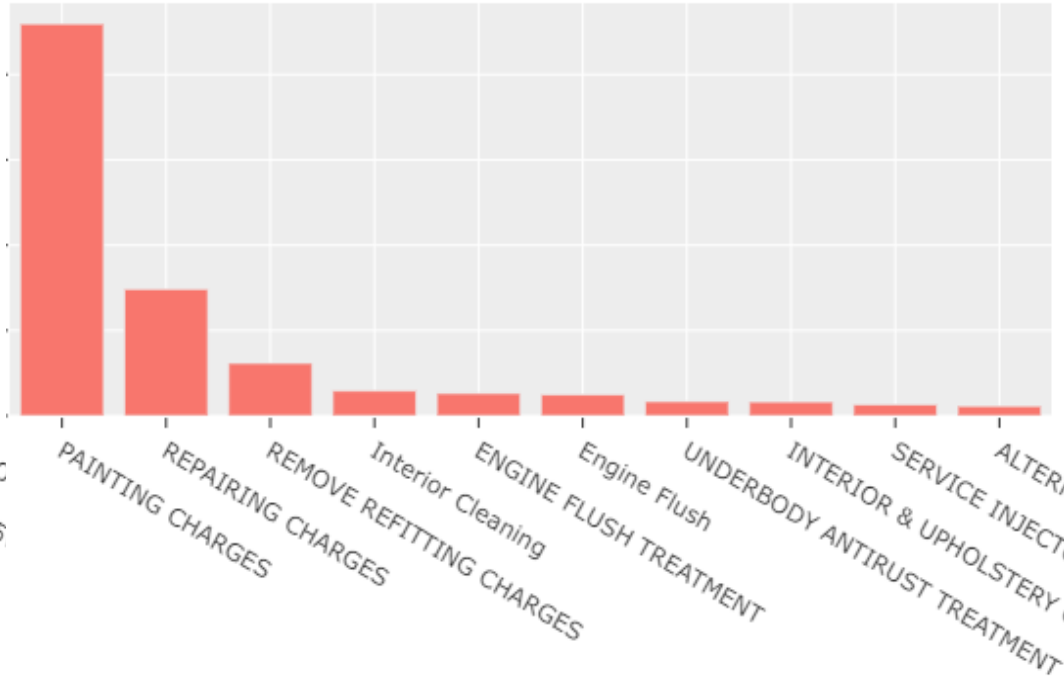


Inventory management analysis for revenue based top 10 parts used and services

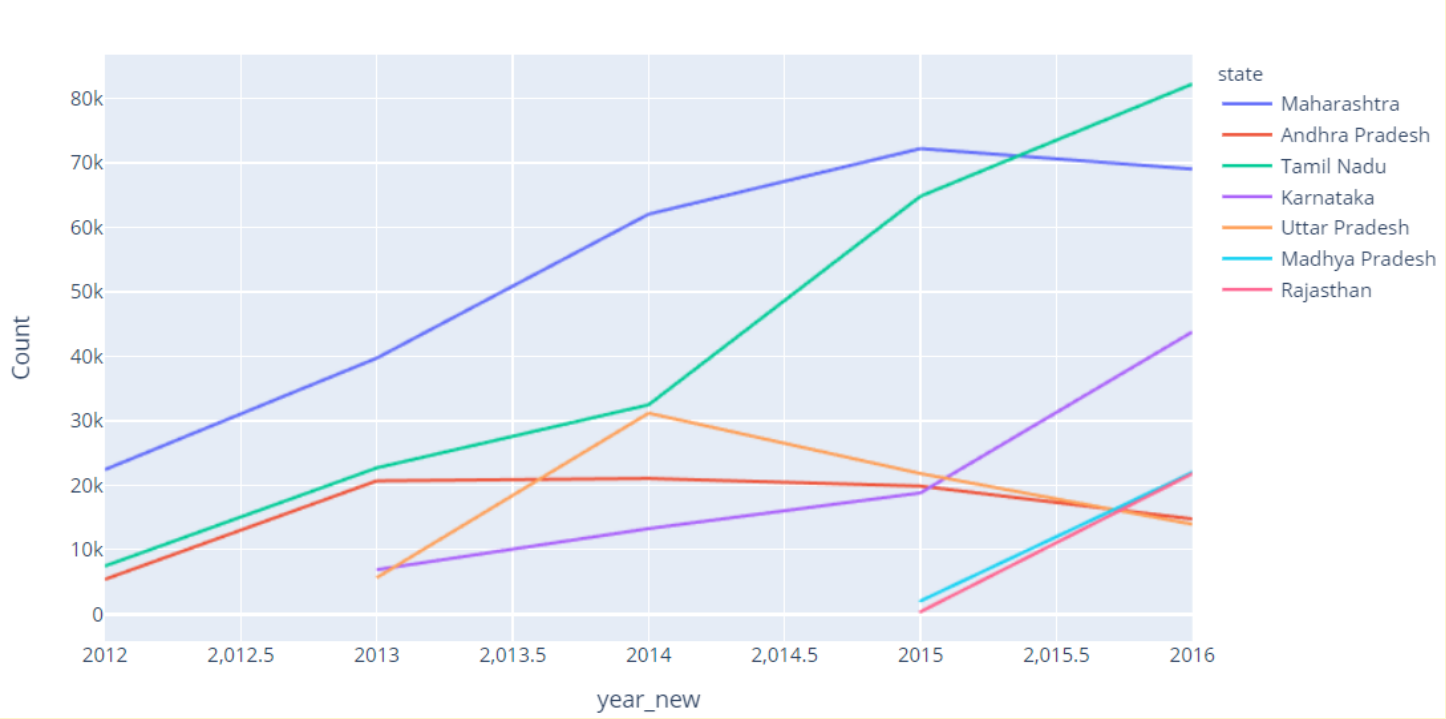
Top 10 most sold parts according to revenue



Revenue earned from external services

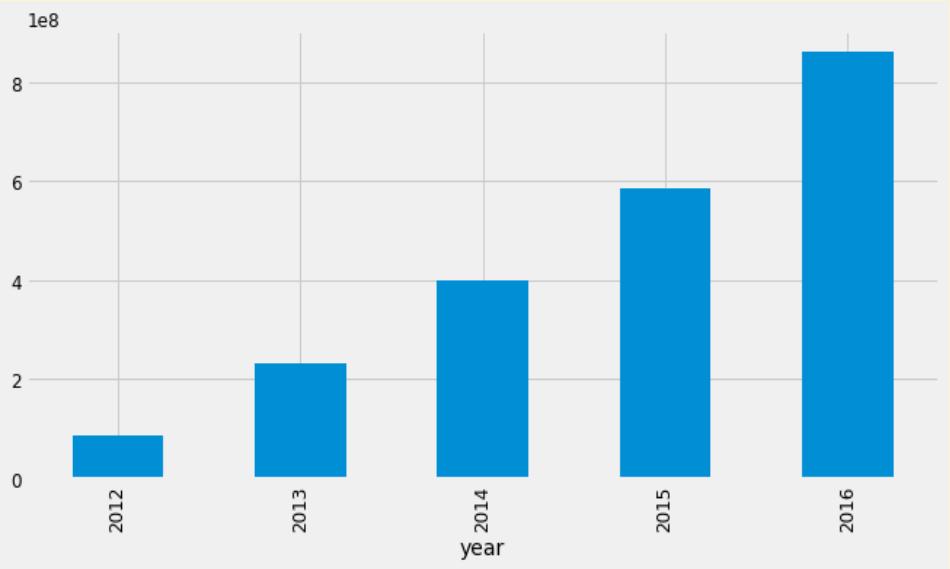


Revenue Analysis State wise Invoice count

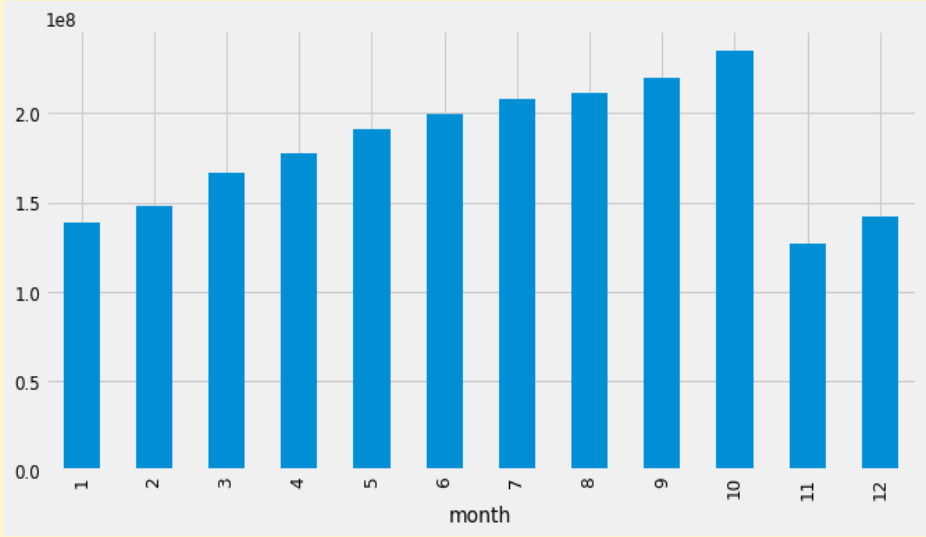


Revenue Analysis trends based on year and months

Year Wise Revenue

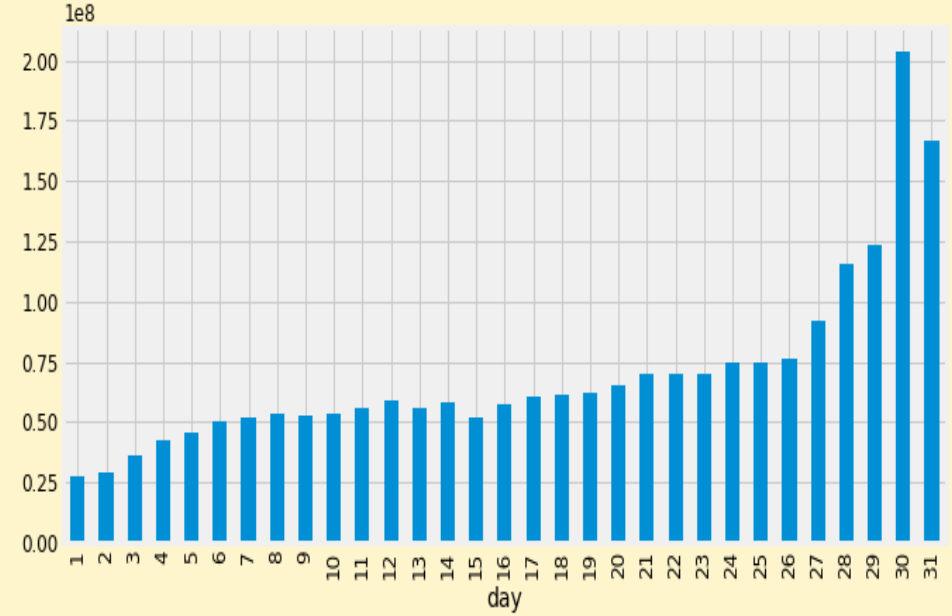


Month wise Revenue

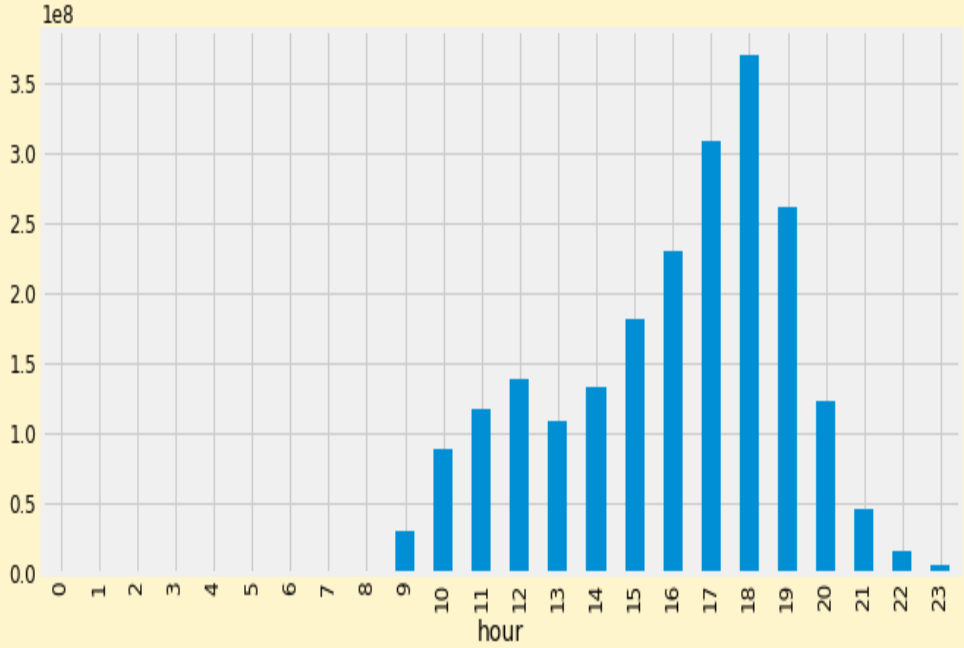


Revenue Analysis hourly and daily sales

Daily Sales

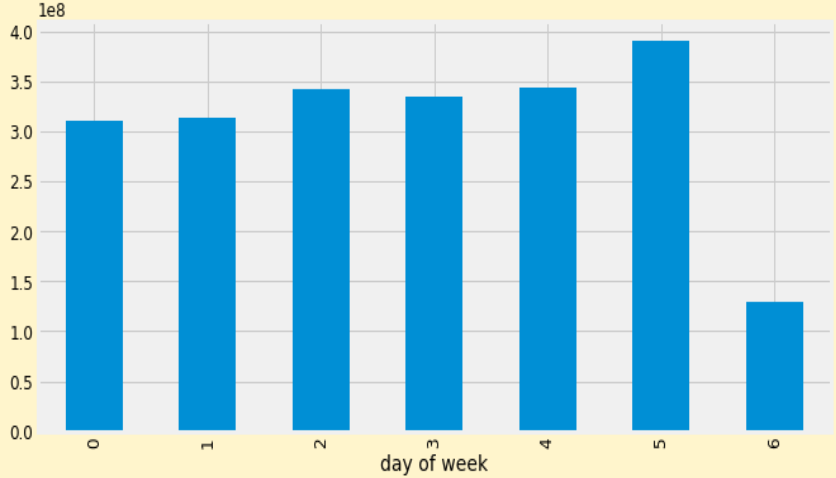


Hourly Sales

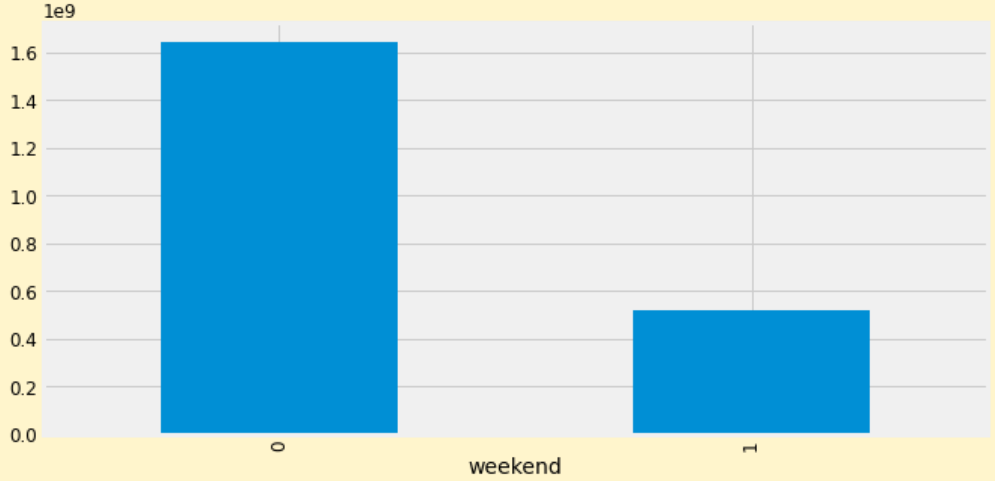


Revenue Analysis daily and weekends

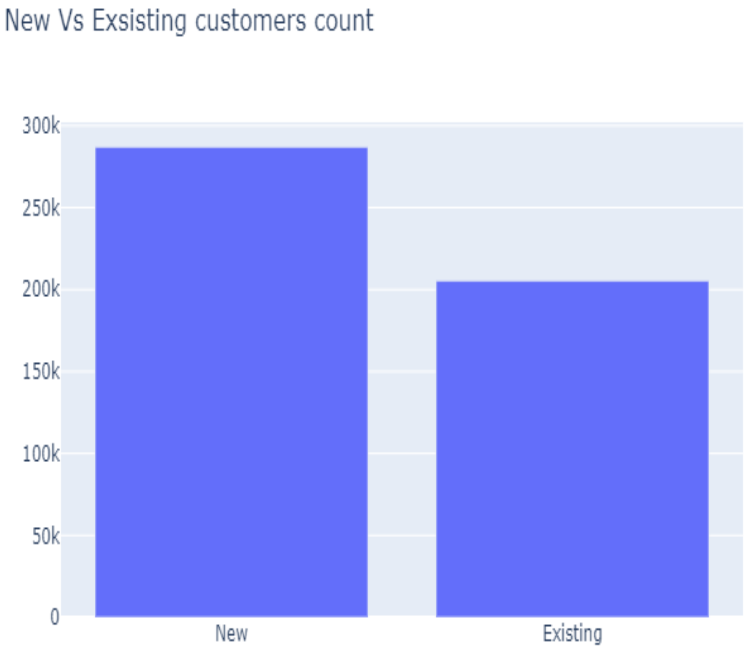
Daily Sales Analysis

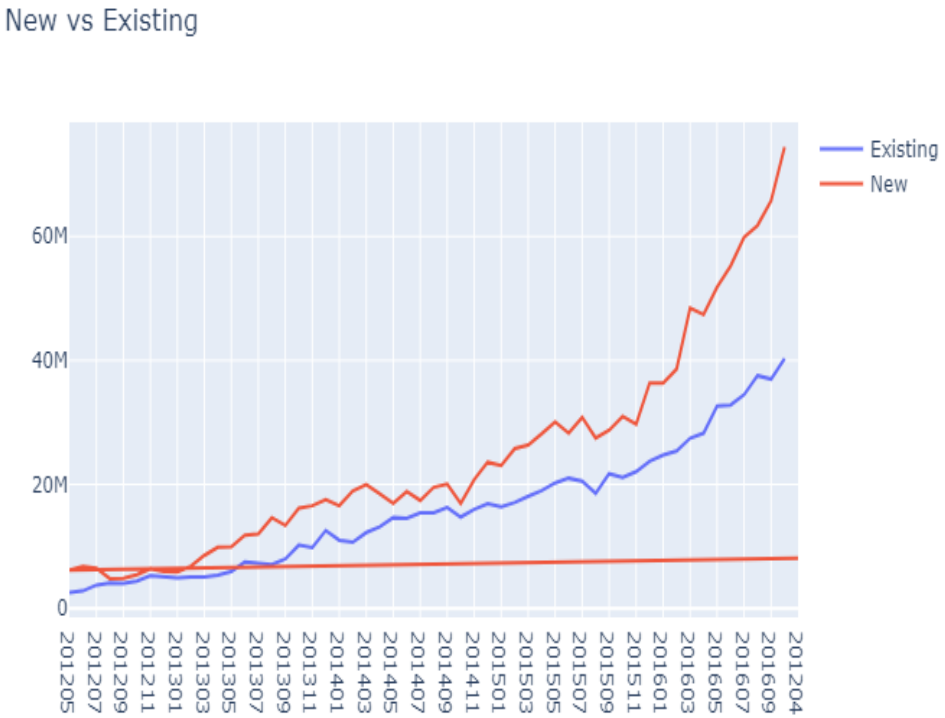


Week day vs Weekend Sales

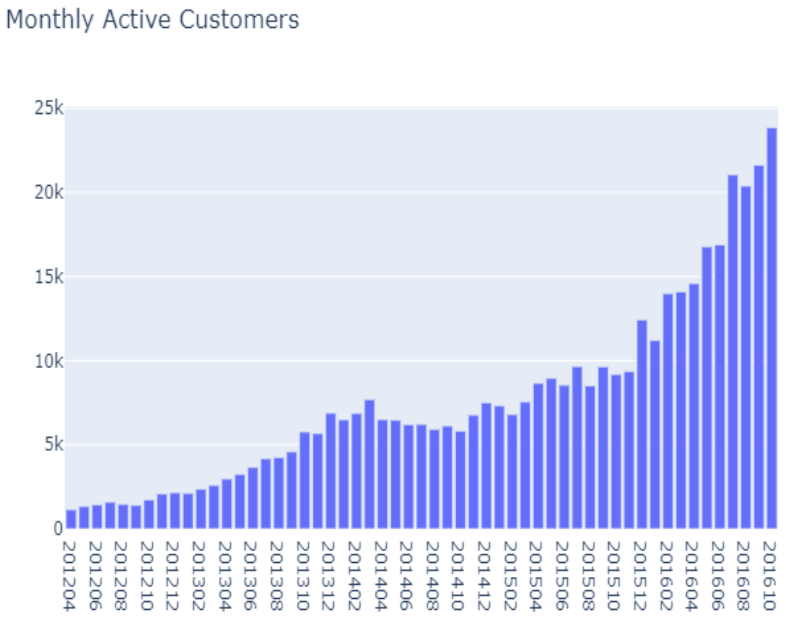


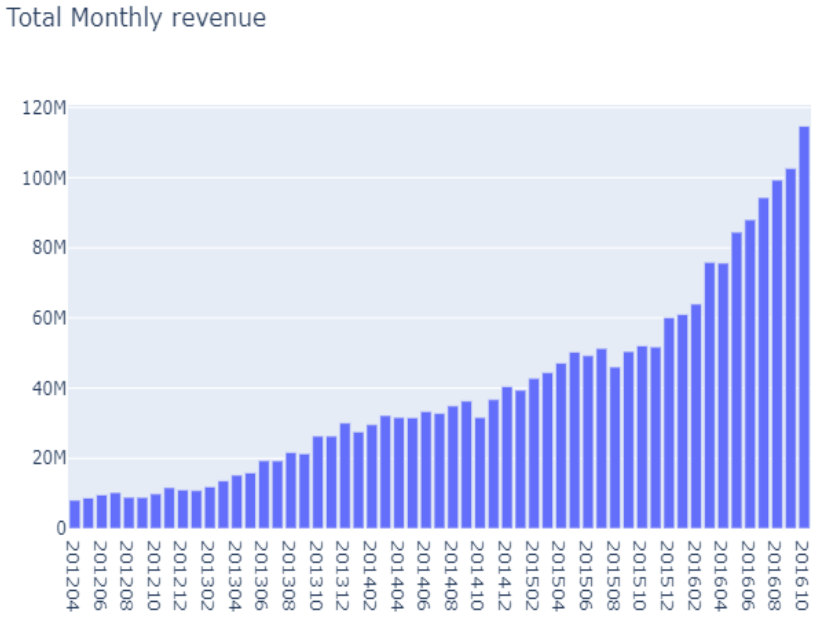
Customer segmentation Analysis Sales contribution from new vs existing customer



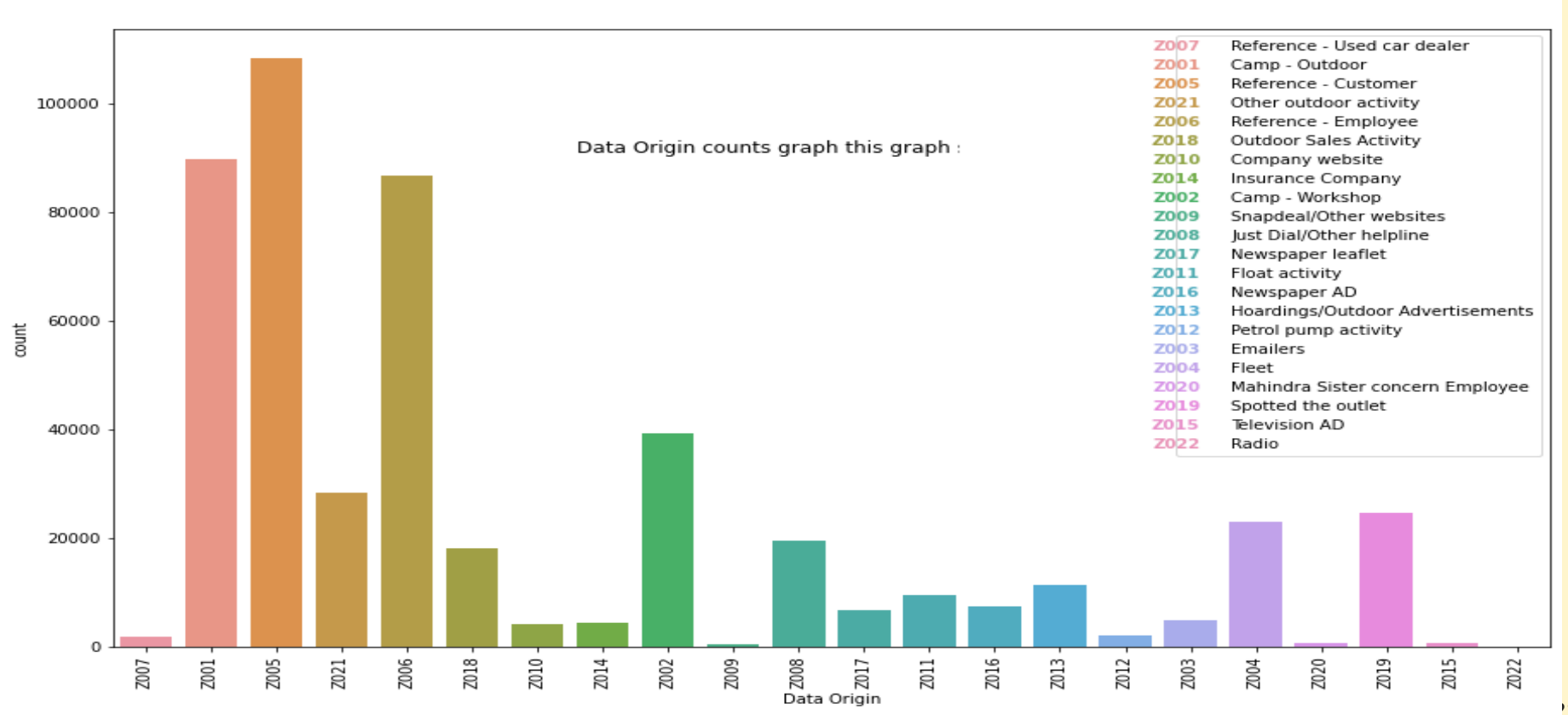


Customer segment analysis Revenue Trends and active customers





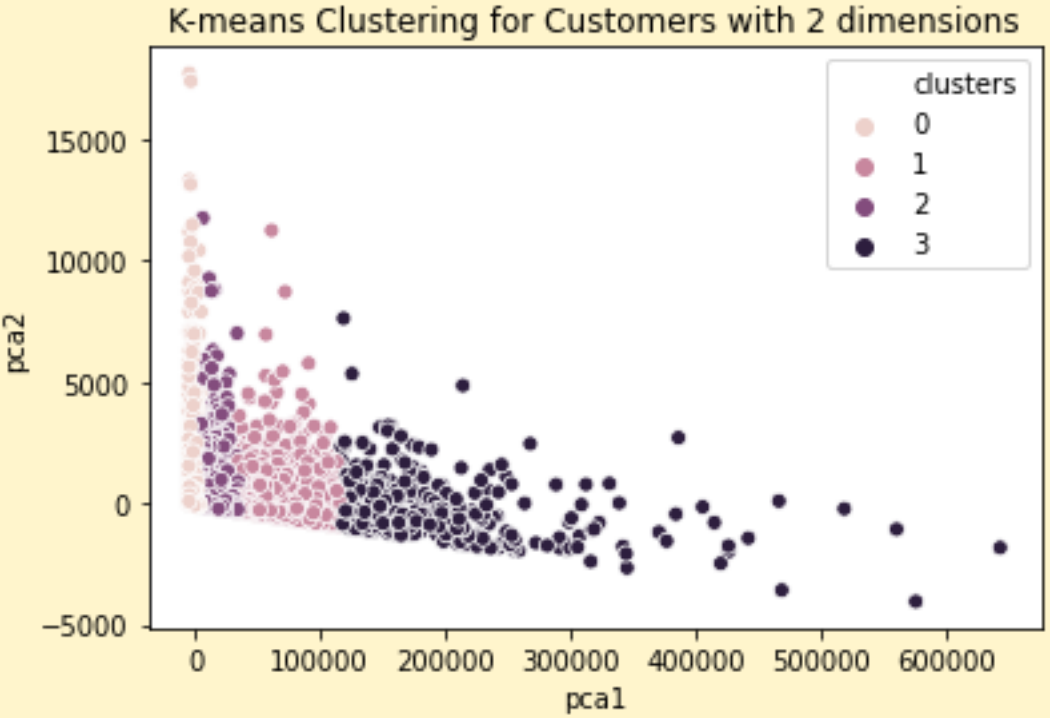
Custer reference sources



Market Segmentation

Lets Start with clustering of data

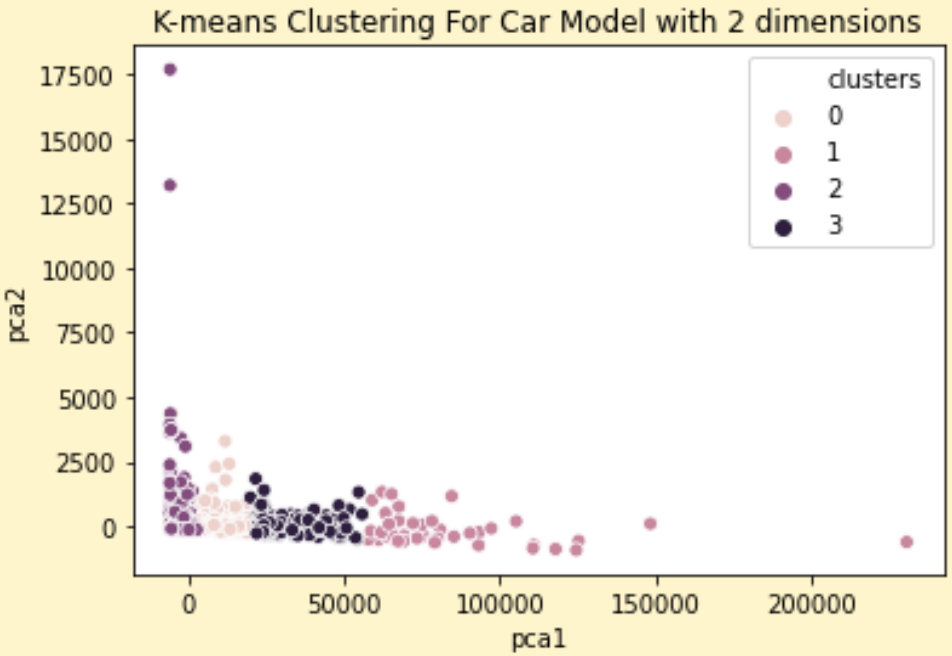
Customer clustering based on revenue



clusters	Service Hour	Total Amt Wtd Tax.
0	67.075737	2198.219239
1	562.379392	61303.102984
2	209.657500	17274.993898
3	1211.764723	181336.634630

0-low revenue earning cluster of customer
1-high revenue earning cluster of customer
2-medium revenue earning cluster of customer
3-very high revenue earning cluster of customer

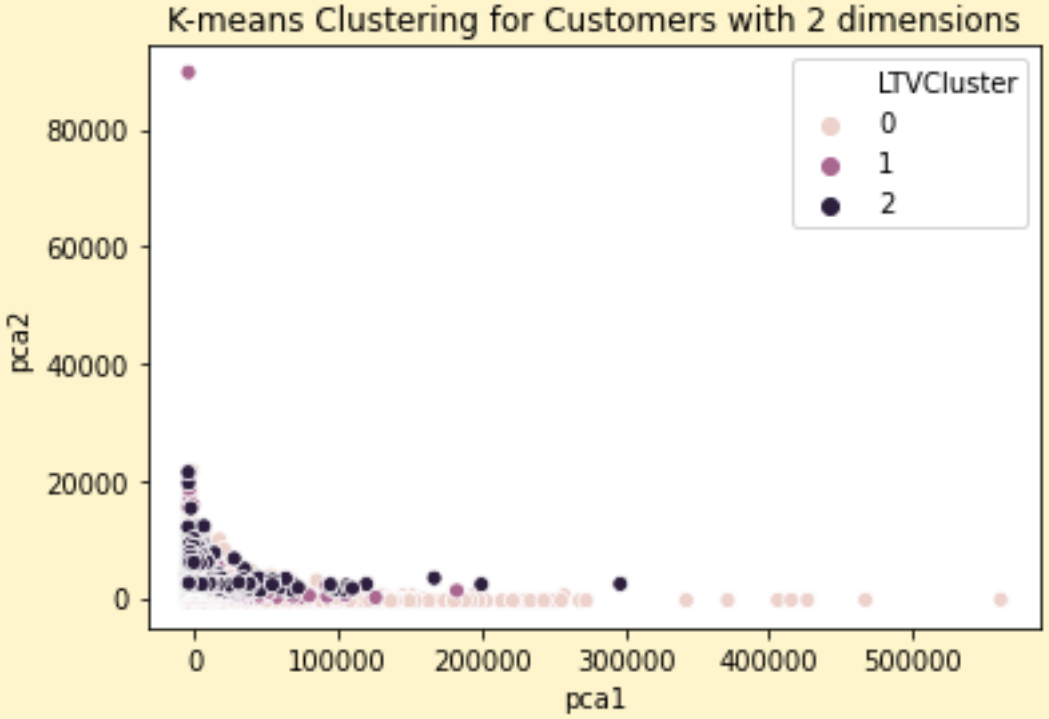
Customer clustering based on Clustering for Revenue generated from different Car Models



clusters	Car_Model	Car_Age	order_Accidental
0	916	4.411572	0.701965
1	40	5.550000	0.800000
2	5272	4.438164	0.044765
3	216	4.912037	0.865741

0 - Medium Revenue Earning Cluster of Car Models
1 - Very High Revenue Earning Cluster of Car Models
2 - Low Revenue Earning Cluster of Car Models
3 - High Revenue Earning Cluster of Car Models

Customer LTV custer



LTVCluster	LifetimeValue	Recency	Frequency
0	1487.849225	148.199911	1.506938
1	1983.534525	97.450118	2.008983
2	2034.003130	94.080370	2.060100

0-low lifetime value of a customer
1-medium lifetime value customer
2-high lifetime value of a customer

Data Preprocessing

Lets Start with preprocessing of data

Data Preprocessing

- ❑ Dropped columns
- ❑ Clean state and city names
- ❑ Treated null values in model, data origin and Partner type column
- ❑ Bucketed states in zones
- ❑ Created year, month and season features
- ❑ Calculated total service hours
- ❑ Calculated age of the cars
- ❑ Scraped car prices for bucketing models into Categories

Customer Lifetime Value

- ✓ Lifetime value of the customer is a metric that represents the total amount of a money a customer is expected to spend over the lifetime of the car.
- ✓ $CLTV = Customer\ Frequeny\ * Total\ amount\ Spent\ /\ Number\ of years$
- ✓ We have calculated Customer LTV for the year 2016.

Modelling

Lets Start with Model Building

Modelling

Models	Accuracy on train	Accuracy on test
Decision Tree Classifier	0.995	0.857
Random Forest Classifier	0.9222	0.9187
Logistic Regression	0.921	0.918
Decision Tree Classifier with Grid search	0.895	0.857

Classification Output:

Low LTV group (<1700)

Mid LTV group (1700 to 2000)

High LTV group (>2000)

Cross Valdiation

We perform cross validation based using voting classifier

Models used for estimators Decision Tree Classifier + Logistic Regression	Accuracy of Voting(Hard)	Accuracy of Voting(Soft)
Train set	0.922	0.995
Test Set	0.918819	0.859333

Actionable Insights

- ✓ States like Karnataka ,Tamil Nadu and Madhya Pradesh had huge spike in customer base every year. Implement their marketing strategies in states with low customer count
- ✓ Investigate the reasons for dip in revenue of Uttar Pradesh and Telangana
- ✓ Eastern states have a significantly lower share of the revenue. This is due to low customer count from these states. More plants could be opened up in this states and marketing strategies similar to top states could help increase the customer base in eastern states.
- ✓ Majority of cars visiting the workshop are of these 4 brands – Maruti Suzuki, Hyundai, M&M and Tata Motors. Inventory should be stocked up accordingly.
- ✓ Though invoice of Accidental cases are less, they contribute the largest share to the revenue. Average service time is also the highest. To reduce the same, spare parts of top brands stated above could be stocked up in advance.
- ✓ The top marketing source has been Customer Reference. We could arrange rewards programs for customers sharing positive testimonials to increase the customer base.

Thank you