



Sales Forecasting Using SARIMA and ETS Model

Novia Khusna Faizah



Table of contents



01

Dataset
Introduction

02

Objective

03

Exploratory Data
Analysis

04

Modelling

05

Model Evaluation

06

Conclusion and
Recommendations





01

Dataset Introduction



Dataset Introduction



Dataset

<https://www.kaggle.com/datasets/rohitsahoo/sales-forecasting>



Our idea

The dataset used is Superstore data which contains Superstore sales transactions from 2015 to 2019





02

Objective



Objective

The objective of this project is to find a better sales projection model between the SARIMA and ETS models in order to carry out sales projections for the next 12 months.



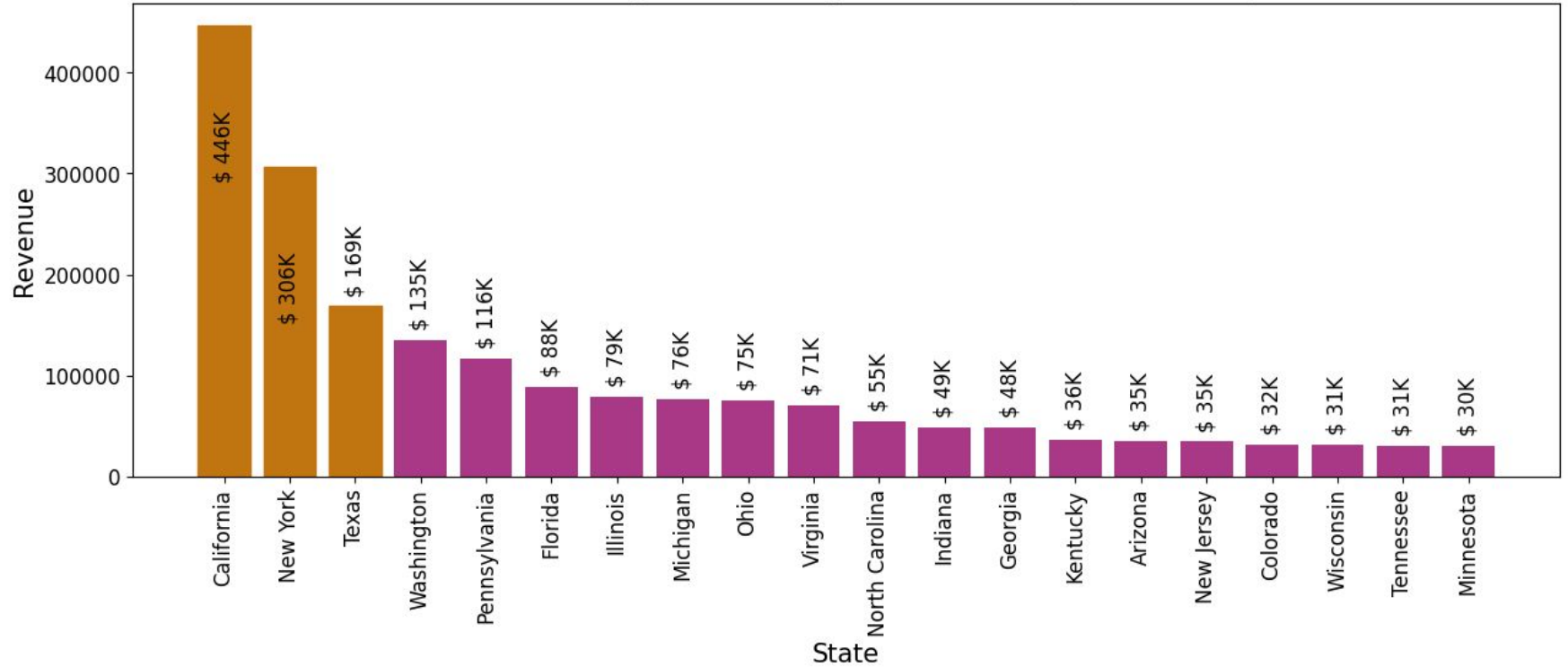


03

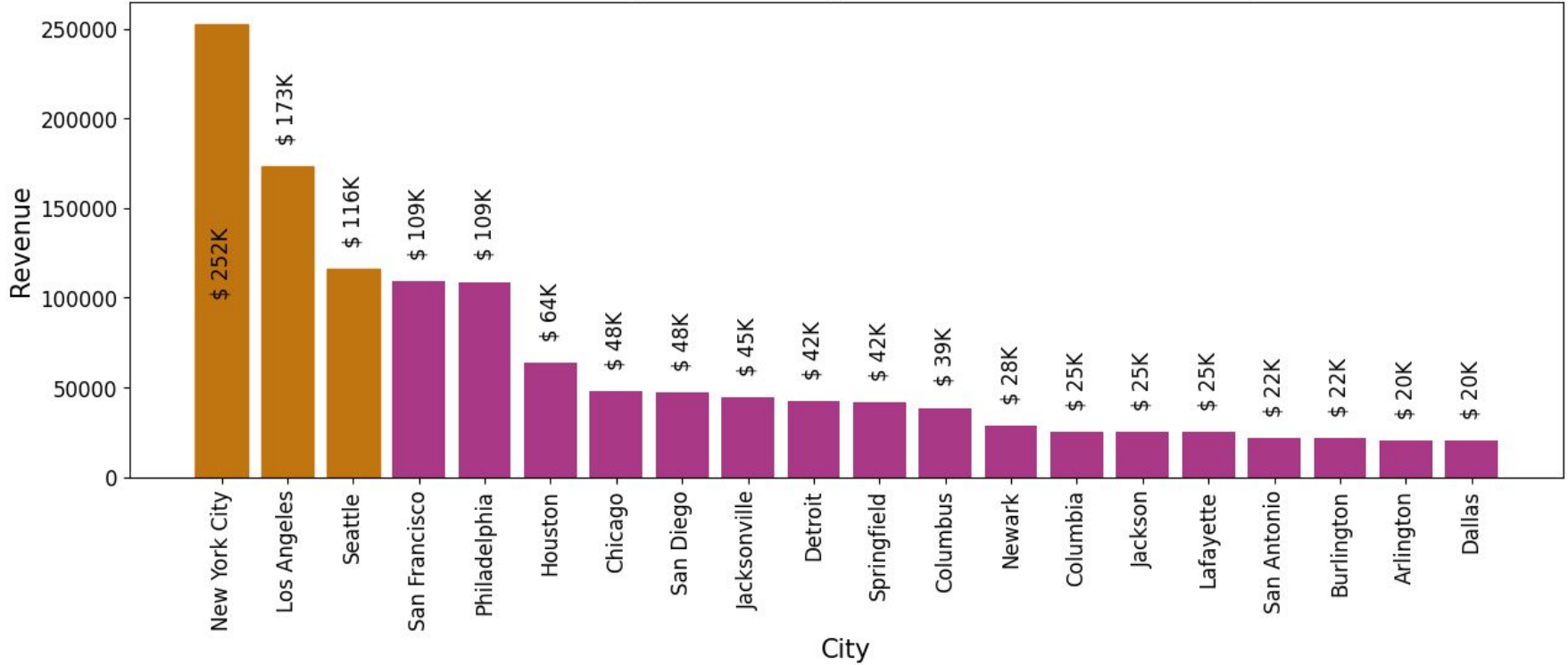
Exploratory Data Analysis



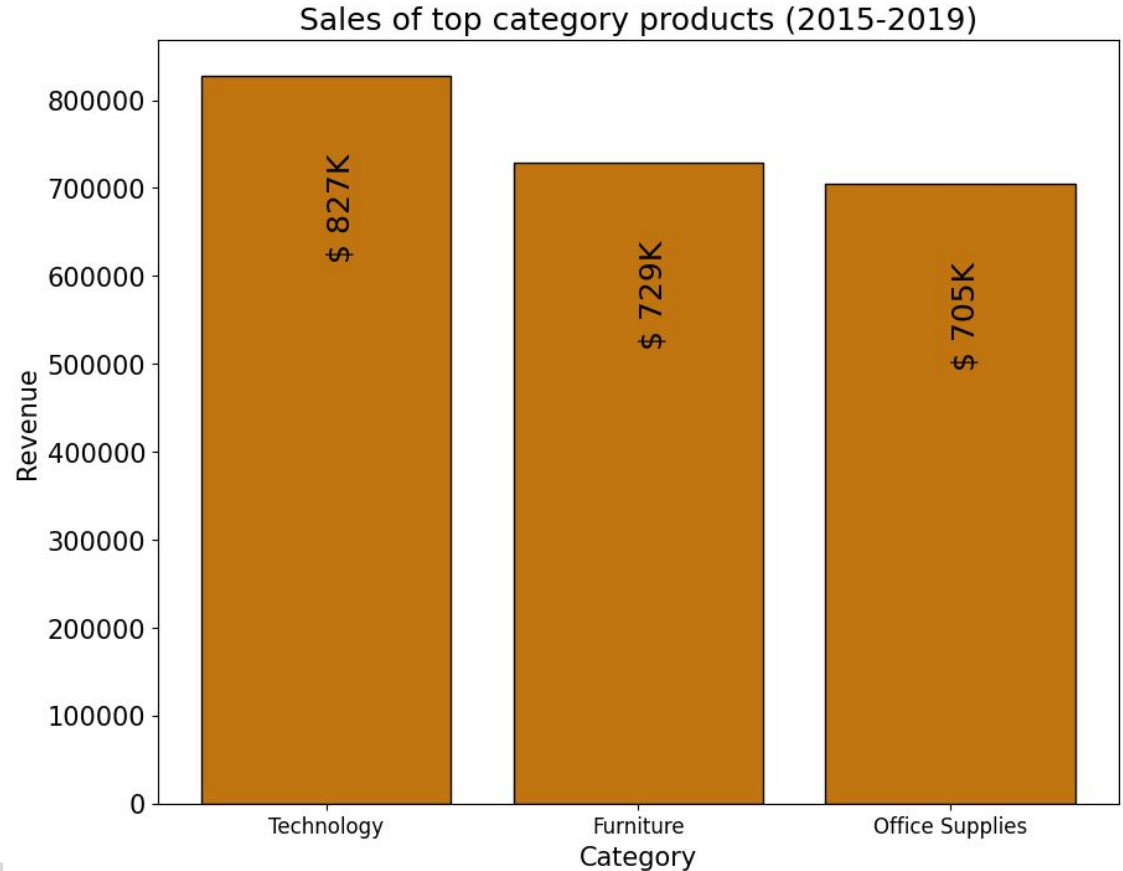
State which generated Highest Revenue (2015-2019)



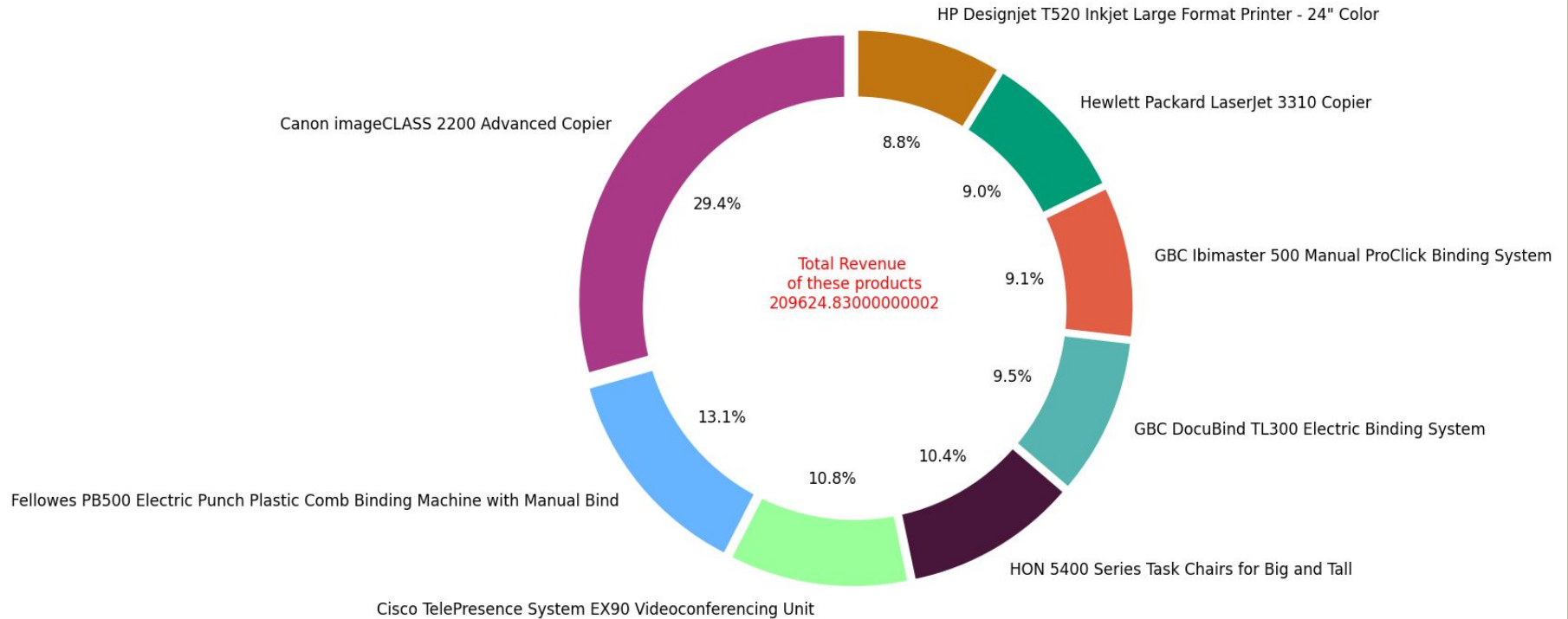
Cities which generated Highest Revenue (2015-2019)

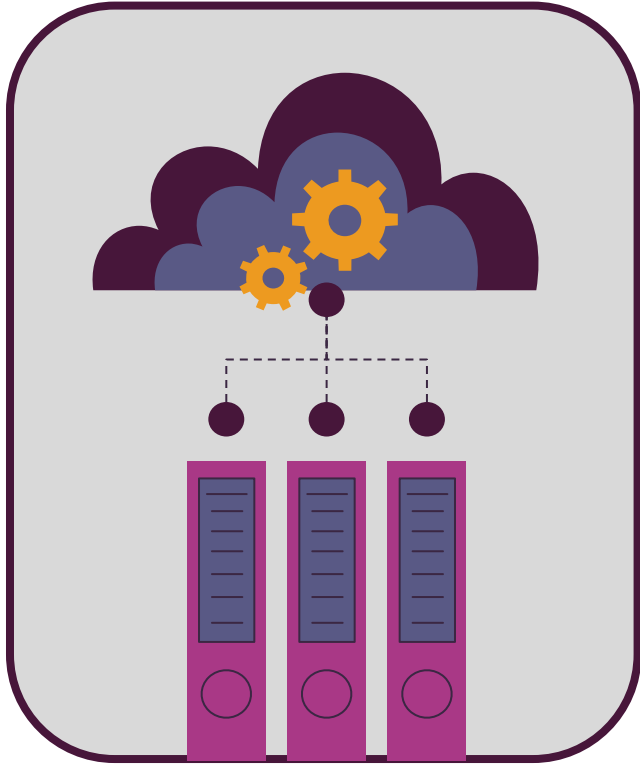


There are three categories of products sold, technology occupies the highest position in the product category with the highest sales with a nominal value of \$ 827,426



Top Sales by Product



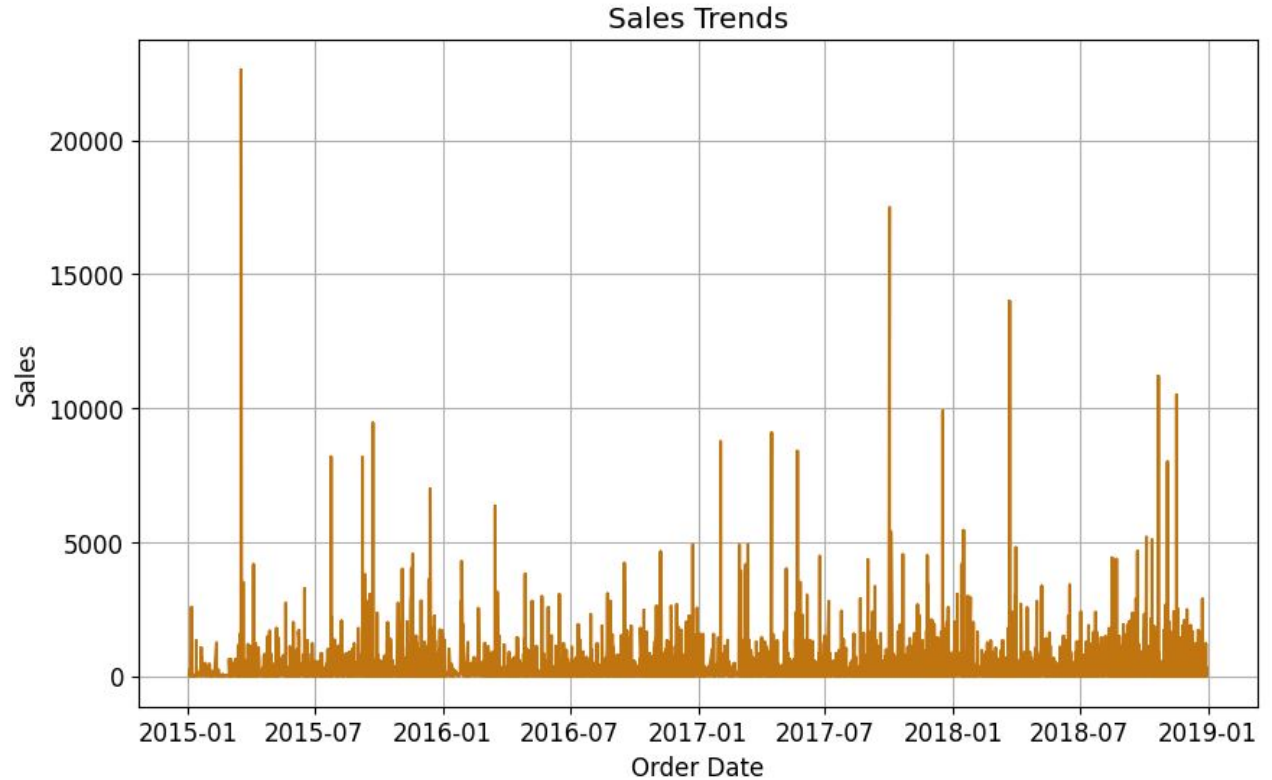


04

MODELLING



A time series is said to be stationary if it has a variance and mean that do not change, or tend to be constant. If we look at the sales distribution from 2015 - 2019, it is still unclear whether the data is stationary or not.



Stationarity Testing

ADF Testing

KPSS Testing

Goal	Testing whether data is not stationary	Testing whether data is stationary
Hypothesis	Ho : Time series is not stationary H1 : Time series is stationary	Ho : Time series is stationary H1 : Time series is not stationary
Interpretation	P-Value < 0.05, the data stationary	P-Value > 0.05, the data stationary



Stationarity Testing



ADF Testing

ADF Statistic: -99.08279579913294

P-value: 0.0

Critical Values: {'1%': -3.4310175185079284, '5%': -2.861835002806603, '10%': -2.566927024865997}

Time series is stationary.

KPSS Testing

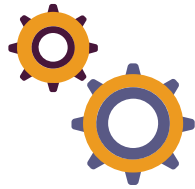
KPSS Statistic: 0.1118274706161094

P-value: 0.1

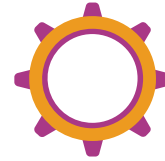
Critical Values: {'10%': 0.347, '5%': 0.463, '2.5%': 0.574, '1%': 0.739}

Time series is stationary.





Seasonal ARIMA



SARIMA (Seasonal Autoregressive Integrated Moving Average) is a statistical model used to model time series, especially those that show seasonal or recurring patterns in the data.

Main component:

AR (Autoregressive): P

I (Integrated): D

MA (Moving Average): Q

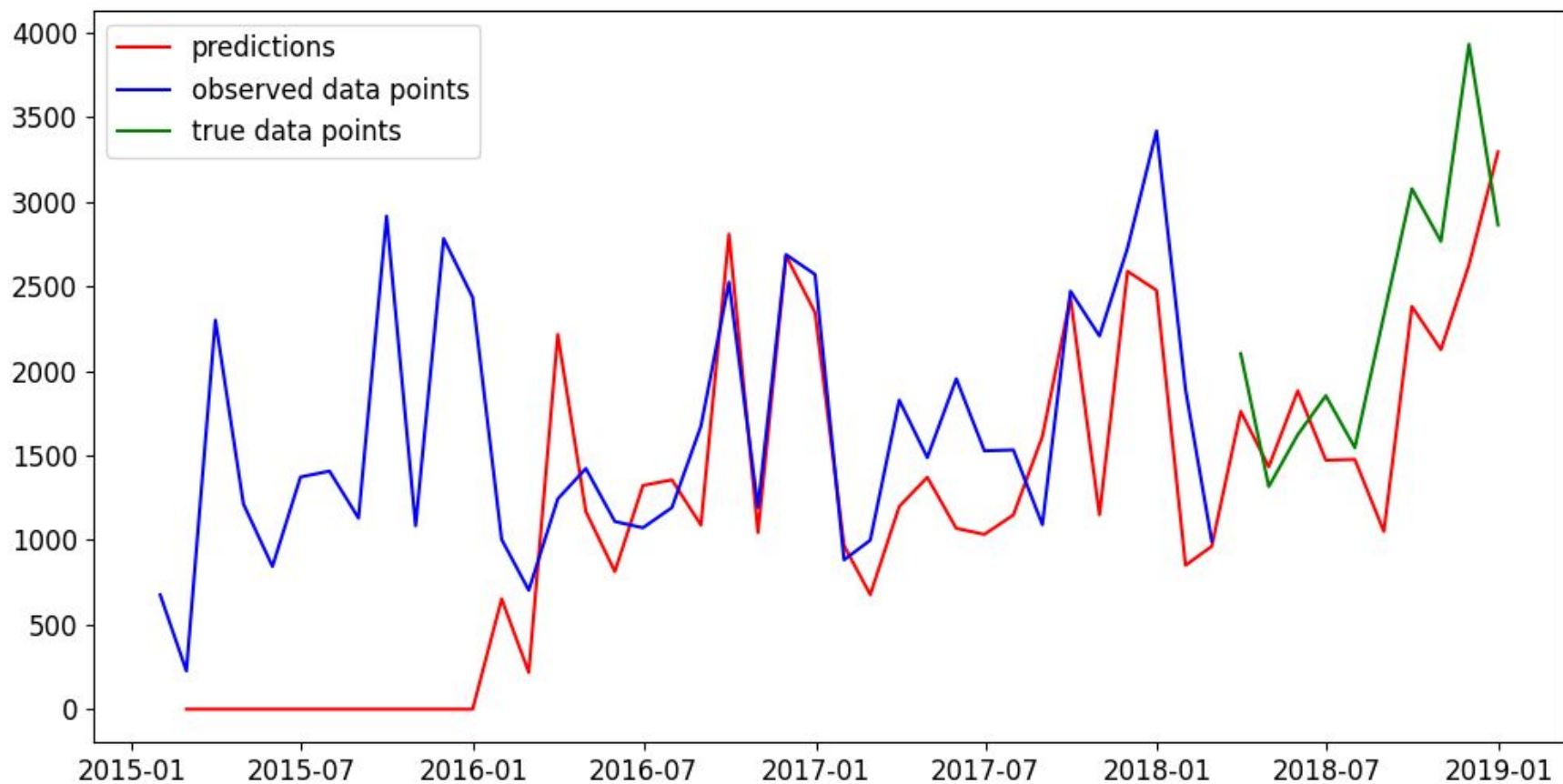
S (Seasonal): s

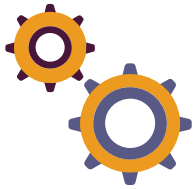
Steps:

1. Split Train and Test data
2. Finding the best parameters (P, D, Q, s)
3. Perform testing on the best parameters
4. Model Evaluation

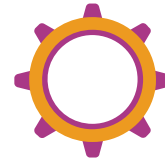


SARIMA Forecast





Exponential Smoothing (ETS)



The ETS model (Error, Trend, Seasonal) is a time series projection model designed to handle various patterns in time series, including linear or non-linear trends, as well as seasonal components that can appear in time series data.

Main component:

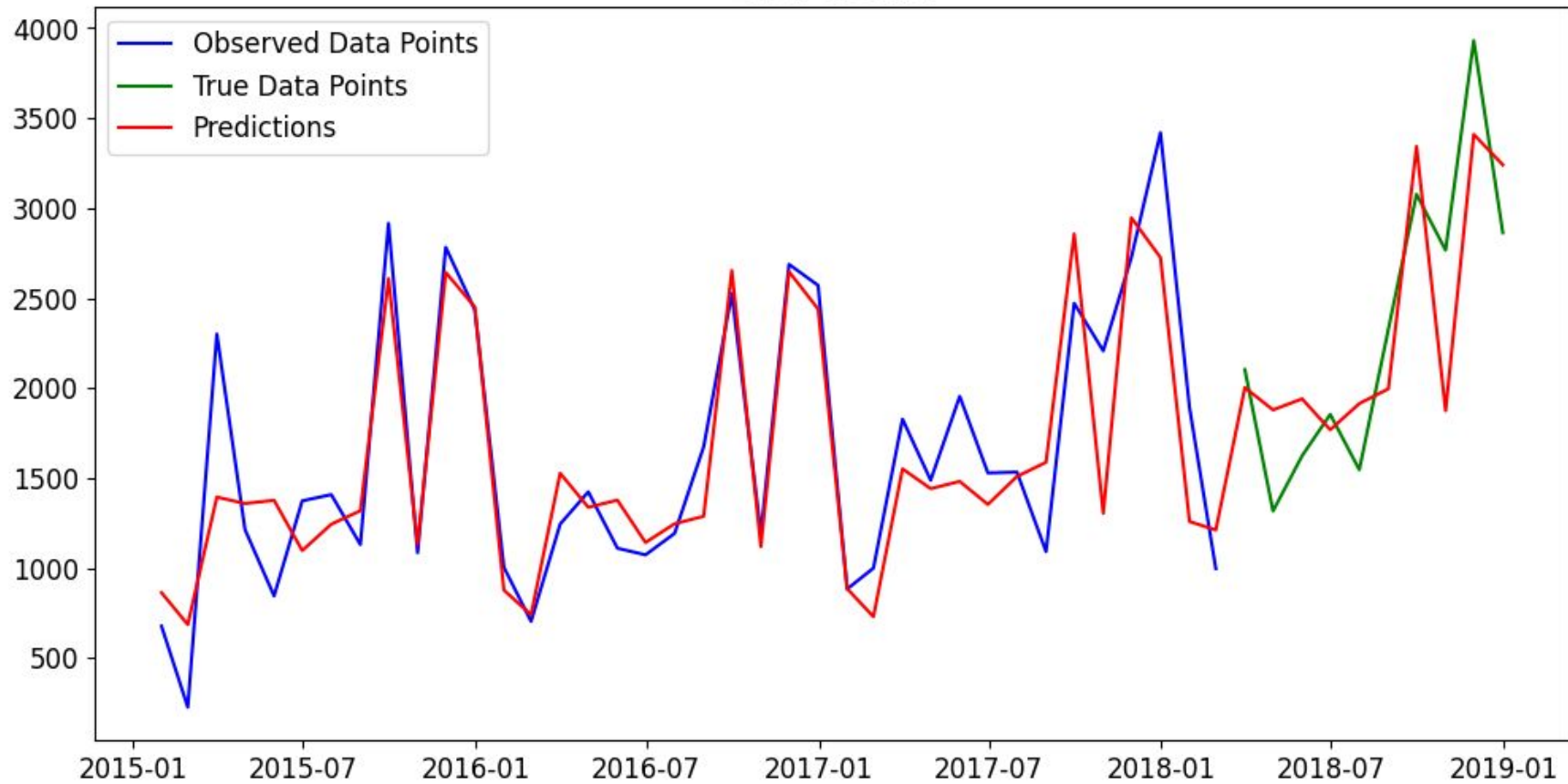
- E (Error)
- T (Trend)
- S (Seasonal)

Steps:

1. Split Train and Test data
2. Select the method to be used
3. Perform testing on testing data
4. Model Evaluation



ETS Forecast

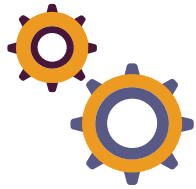




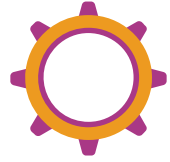
05

MODEL EVALUATION





Perbandingan SARIMA VS ETS



MODEL SARIMA

MAE: 721.0311043781512
MAPE: 0.44194535285045217
RMSE: 1003.5981041612715
MSE: 1007209.1546760985
R-Squared: -0.5583488455607937

MODEL ETS

MAE: 285.0785784286879
MAPE: 0.20454308533167942
RMSE: 369.51375809014723
MSE: 136540.41741790387
R-Squared: 0.7887453654807128

1. The smaller the MAE, MAPE, MSE, and RMSE indicate that the model provides better prediction, because it results with smaller deviations or errors
2. The R-Squared value gets better as it approaches 1
3. ETS Model is the better model to predict Sales



06

CONCLUSION & RECOMMENDATIONS

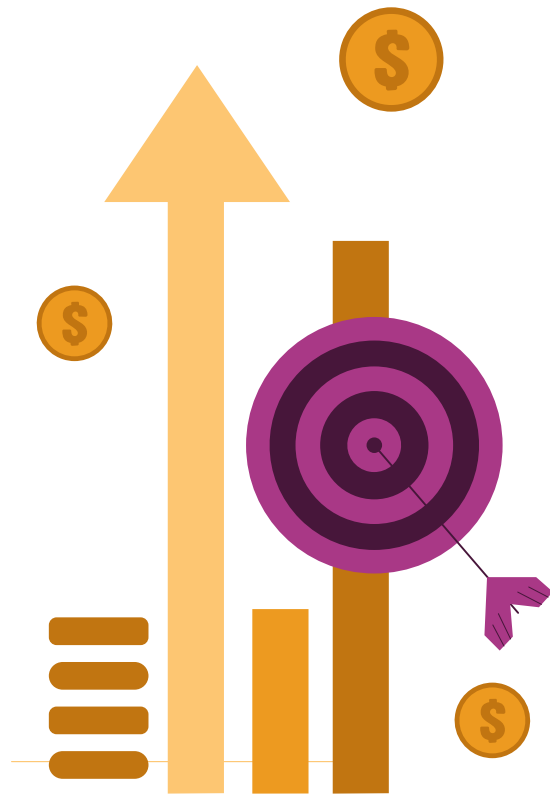


Conclusions

1. If you look at the visualization of the model results, the ETS model provides better predictions than the SARIMA model, where there are several periods using the SARIMA model that experience under forecasting.
2. Based on this evaluation, consistently, the ETS model performs better than the SARIMA model in terms of MAE, MAPE, RMSE, MSE, and R-Squared. Therefore, based on these metrics, ETS can be considered the better choice between the two models

Recommendations

1. The more data used on the data, the more accurate the model will provide performance
2. Test with other models to provide the best predictions



References

<https://www.kaggle.com/datasets/rohitsahoo/sales-forecasting>

<https://towardsdatascience.com/how-to-forecast-sales-with-python-using-sarima-model-ba600992fa7d>

<https://www.kaggle.com/code/samruddhim/part-1-exploratory-data-analysis>

<https://www.kaggle.com/code/samruddhim/part-2-time-series-analysis>

https://medium.com/@tushar_aggarwal/master-the-power-of-sarima-a-comprehensive-step-by-step-guide-9aded3088949

<https://medium.com/analytics-vidhya/time-series-forecasting-models-726f7968a2c1>



Thanks!



CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)

Please keep this slide for attribution

