BIKE DEMAND FORECAST

BY
KEDAR WAGHOLIKAR

Roadmap

- Introduction
- Objective
- Steps for forecasting
- Result
- Key Takeaways
- Managerial Insights

Introduction

- **Data** : Demand for bikes for bike sharing problem

- Two types of **customer segments**:
  - Casual Customers : Not registered but avail bike ride on spot payment basis
  - Registered Customers : Registered and already paid in advance for membership subscription

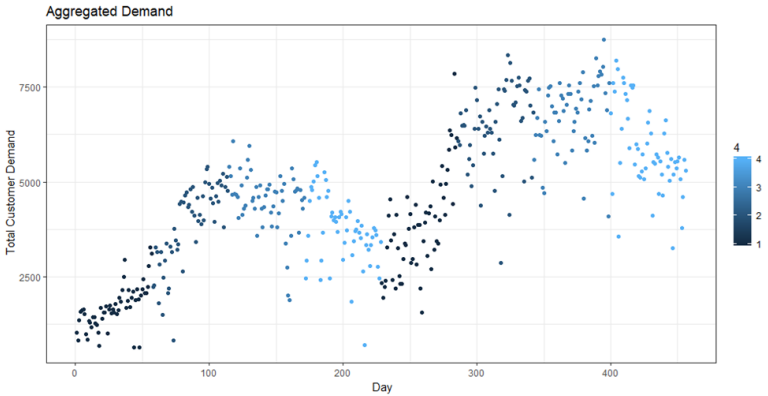
- **Variables used**: Month, Day, Season, Holiday, Temperature, Humidity, Wind speed, Total demand on the same day, Casual and Registered users demand

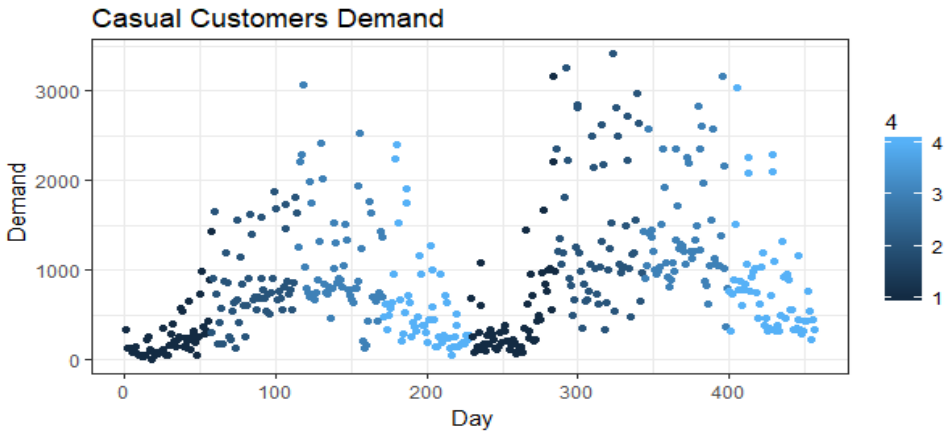
Objective

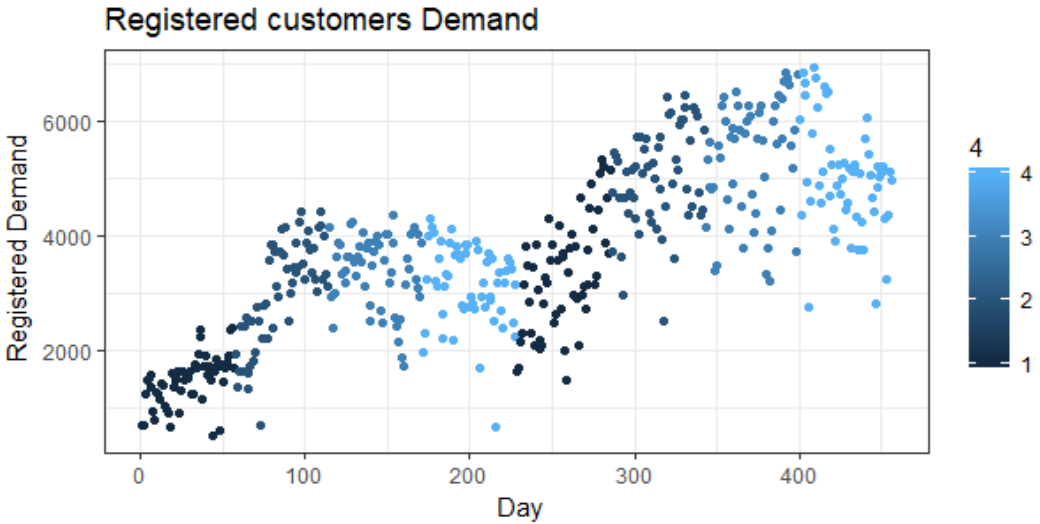
- To enhance **customer experience** and **improve the service**
- **Optimize** the number of bikes available at a given point of time
- **Satisfy customer needs** when required and maintain **profitability**
- Predict bike demand using **machine learning algorithms** based on **past data**
- Compared the **aggregated and disaggregated demand** for the bikes

Steps for forecasting- Visualization

Overall trend for Demand

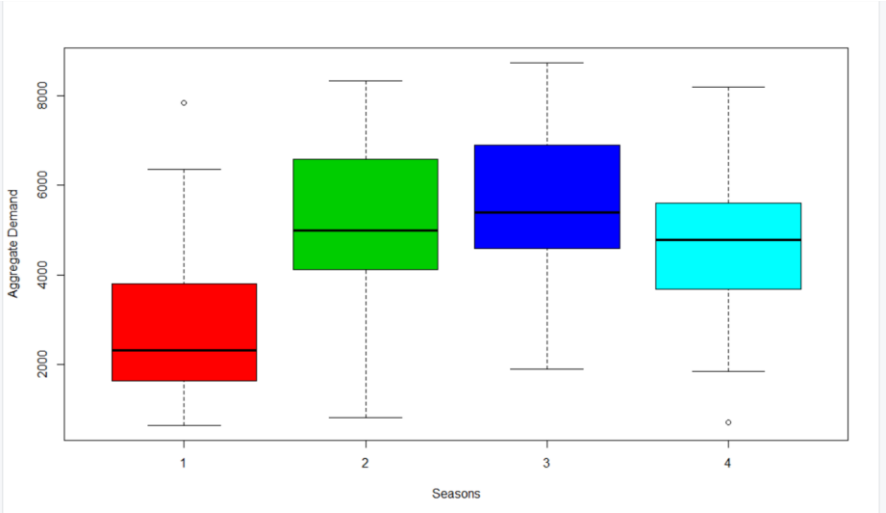



Casual Customers Demand
Demand
3000
2000
1000
0
0
100
200
300
400
Day
4
4
3
2
1


Registered customers Demand
Registered Demand
6000
4000
2000
0
100
200
300
400
Day
4
4
3
2
1

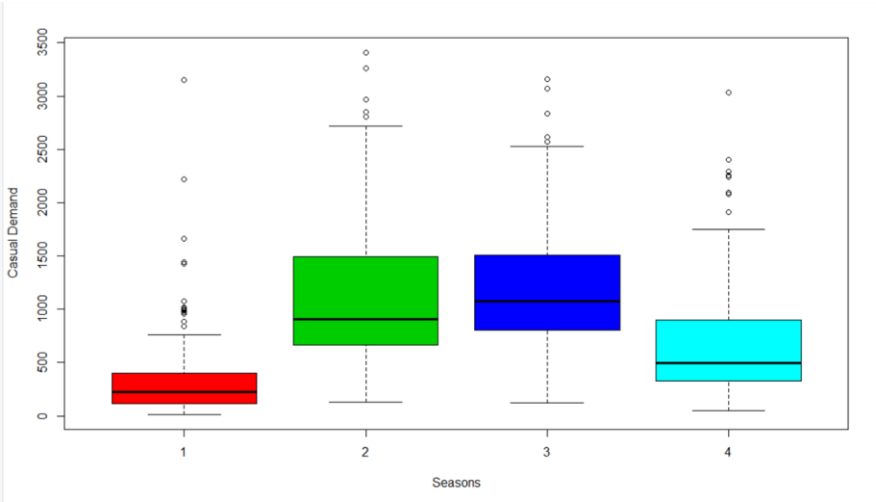
Seasonal Variations

Total Demand



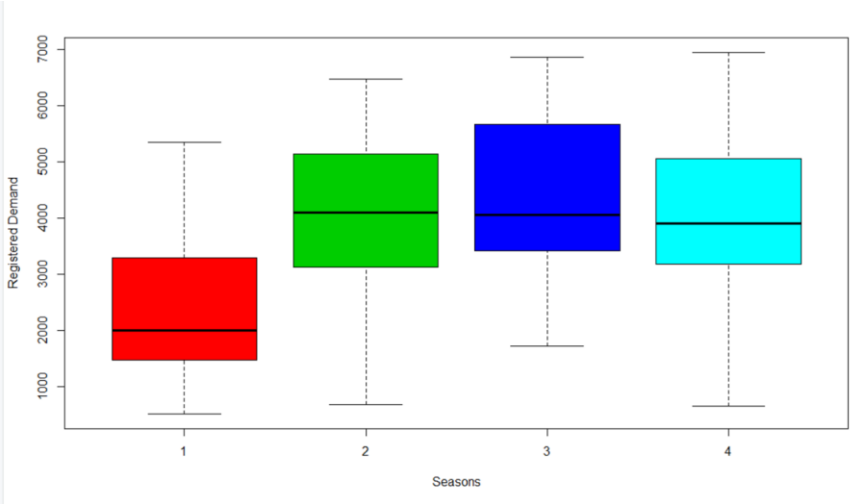
Seasonal Variations

Casual Customer Demand

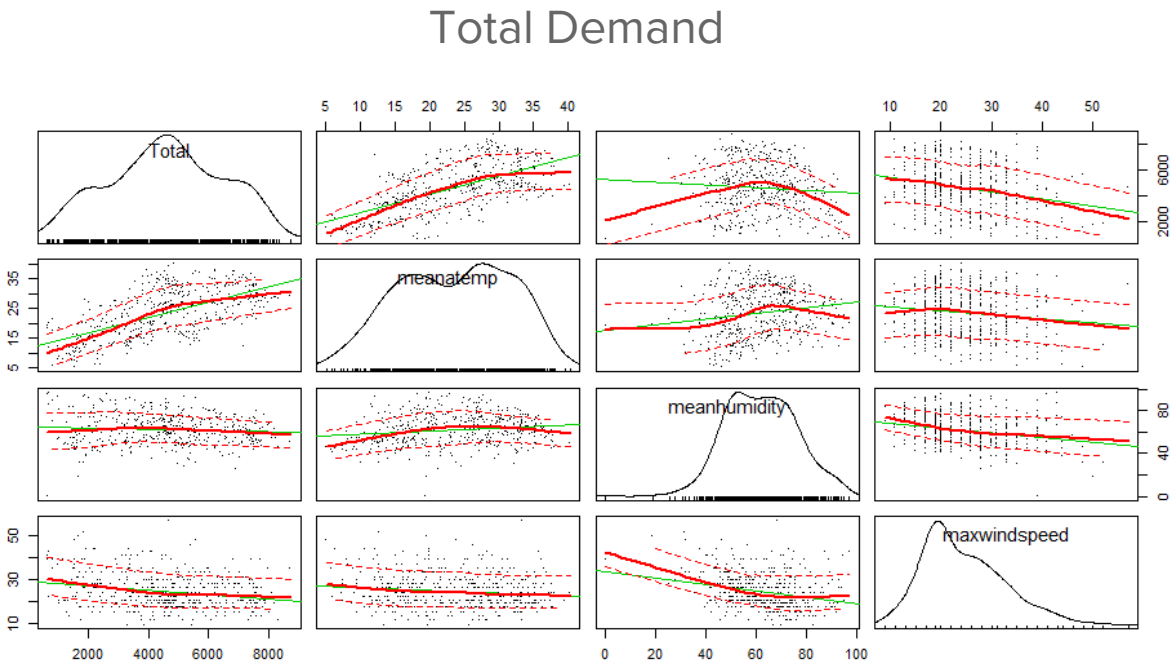


Seasonal Variations

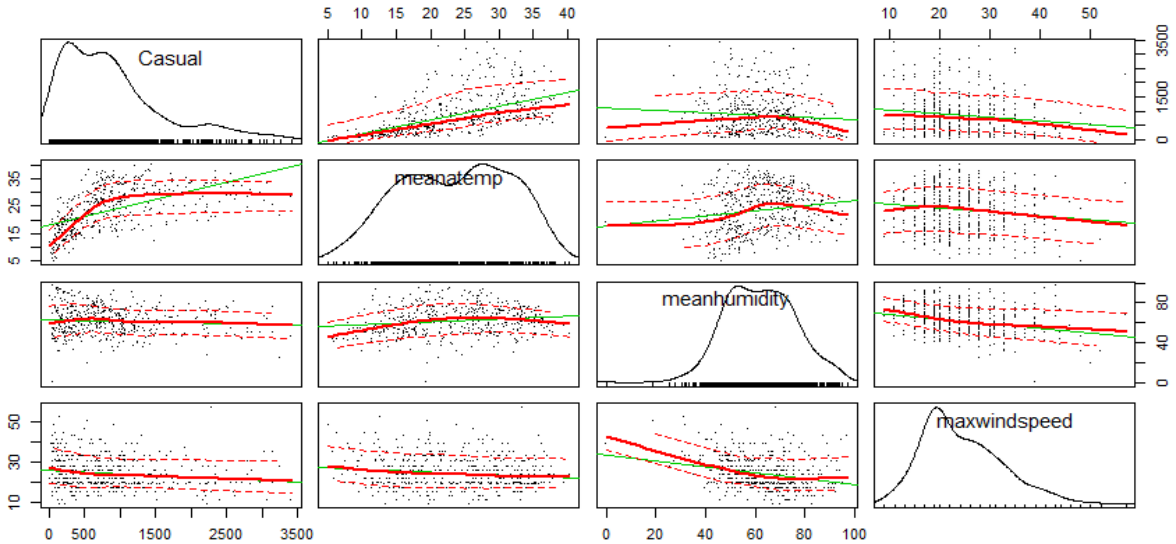
Registered Customers demand



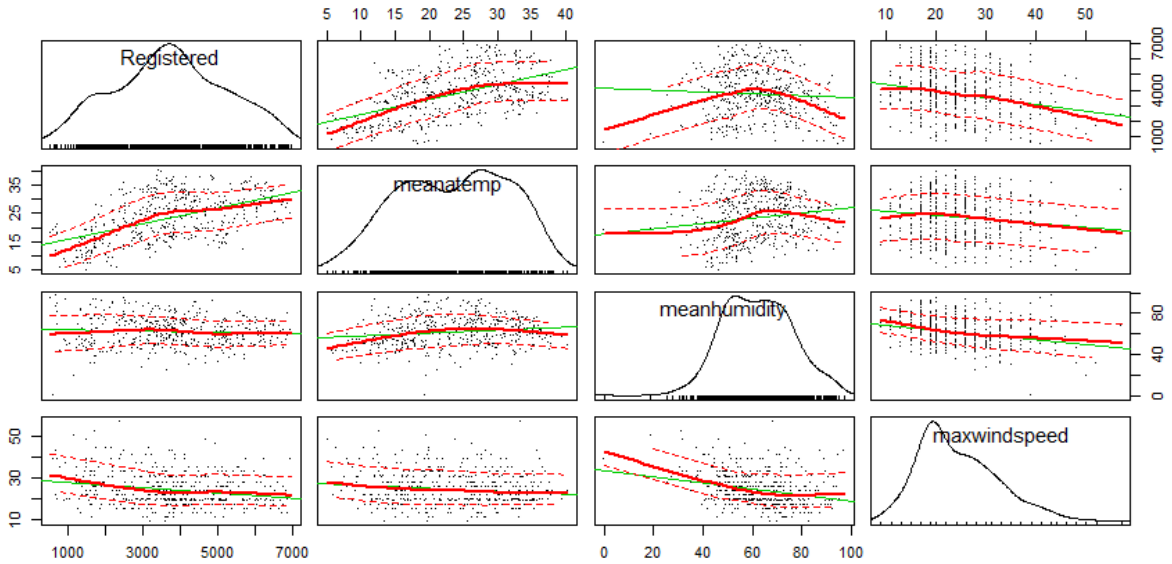
Visualization for meanatemp, meanhumidity, and maxwindspeed variables



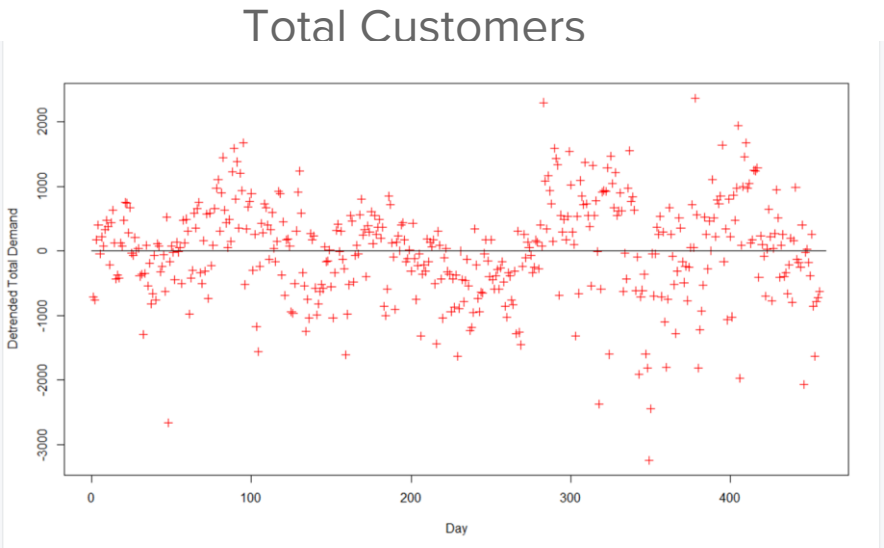
Casual Customers



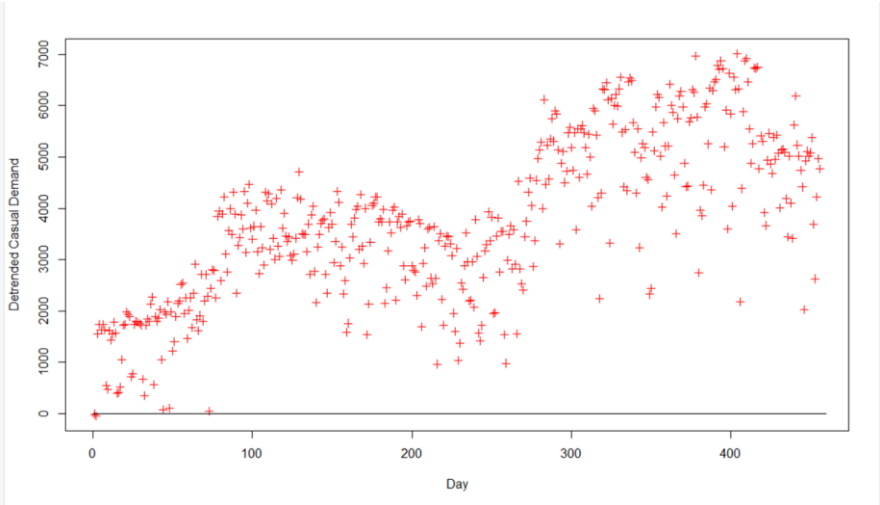
Registered Customers



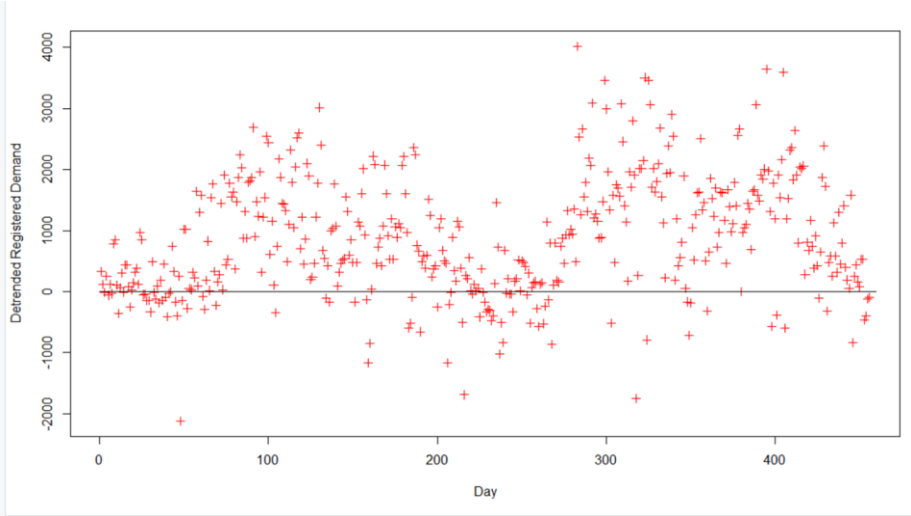
Detrended Data- after removal of seasonal variation



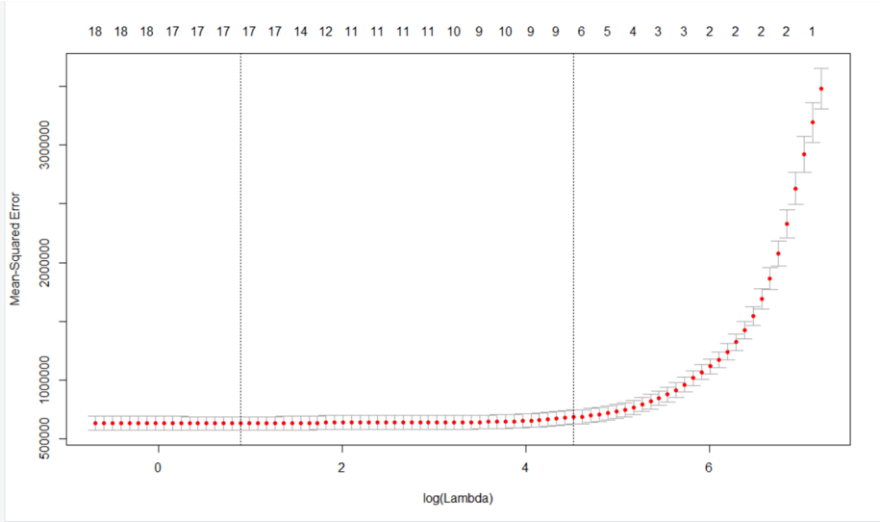
Casual Customers



Registered customers

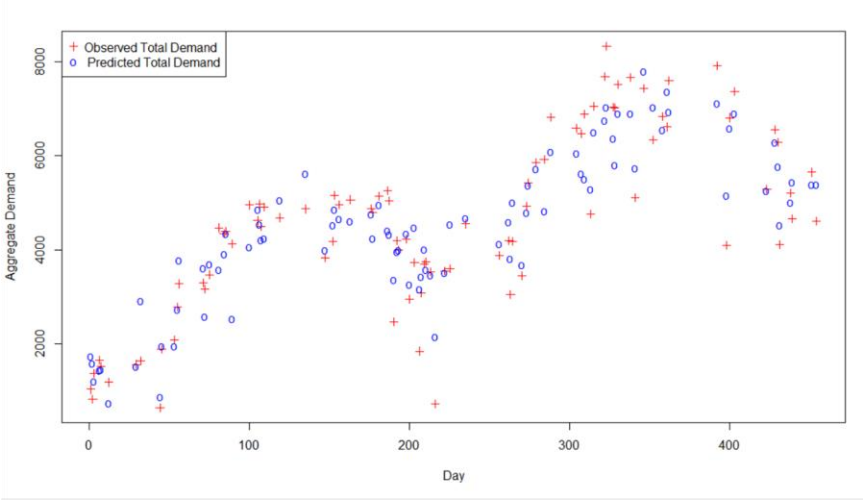


Trend of RMSPE with increasing lambda(penalizing factor)

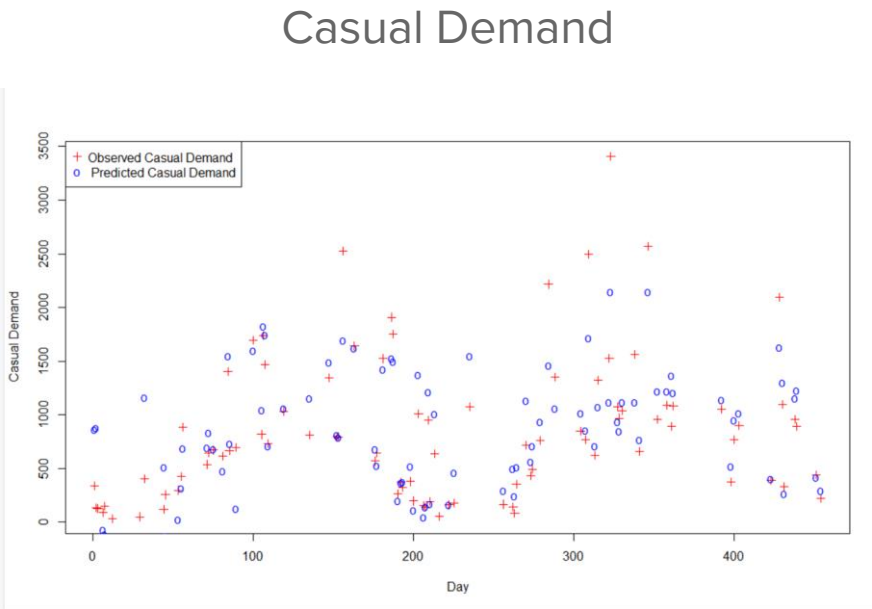


Predicted Total Demand Plot after LASSO

Total Demand

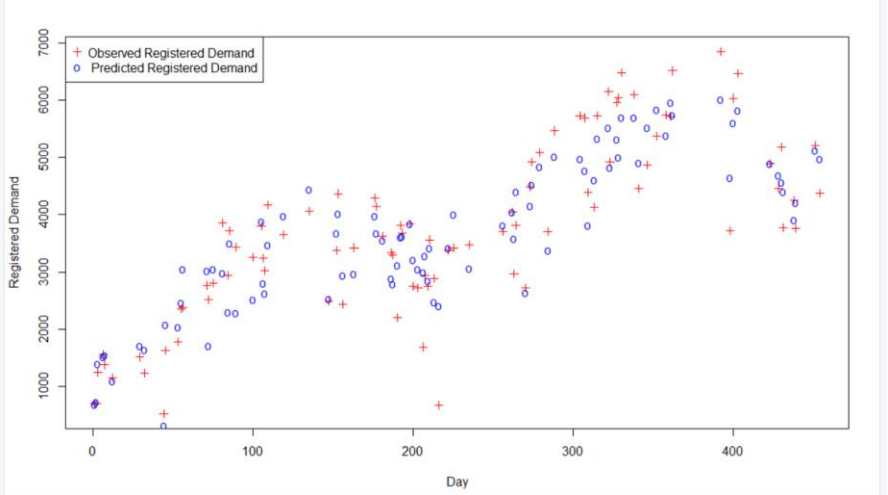


Predicted Causal Demand Plot after LASSO



Predicted Registered customers Demand Plot after LASSO

Registered Customers



Steps for forecasting

Variable Selection: Least Absolute Shrinkage and Selection Operator (LASSSO)

- Regression method for variable selection through penalization for high values of regression coefficients
- R package glmnet() is used for estimation of optimal value of lambda
- The non-zero values of beta parameters in the estimated model indicate the selected variables

Variables not selected:
Total: Month and working day
Casual: month, season, maxatemp, maxhumidity, minhumidity, minwindspeed, swindspeed
Registered: maxhumidity, minhumidity, minwindspeed

Steps for forecasting Continued ...

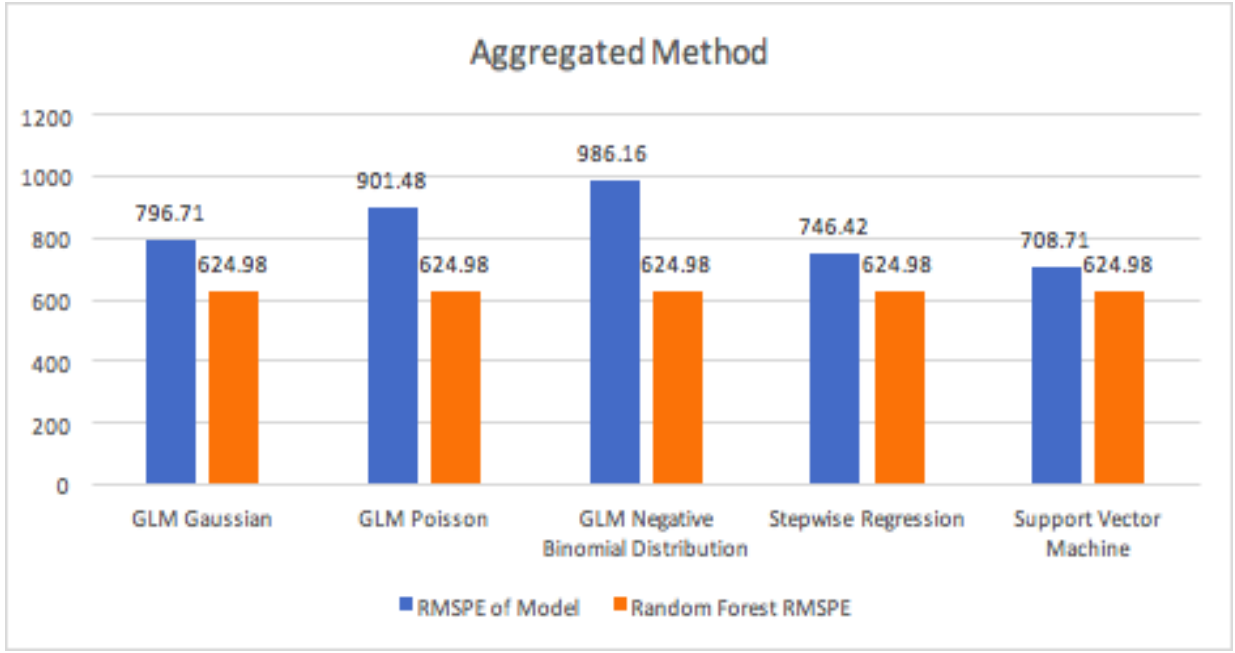
- Two Forecasting Approaches (3 models)
  - Aggregated (Total Customer Demand)
  - Disaggregated (Casual + Registered Customer Demand)

- Machine Learning Models
  - Generalised Linear
    - Gaussian Distribution
    - Poisson Distribution
    - Negative Binomial Distribution
  - Stepwise Regression (Forward Approach)
  - Random Forest
  - Support Vector Machine (SVM)

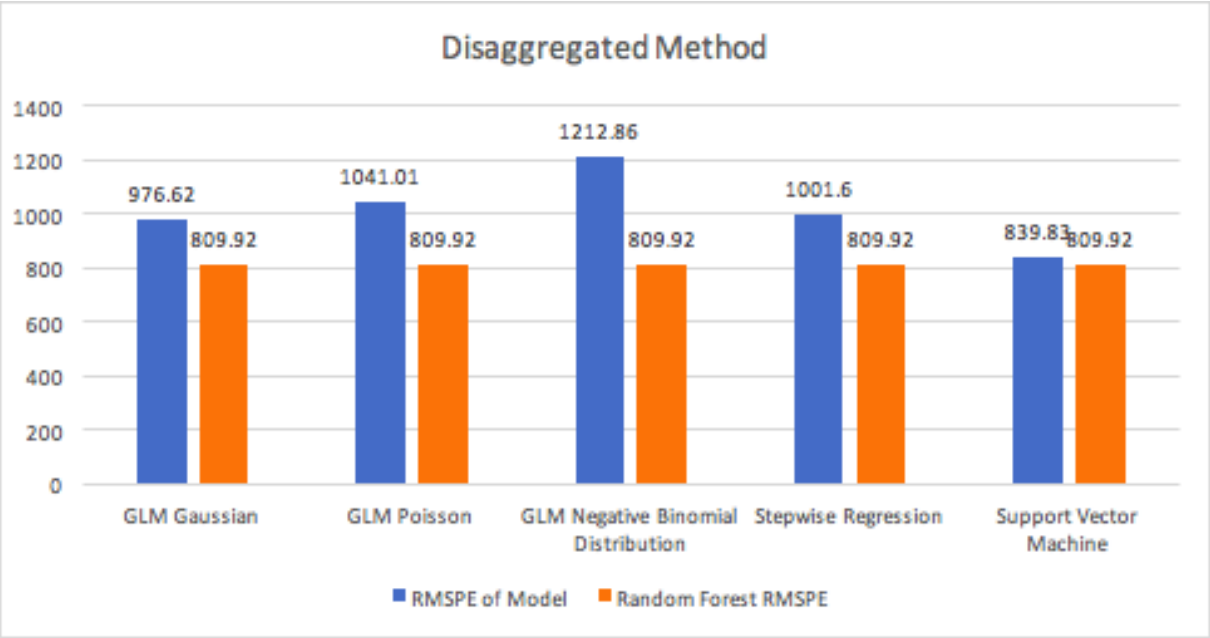
- Train(80%) and Test(20%) the models for 1000 randomly selected samples.

- Calculate Mean RMSPE from the RMSPE per model per algorithm per sample.

Results - Aggregated Method



Results - Disaggregated Method



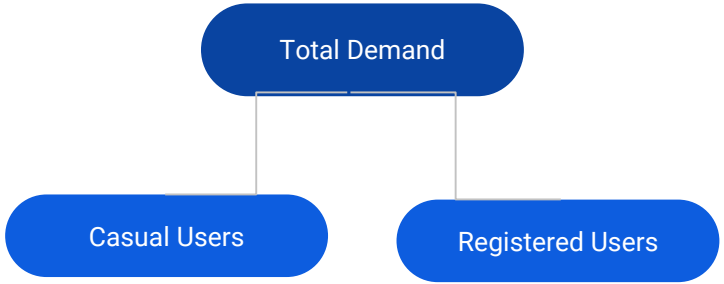
Key Takeaways

Aggregated vs Disaggregated approach

RMSPE = 624.98

More accurate results

RMSPE = 809.92

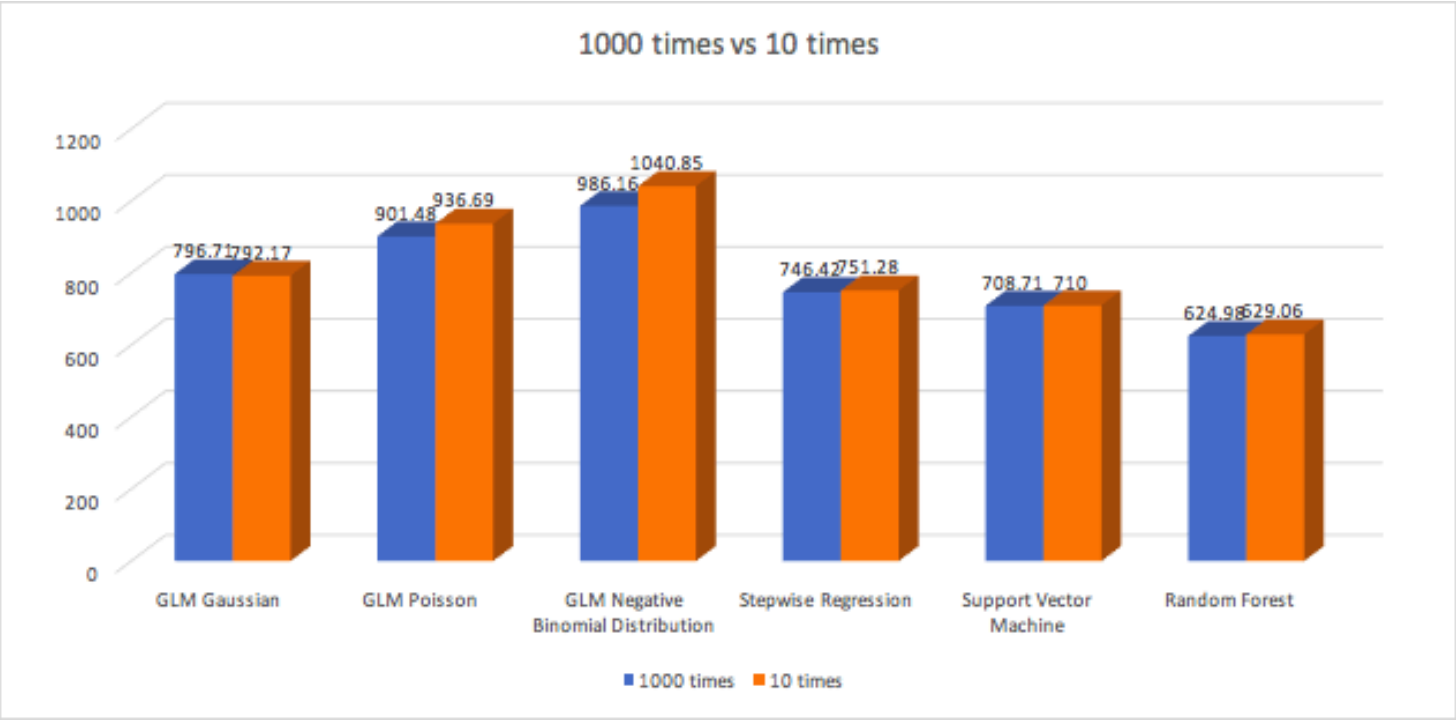


RMSPE = 304.53

RMSPE = 505.38

Scope for aggregation bias

Key Takeaways



Managerial Insights

Supply Regulatory

Capacity and Collaboration service

Effective transportation and logistics

Less investment in the Inventory

Cost effectiveness

Minimizing Safety stocks level

Form better forecast overtime

Optimal production processes

Customer Profitability

Effective customer

Customer Satisfaction

THANK YOU

Sources

https://www.spensatech.com/ap/