Sales Forecasting

Team-278

Index

1

INTRODUCTION

Introduction

- Company manufactures electric fans and sells them to consumers in various regions.
- Due to poor forecasting, the demand and supply ratio is poor.
- The task is to improve forecasting using advanced time series and machine learning algorithms.
- We have used data visualization tools and developed various time series, forecasting models.

2

WORKFLOW

WorkFlow

1 STEP

Raw Data

2 STEP

Exploratory Analysis

3 STEP

Approach and Model training

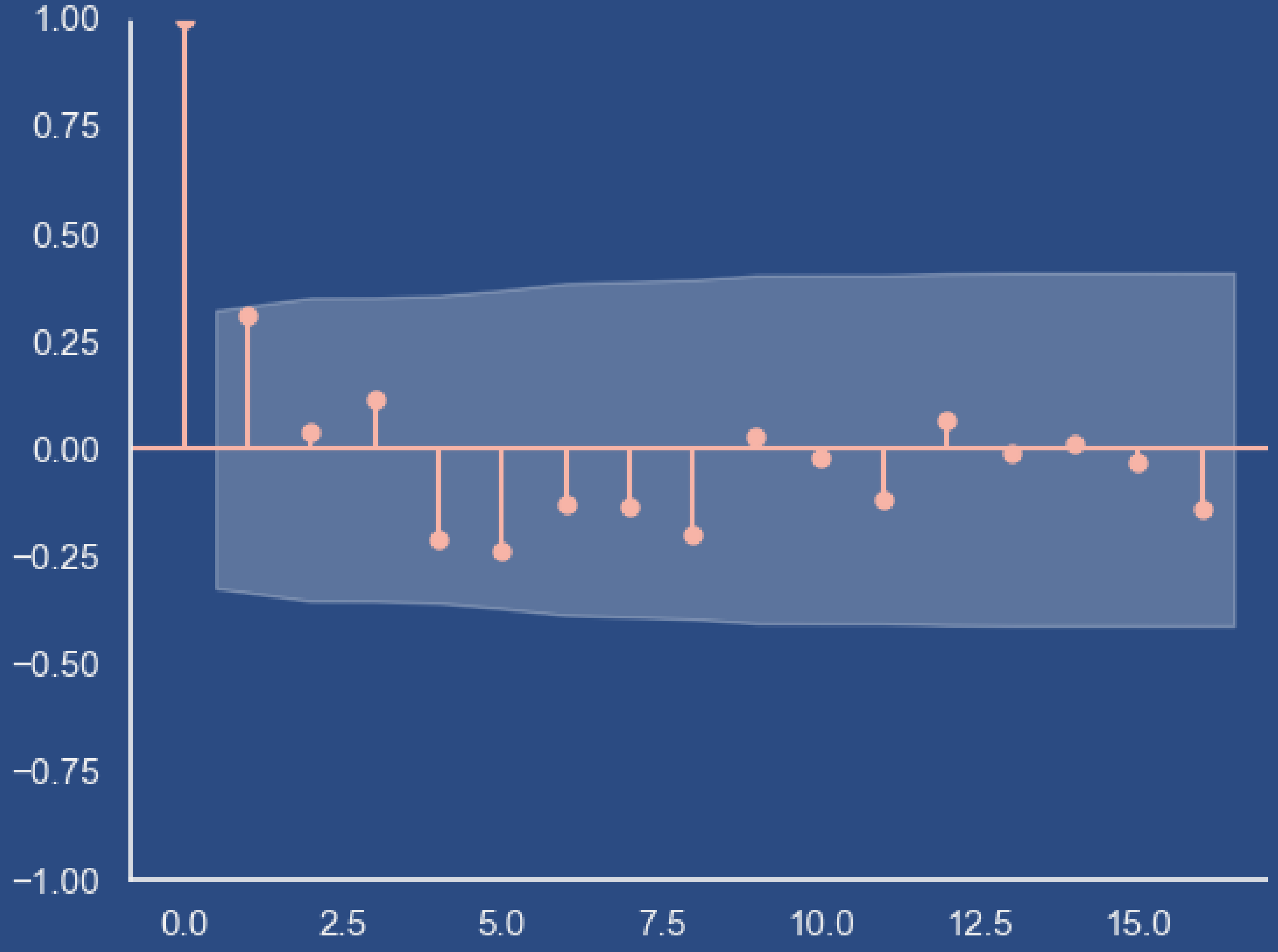
4 STEP

Forecasting & final analysis

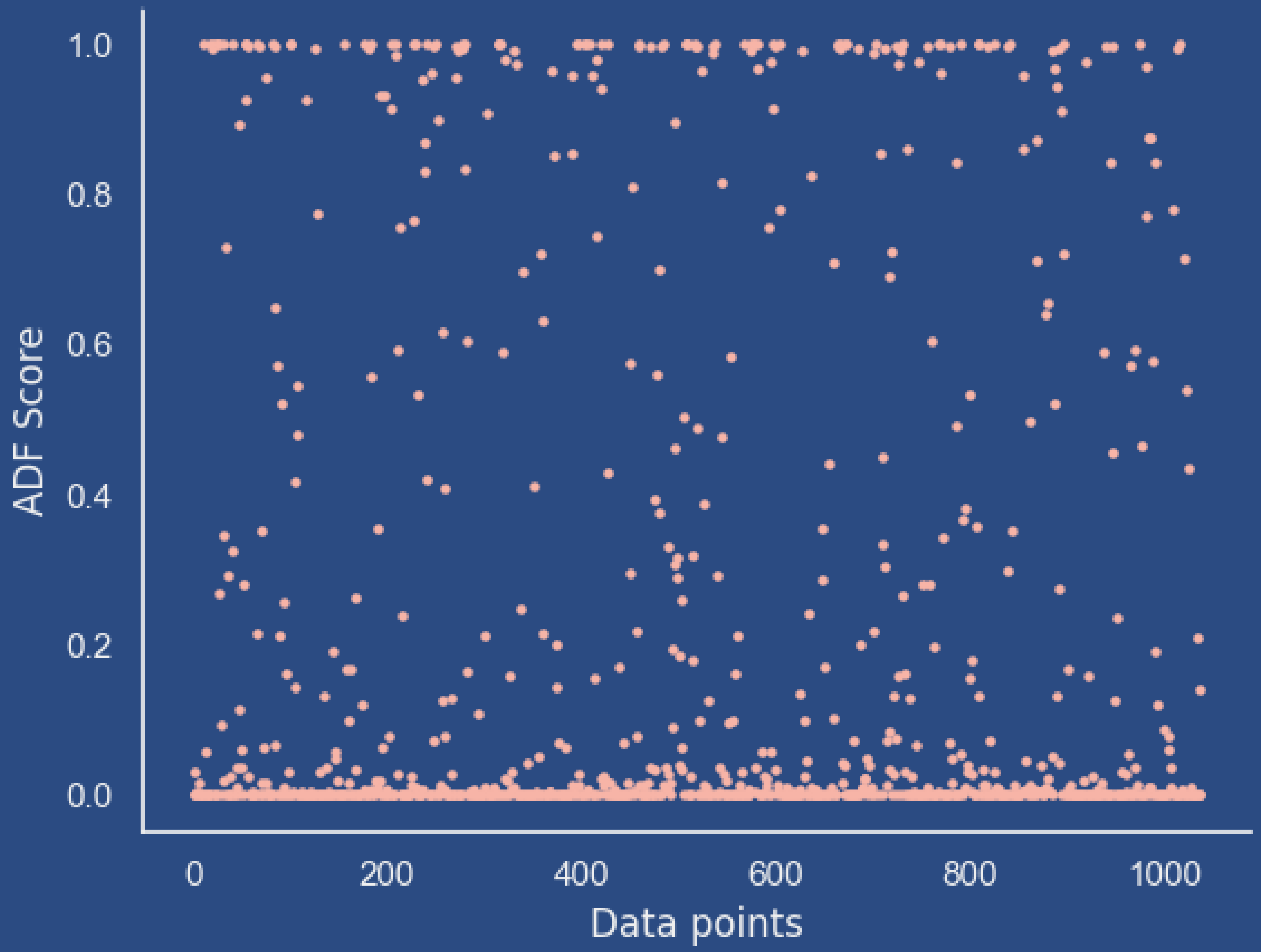
3

EXPLORATORY ANALYSIS

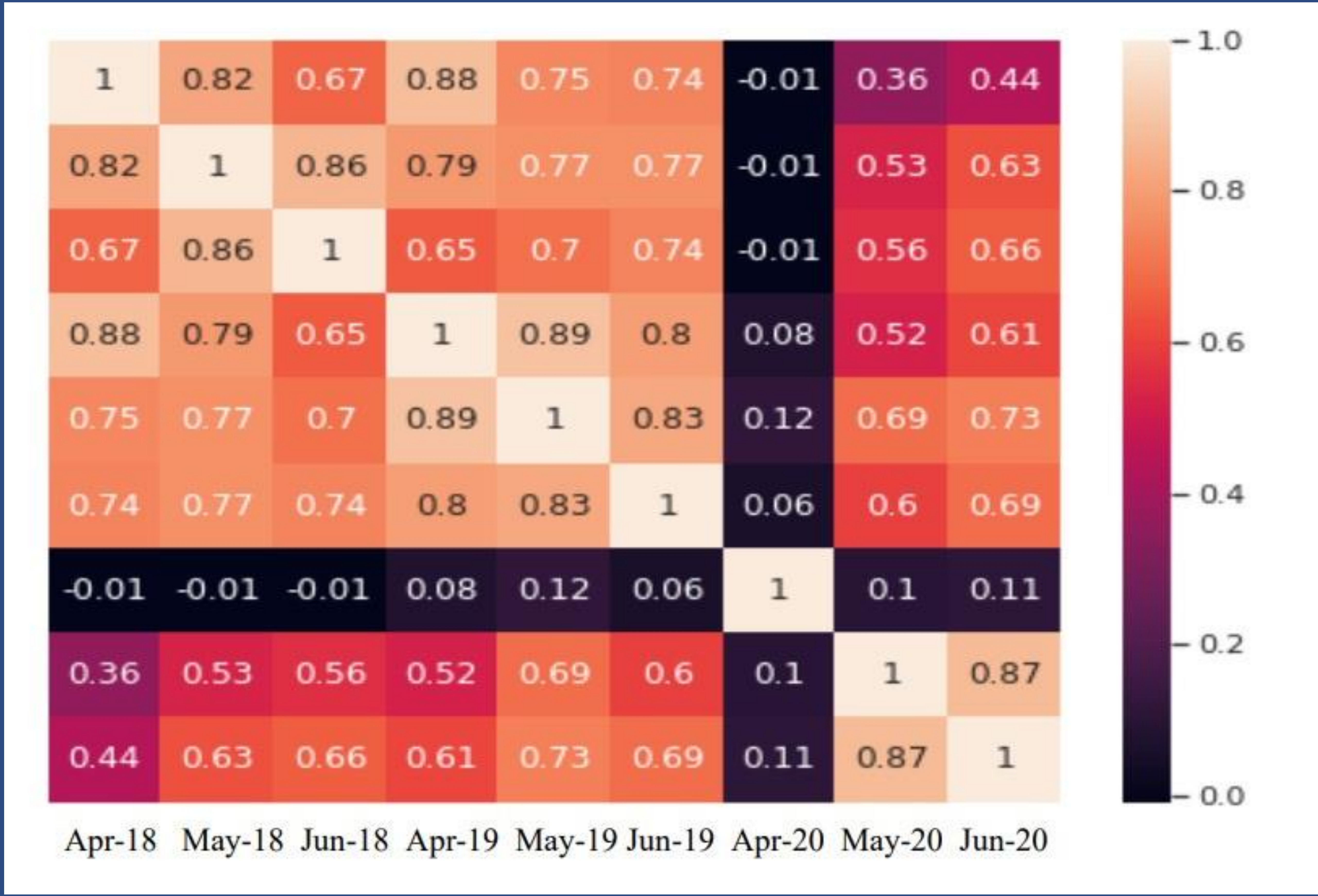
Autocorrelation Plot



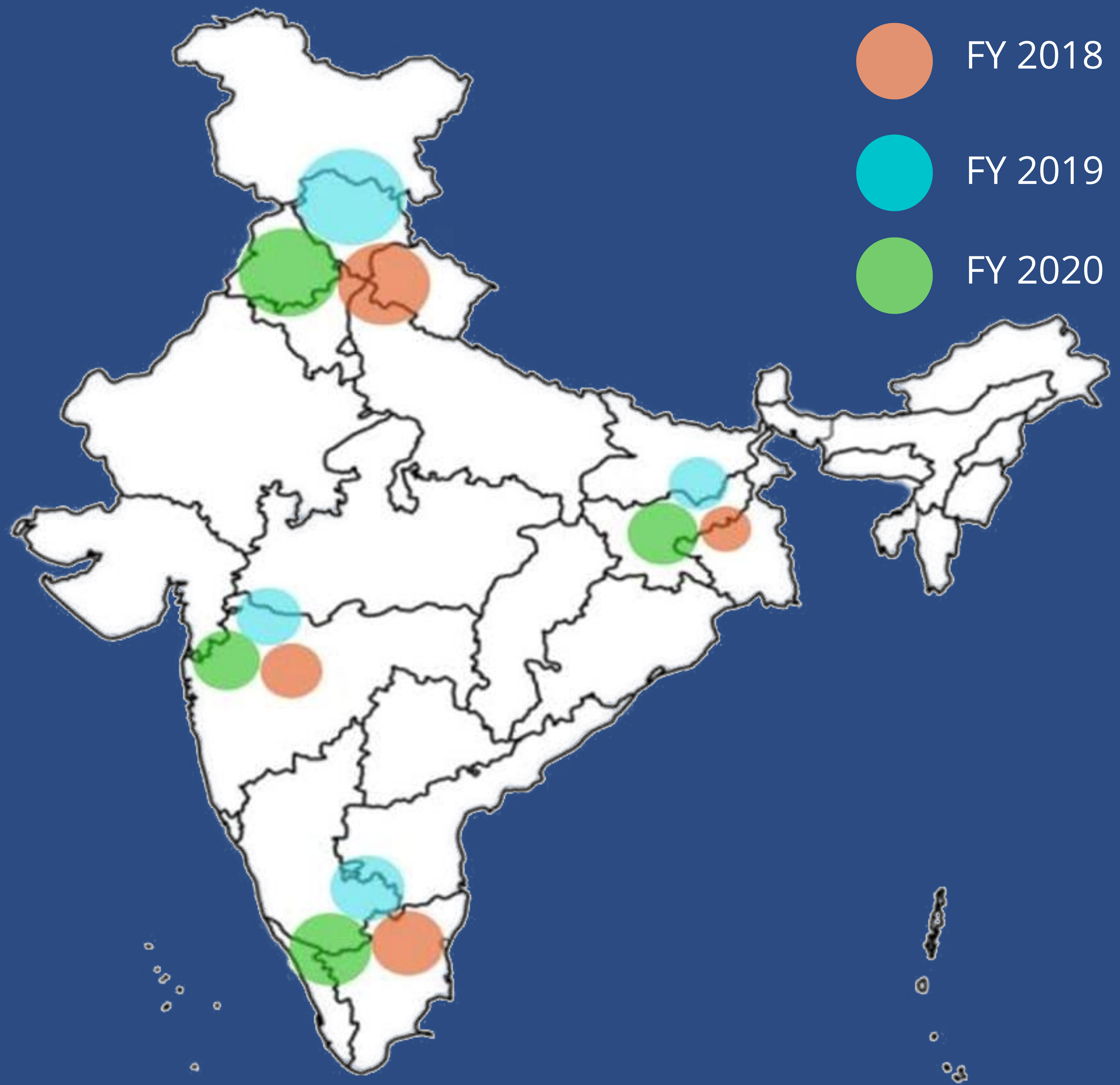
Augmented Dickey-Fuller (ADF) Test



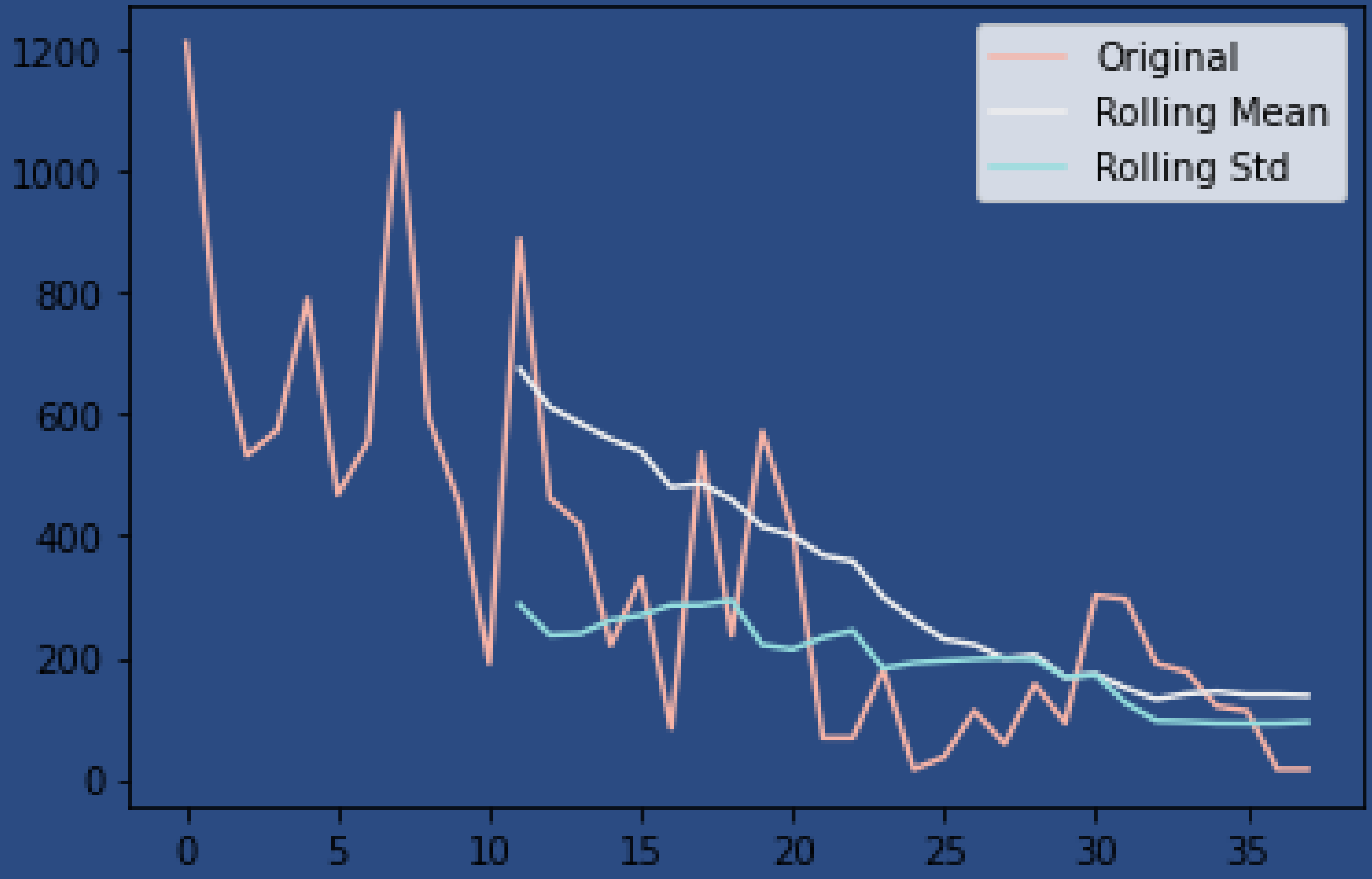
Heat Map


1 0.82 0.67 0.88 0.75 0.74 -0.01 0.36 0.44
0.82 1 0.86 0.79 0.77 0.77 -0.01 0.53 0.63
0.67 0.86 1 0.65 0.7 0.74 -0.01 0.56 0.66
0.88 0.79 0.65 1 0.89 0.8 0.08 0.52 0.61
0.75 0.77 0.7 0.89 1 0.83 0.12 0.69 0.73
0.74 0.77 0.74 0.8 0.83 1 0.06 0.6 0.69
-0.01 -0.01 -0.01 0.08 0.12 0.06 1 0.1 0.11
0.36 0.53 0.56 0.52 0.69 0.6 0.1 1 0.87
0.44 0.63 0.66 0.61 0.73 0.69 0.11 0.87 1
Apr-18
May-18
Jun-18
Apr-19
May-19
Jun-19
Apr-20
May-20
Jun-20
1.0
0.8
0.6
0.4
0.2
0.0

Geovisualization

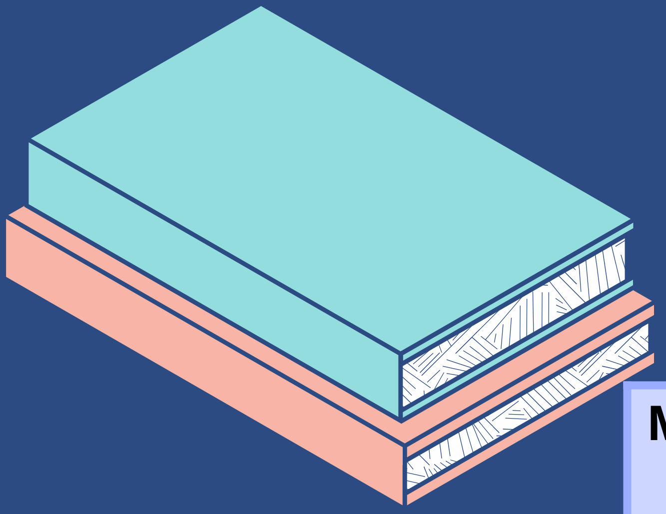


Rolling Mean and Rolling Standard Deviation



4

APPROACH & MODEL



Approach and Models

Models	MAPE	Root Mean Square Error
Random Forest	85%	355
XG Boost	93%	404
ARIMA	102%	473
SARIMA	215%	556
Holt Linear Trend	298%	621
Holt's Winter	365%	745
FB Prophet	655%	934
LSTM	94%	412
Temporal Fusion Transformer	544%	898

5

FINAL APPROACH

Problem

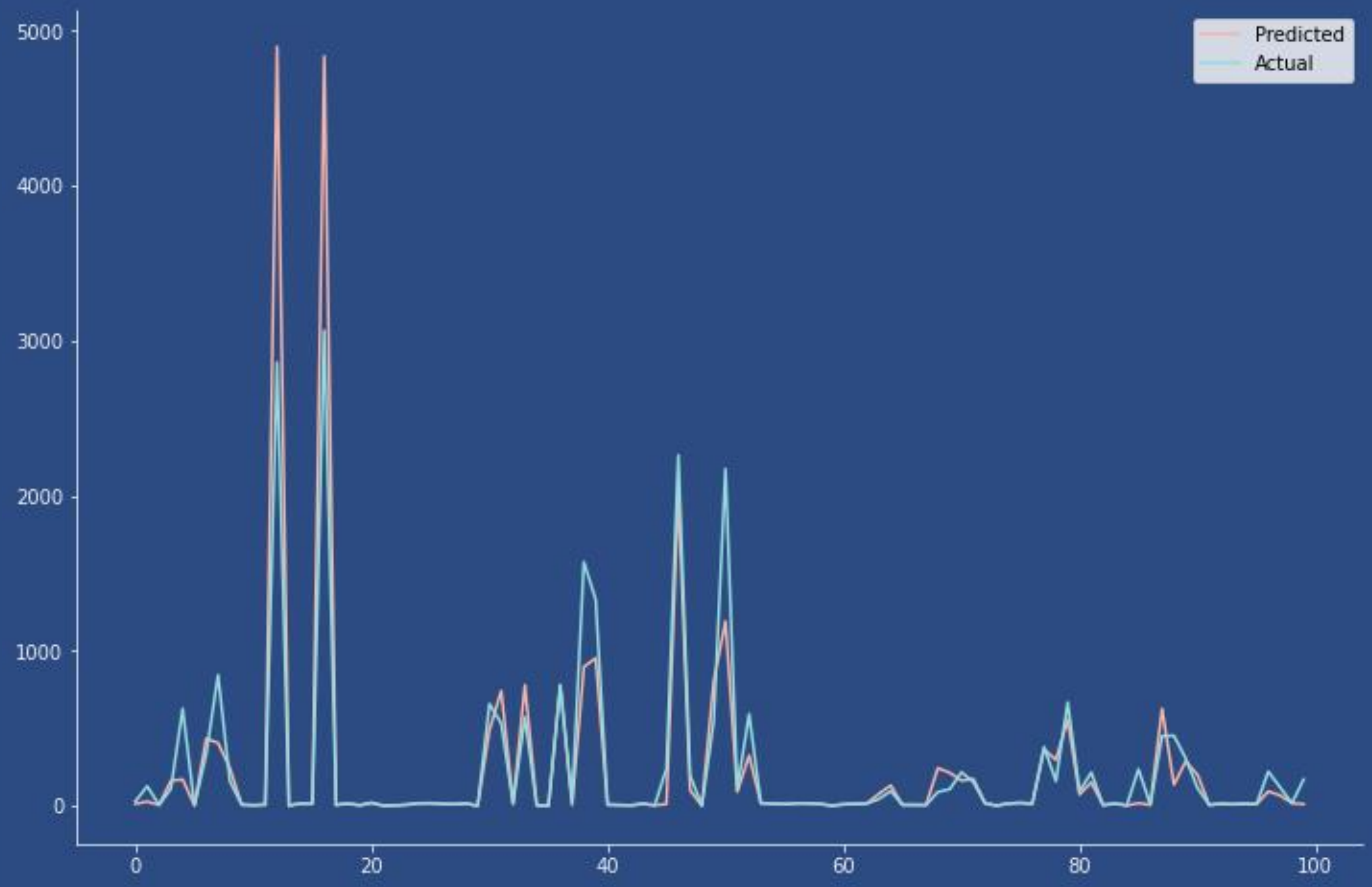
- In our initial approach, we made use of Random Forest, XG boost regressor, LSTM,TFT, ARIMA, SES, SARIMA, SARIMAX, HOLT'S linear trend, HOLT'S winter and FB prophet models.

- They turned out to be much more complex and resulted into a large MAPE. Moreover, the data provided wasn't showing similar trends over a larger time span.

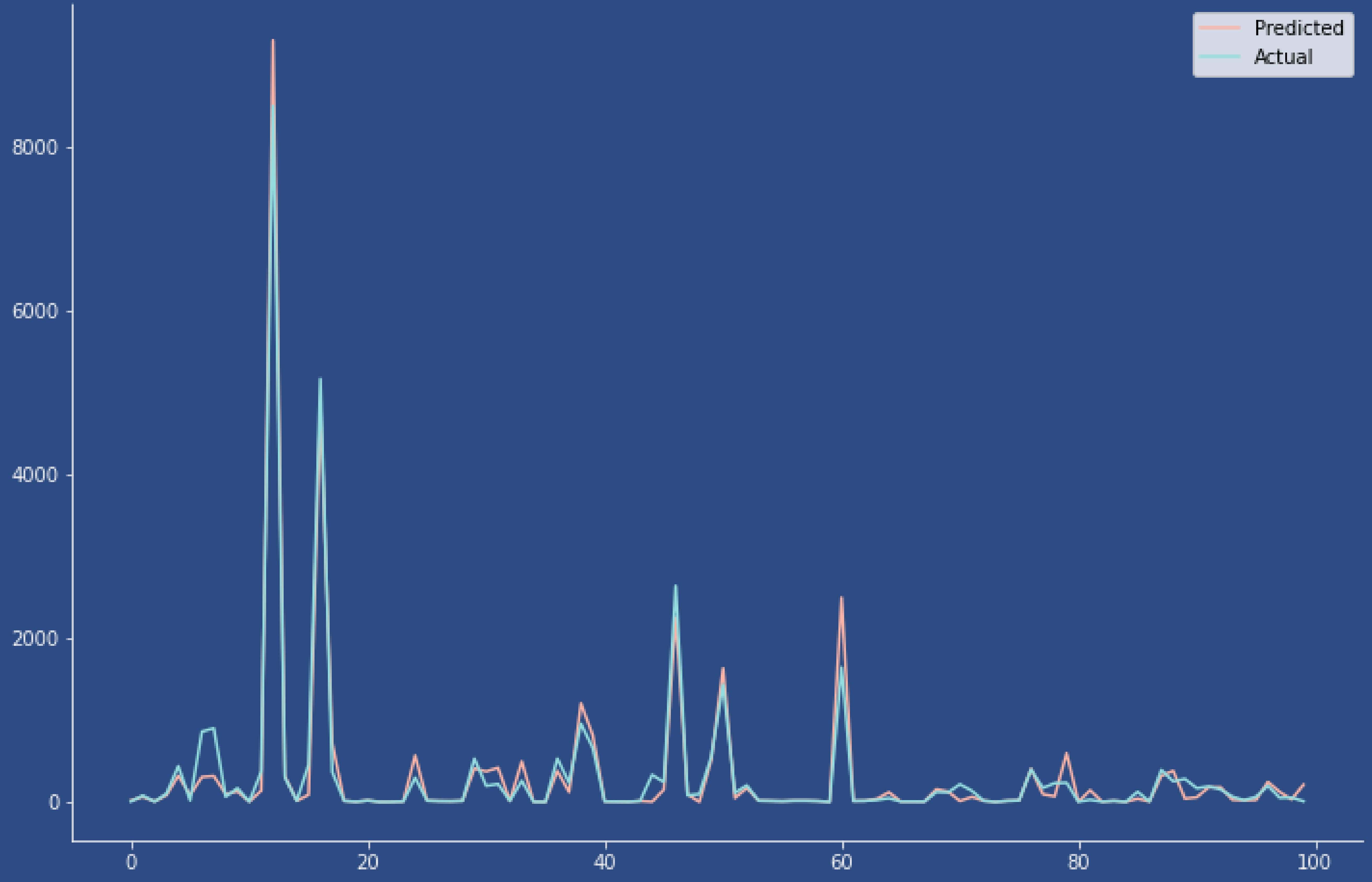
Solution

- Comparing sales of consecutive months shows a strong correlation between them. MAPE reduced significantly when month's sales were predicted by taking the previous month's sales.
- MAPE favors under-forecasting because the percentage error can't exceed 100% for relatively lower forecasts, while there is no upper limit for forecasts that are too high.
- So our final approach was as follows : **Predicted June's sales is the minimum of April and May sales.**

Predicted vs Actual June-18



Predicted vs Actual June-19



Thank You!