



# **SALES FORECASTING FOR SUPERSTORE SUPERMARKET**



# Introduction

Superstore is a very large supermarket based in the United States of America with presence in all the states. They are best known for sale of office supplies, furniture and Technology supplies such as phones. Using Historical data, the company wants to forecast their future sales so as to plan and make informed decisions about future operations, marketing, and resource allocation.

Accuracy in sale prediction helps firms to adjust their strategy accordingly, anticipate future demand and identify potential problems or opportunities. Sale forecasting is a very key task that businesses need to embrace.

# Objectives

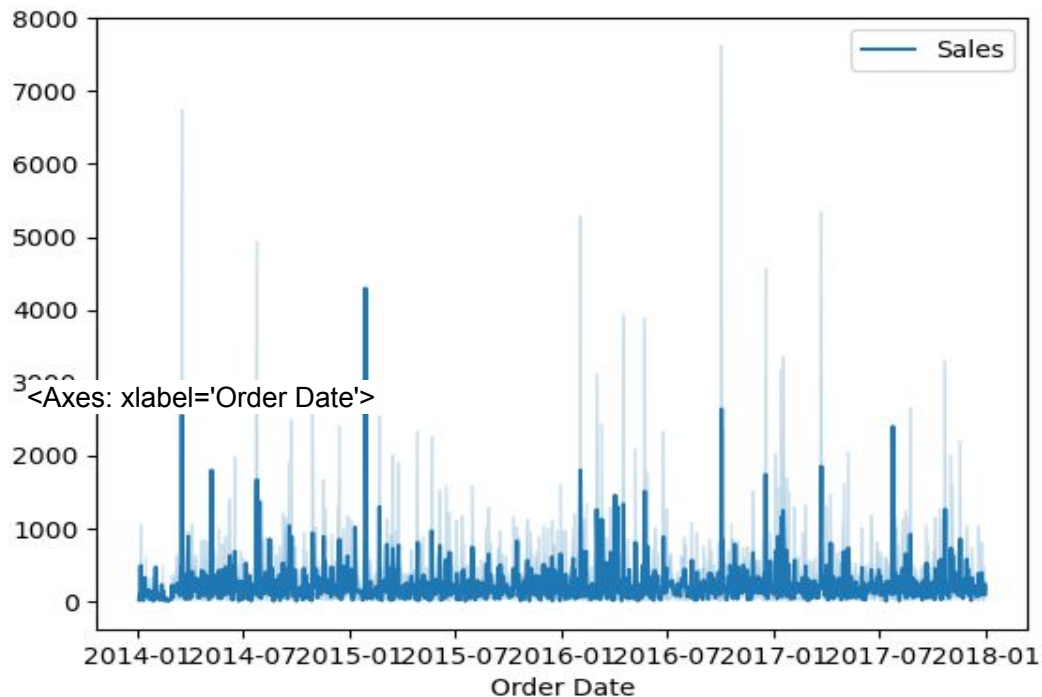
Superstore want to strategize their inventory management, logistics, production and manpower planning for the future. The business problem this project is to forecast future sales data by training supervised machine learning models on historical data.

# Data Understanding

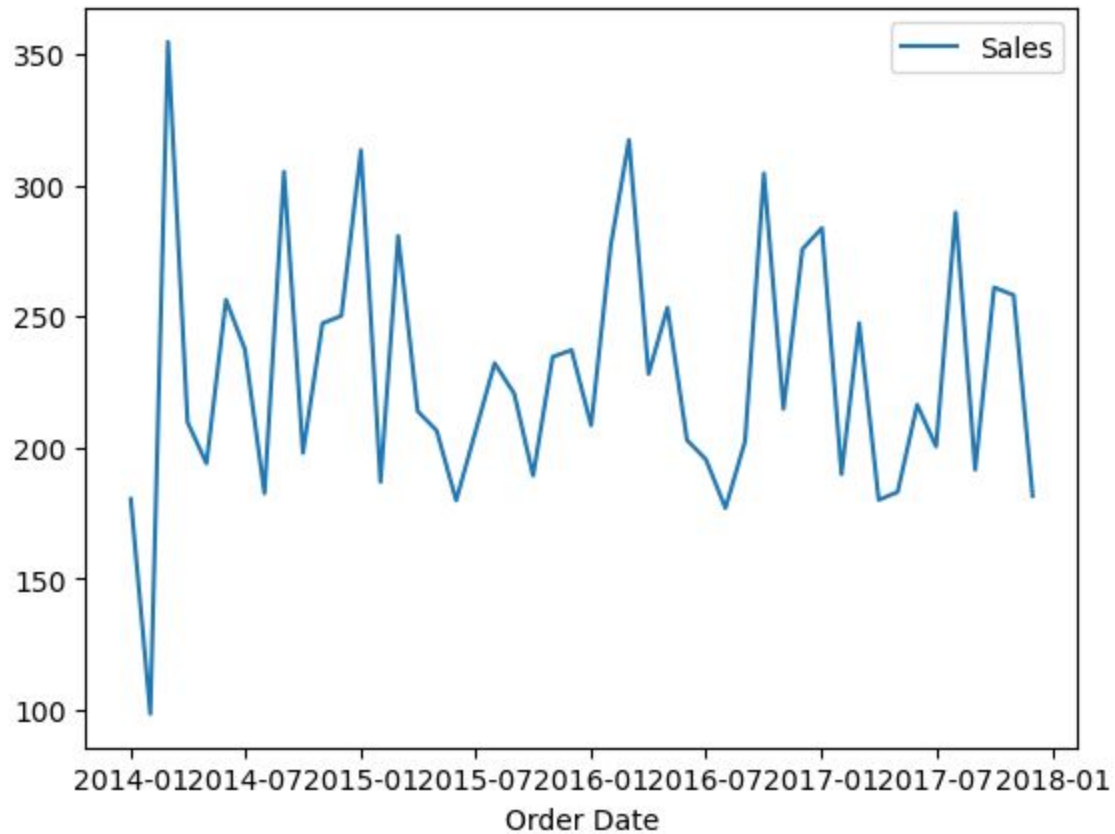
The dataset contains the historical order details of the customers of the superstore in all the stores in the United States. This is monthly data for shipped goods from the month of July in the year 2014 to May 2018.

The dataset contains 9994 rows and 21 Columns

# Exploratory Data Analysis



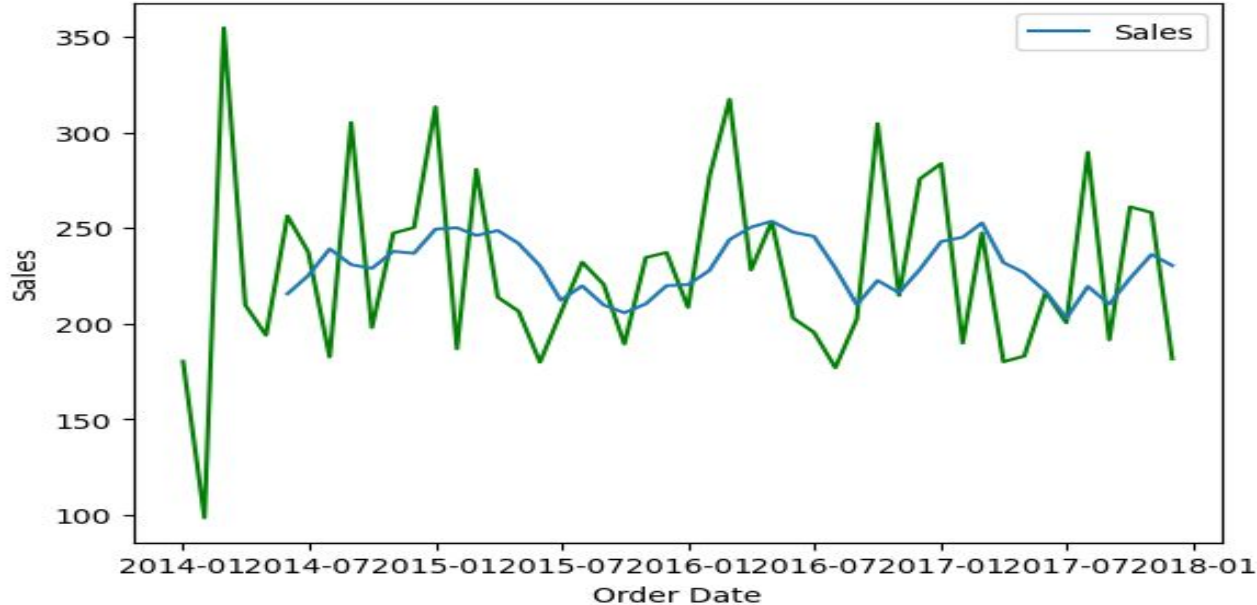
**#Visualizing the dataset using  
lineplot  
sns.lineplot(data)**



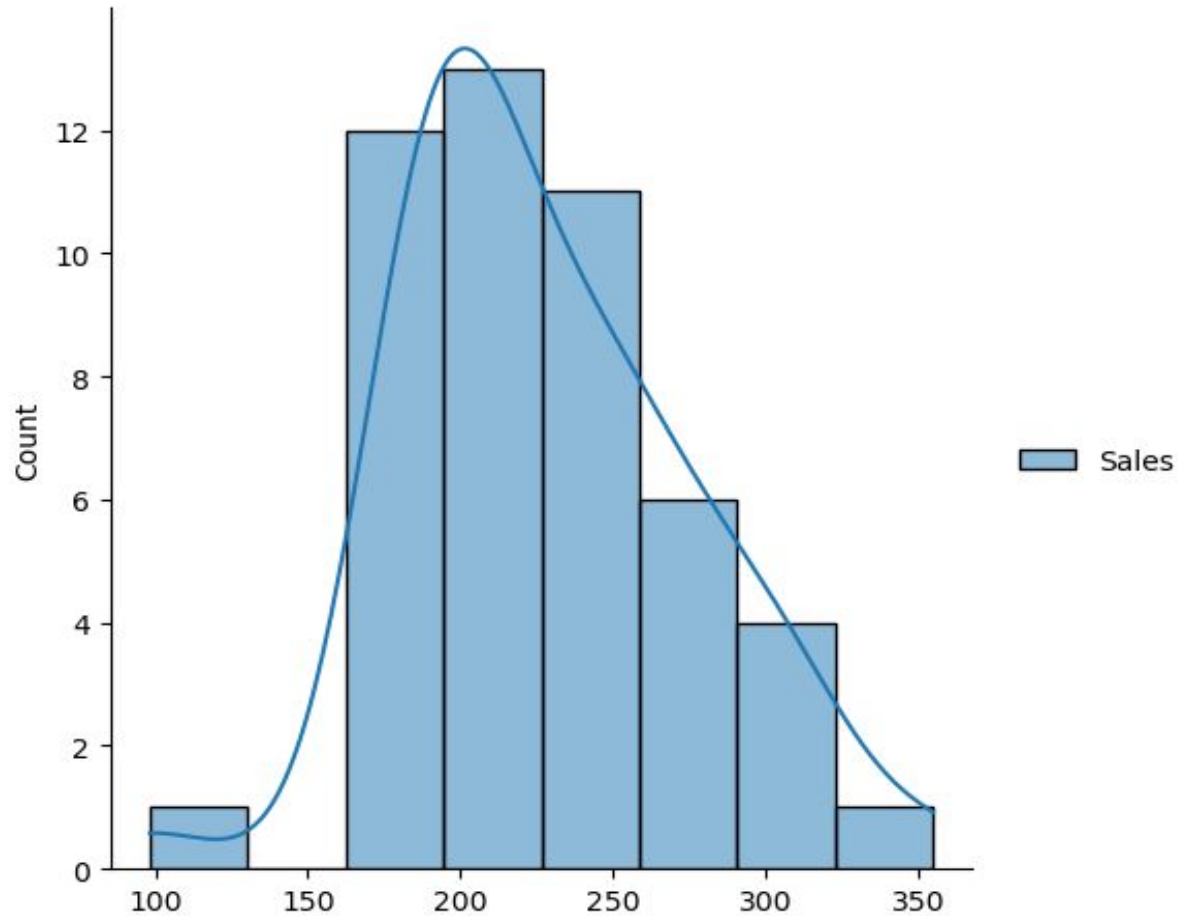
**#Visualizing the monthly  
sales dataset using lineplot  
`sns.lineplot(monthly_data)`**

## Checking for Trend

<Axes: xlabel='Order Date', ylabel='Sales'>



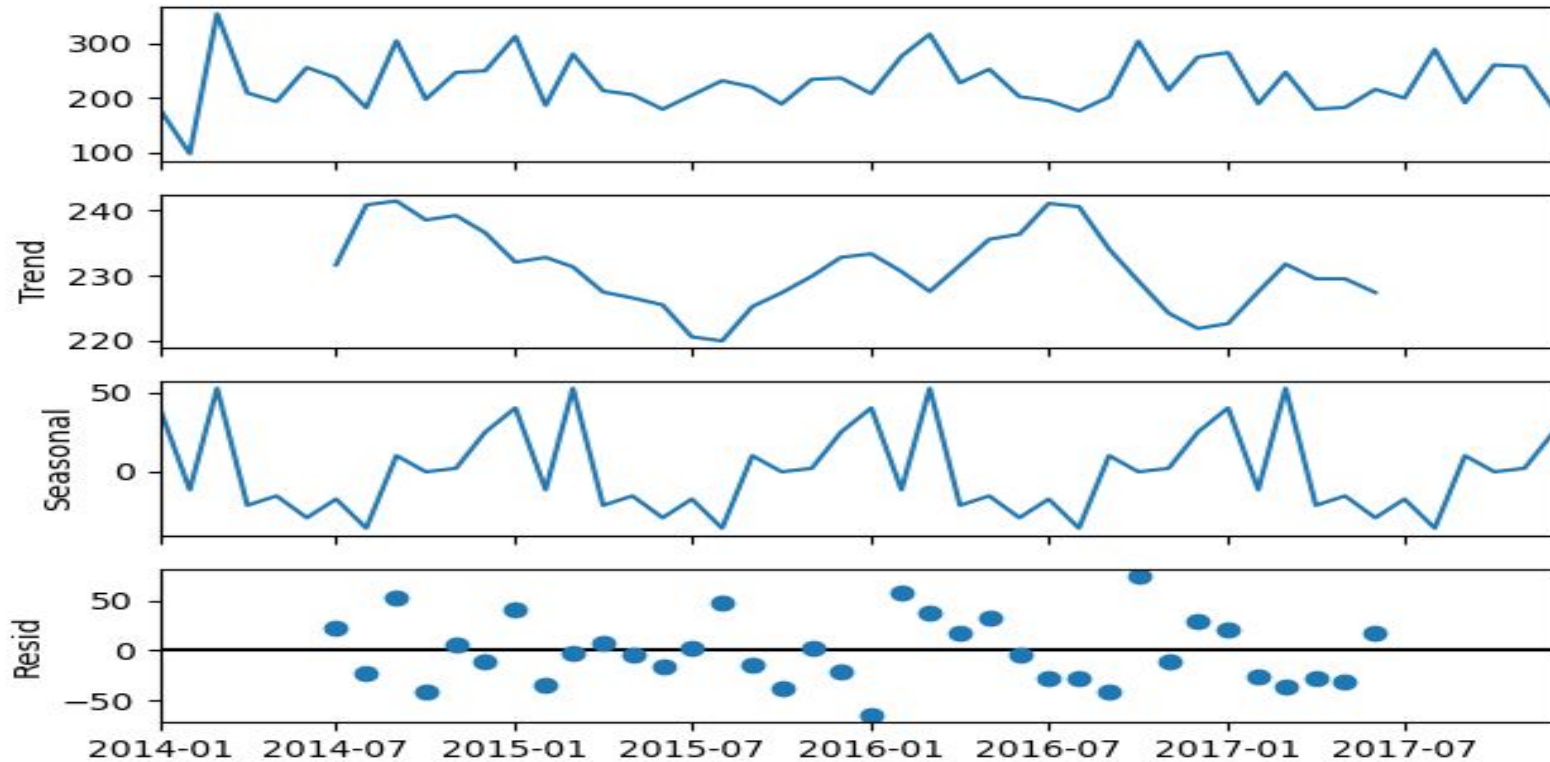
#Checking for trend using  
rolling mean  
# comparing the two  
lineplot



**# visualizing the distribution of  
the dataset  
`sns.displot(monthly_data,  
kde=True)`**



#Checking for stationarity using dickyfuller test on monthly data



```
output = adfuller(monthly_data)
```

Output

```
(-3.2865668298704307,  
0.015489720191097255,  
10,  
37,  
{'1%': -3.6209175221605827,  
'5%': -2.9435394610388332,  
'10%': -2.6104002410518627},  
376.49084160927134)
```

\* $H_0$  : it is non-stationary

\* $H_1$  : it is stationary

If  $p < 0.05$  ; Data is stationary

if  $p > 0.05$ ; Data is not stationary

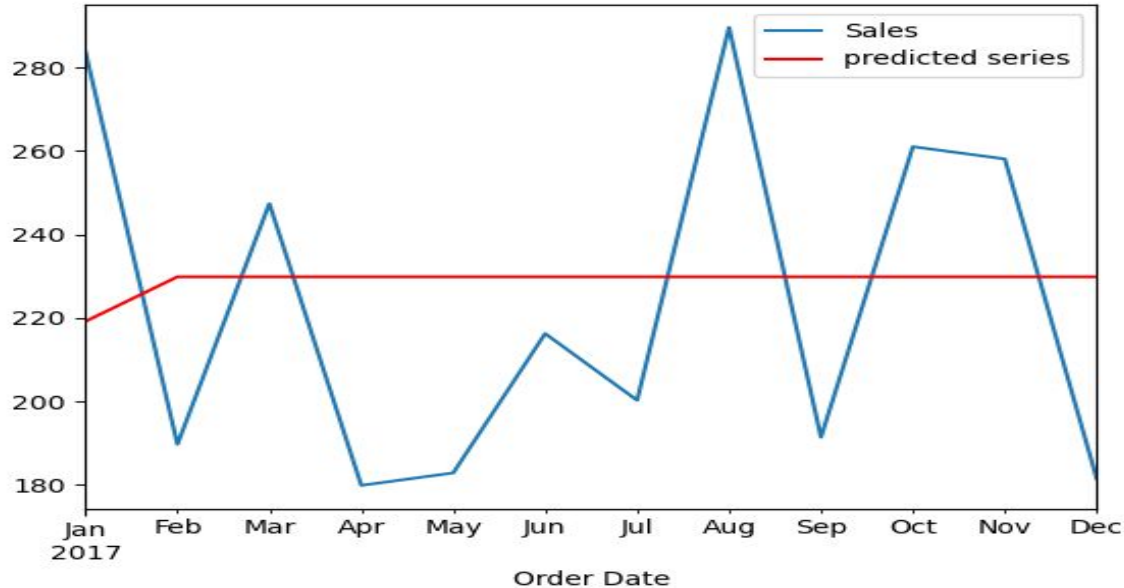
Since the p-value is 0.00020180198458237758 ,which is less than 0.05 we conclude that the data is stationary

# Modelling

## ARIMA Model

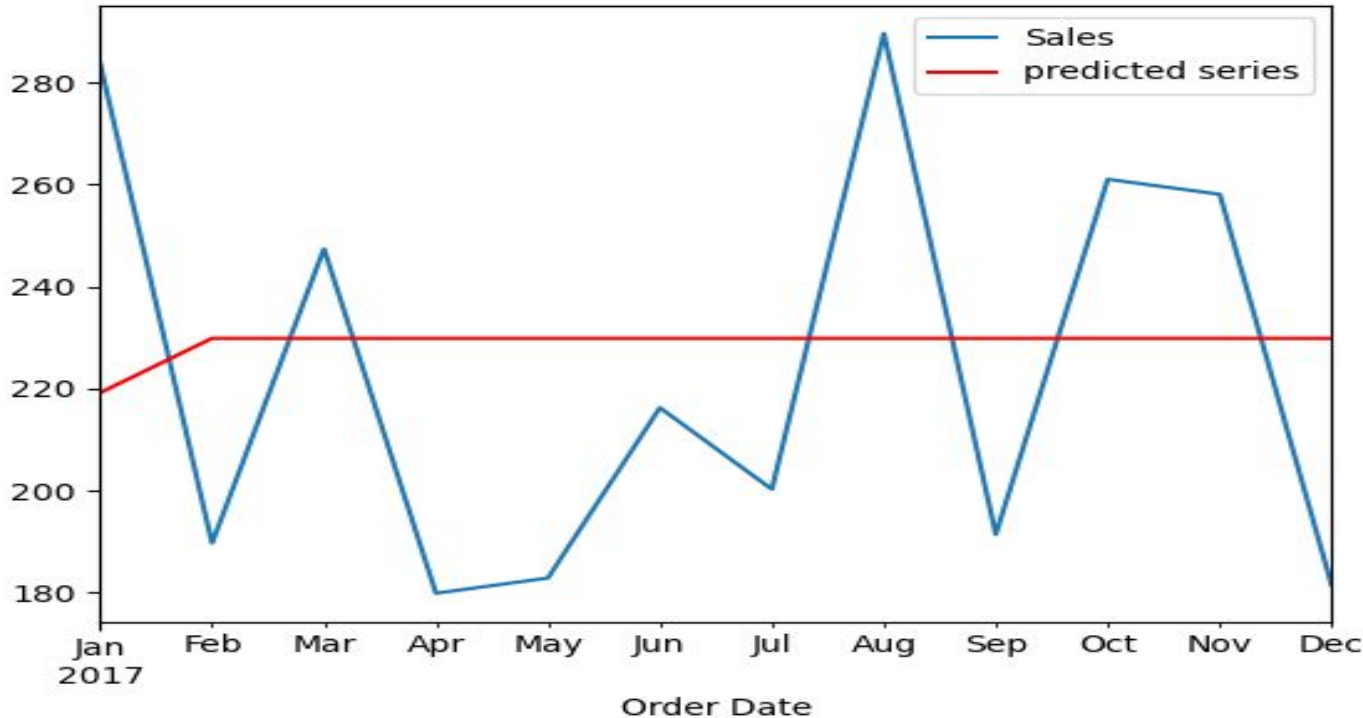
- # plotting the test set(actual) and predict model(forecast) to see a comparisson
- # Plot forecasts against actual outcomes

ARIMA Model - MSE:  
1748.62, RMSE: 41.82



## SARIMA Model

- # plotting the test set(actual) and predict model(forecast) to see a comparisson
- # Plot forecasts against actual outcomes

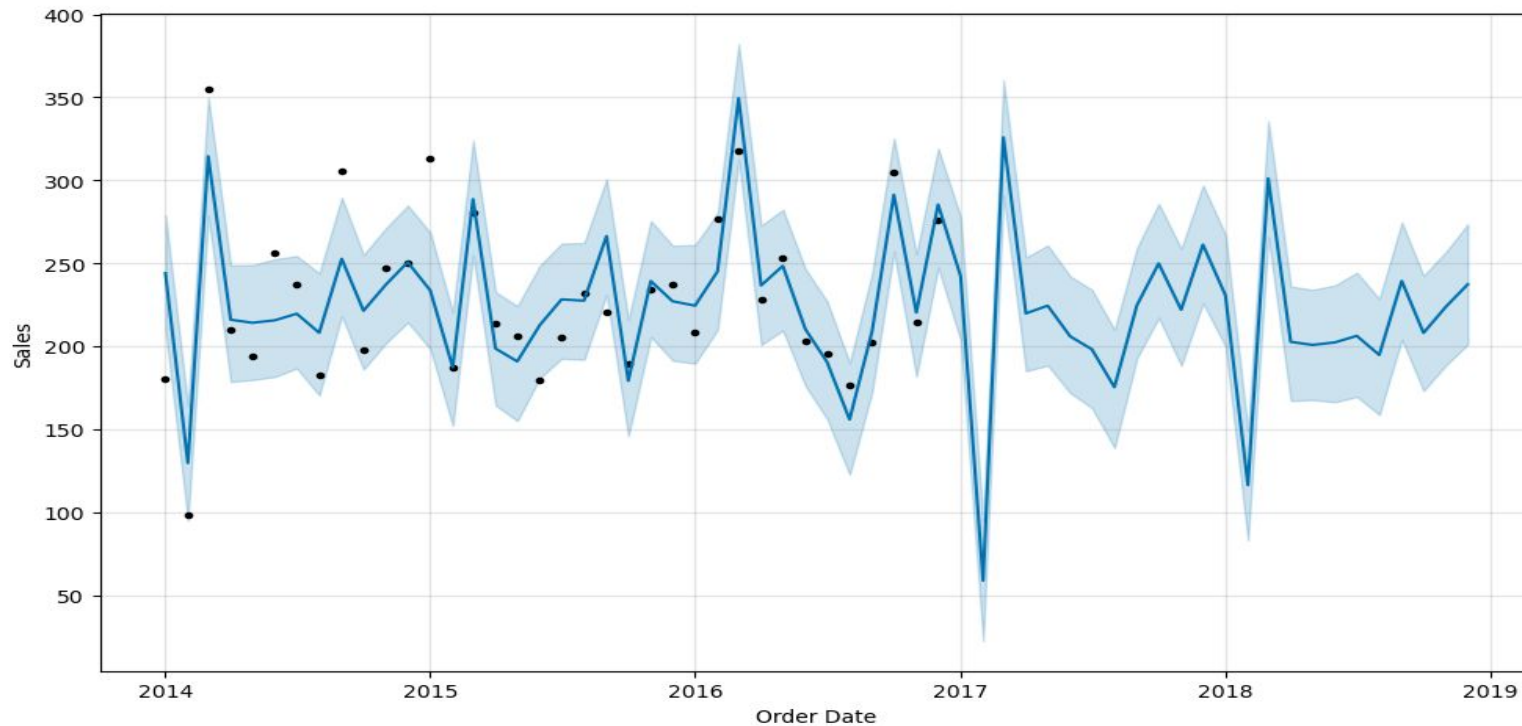


SARIMA Model - MSE:  
1748.62, RMSE: 41.82

## Facebook Prophet Model

# Plot the original data and the forecast

```
fig = prophet_model.plot(forecast, xlabel='Order Date', ylabel='Sales')
```



# Conclusion

After thorough evaluation of the three models, Arima, Sarima and Facebook Prophet using the RMSE of the models ; Facebook prophet had the least RMSE value. This study therefore found the Facebook Prophet model most appropriate to forecast sales for Superstore supermarket. The prediction indicates a deep in sales. The month of January 2018 experienced the lowest deep with sales as low as 110.

## **Recommendations**

This study recommends increased use of digital marketing strategies to promote sales in future. The sales team should also consider discounts in future so as to entice customers thereby increasing sales.

However, since there is probability of decline in sales in the future, the Company should avoid overstocking.