Bank Traffic Predictions

BeBug Team

Presented By:

Ali Ghazal
Mohammed Abuelwafa

Table of contents

01 Problem Statement

02 Descriptive Analytics

03 Methodology

04 Predictive Analytics

01

Problem Statement

Problem Statement

- We have a bank branch with **15** tellers who offer different services.
- Each Teller Serves **X** number of customers.
- The bank is trying to predict the customer traffic for the next month, on a daily basis

Data Characteristics

- We have no-null values.
- We have 10013 data entry.
- The data by nature had **NO** missing values to be handled.

Feature Engineering

We created the following features to ease our analysis:

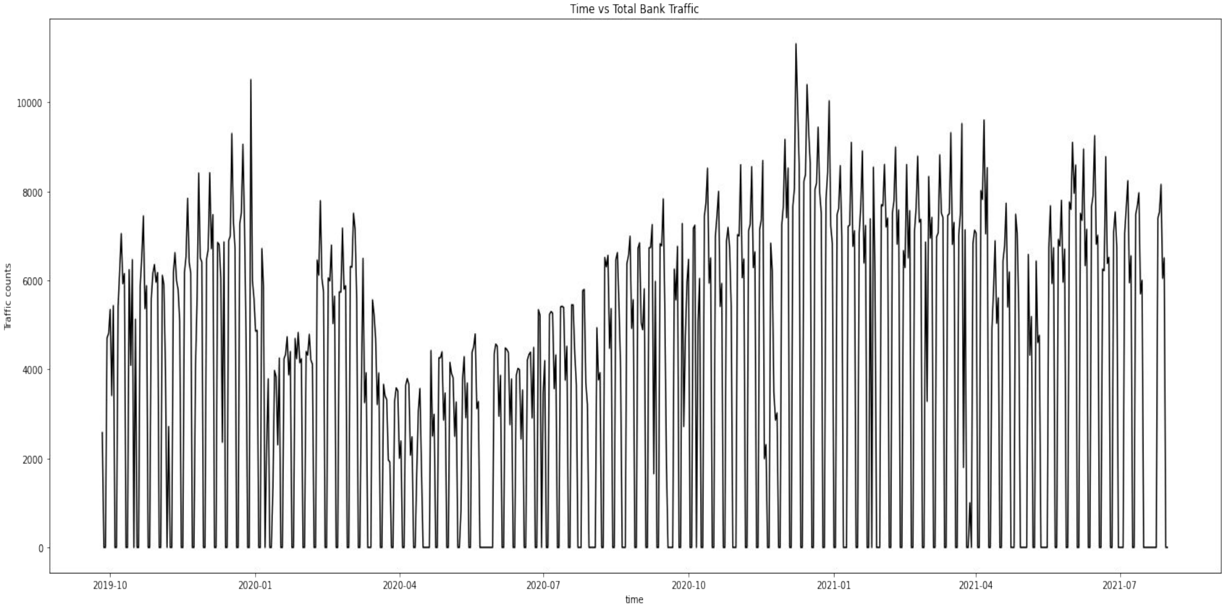
- weekDayInt
- WeekDayCalender
- isOfficialHoliday

02

Business Insights

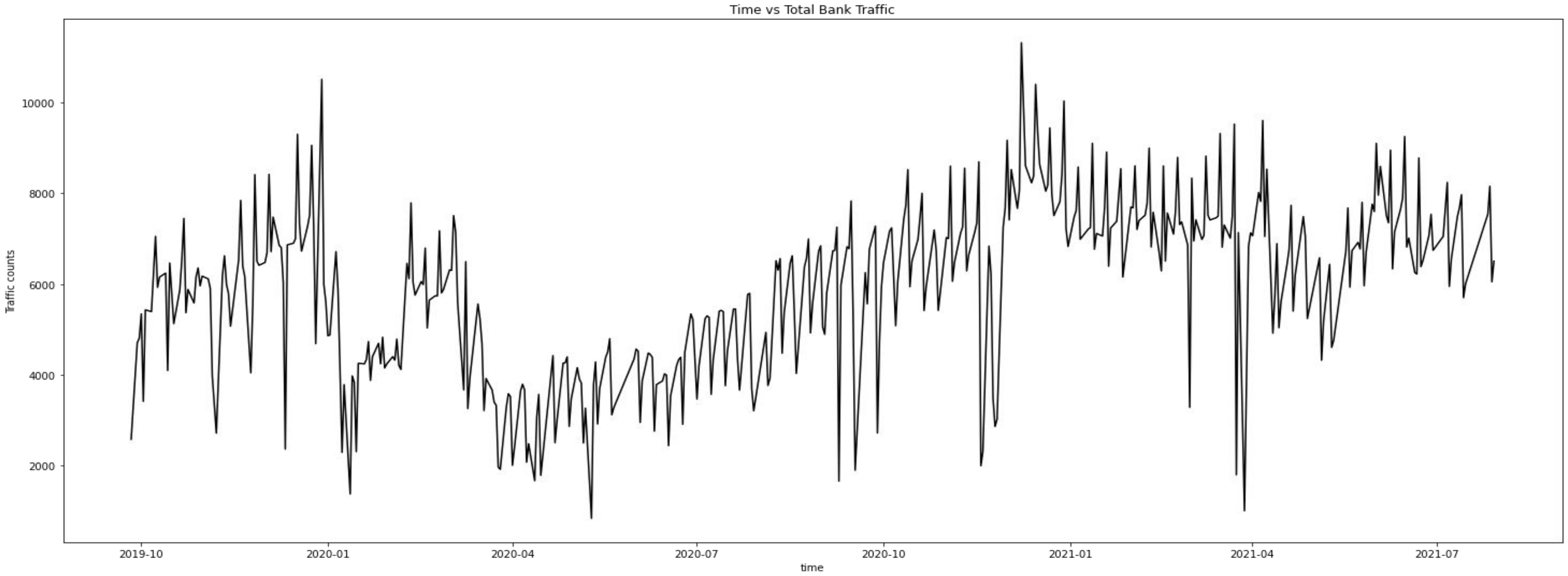
Data Visualizations

- Traffic Data with Holidays



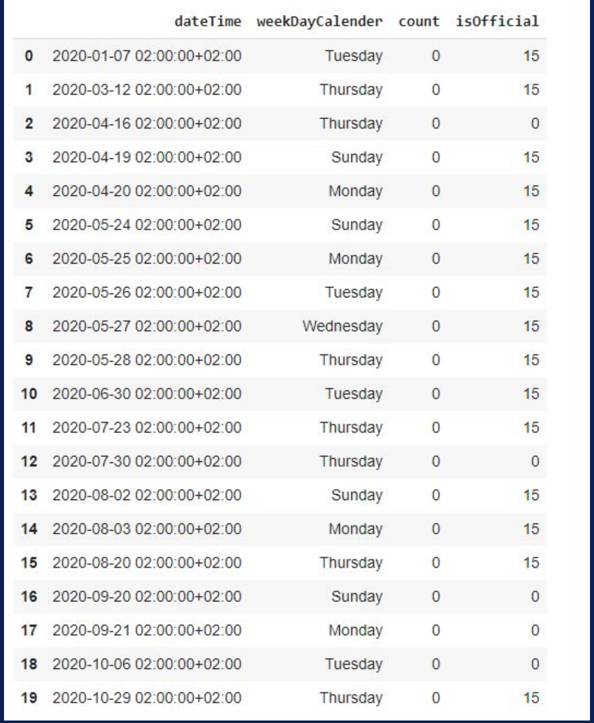
Data Visualizations

- Traffic Data with No Holidays



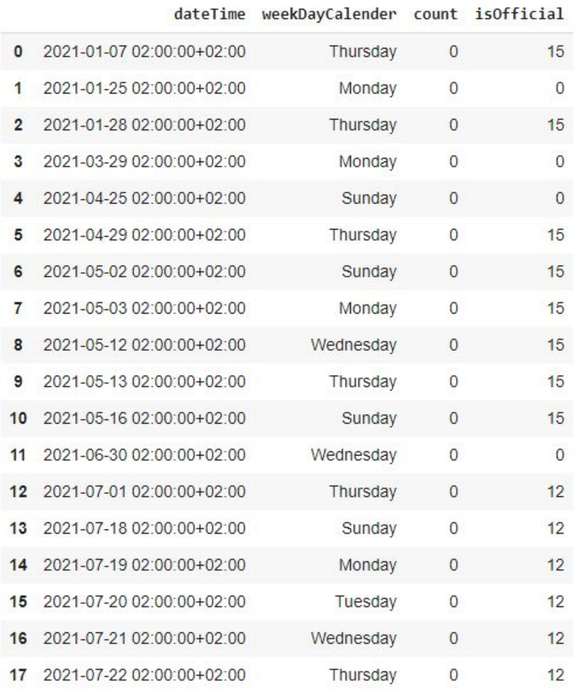
Holiday Analysis

2020



	dateTime	weekDayCalender	count	isOfficial
0	2020-01-07 02:00:00+02:00	Tuesday	0	15
1	2020-03-12 02:00:00+02:00	Thursday	0	15
2	2020-04-16 02:00:00+02:00	Thursday	0	0
3	2020-04-19 02:00:00+02:00	Sunday	0	15
4	2020-04-20 02:00:00+02:00	Monday	0	15
5	2020-05-24 02:00:00+02:00	Sunday	0	15
6	2020-05-25 02:00:00+02:00	Monday	0	15
7	2020-05-26 02:00:00+02:00	Tuesday	0	15
8	2020-05-27 02:00:00+02:00	Wednesday	0	15
9	2020-05-28 02:00:00+02:00	Thursday	0	15
10	2020-06-30 02:00:00+02:00	Tuesday	0	15
11	2020-07-23 02:00:00+02:00	Thursday	0	15
12	2020-07-30 02:00:00+02:00	Thursday	0	0
13	2020-08-02 02:00:00+02:00	Sunday	0	15
14	2020-08-03 02:00:00+02:00	Monday	0	15
15	2020-08-20 02:00:00+02:00	Thursday	0	15
16	2020-09-20 02:00:00+02:00	Sunday	0	0
17	2020-09-21 02:00:00+02:00	Monday	0	0
18	2020-10-06 02:00:00+02:00	Tuesday	0	0
19	2020-10-29 02:00:00+02:00	Thursday	0	15

2021



	dateTime	weekDayCalender	count	isOfficial
0	2021-01-07 02:00:00+02:00	Thursday	0	15
1	2021-01-25 02:00:00+02:00	Monday	0	0
2	2021-01-28 02:00:00+02:00	Thursday	0	15
3	2021-03-29 02:00:00+02:00	Monday	0	0
4	2021-04-25 02:00:00+02:00	Sunday	0	0
5	2021-04-29 02:00:00+02:00	Thursday	0	15
6	2021-05-02 02:00:00+02:00	Sunday	0	15
7	2021-05-03 02:00:00+02:00	Monday	0	15
8	2021-05-12 02:00:00+02:00	Wednesday	0	15
9	2021-05-13 02:00:00+02:00	Thursday	0	15
10	2021-05-16 02:00:00+02:00	Sunday	0	15
11	2021-06-30 02:00:00+02:00	Wednesday	0	0
12	2021-07-01 02:00:00+02:00	Thursday	0	12
13	2021-07-18 02:00:00+02:00	Sunday	0	12
14	2021-07-19 02:00:00+02:00	Monday	0	12
15	2021-07-20 02:00:00+02:00	Tuesday	0	12
16	2021-07-21 02:00:00+02:00	Wednesday	0	12
17	2021-07-22 02:00:00+02:00	Thursday	0	12

Whoa!

It is Present in the data that:

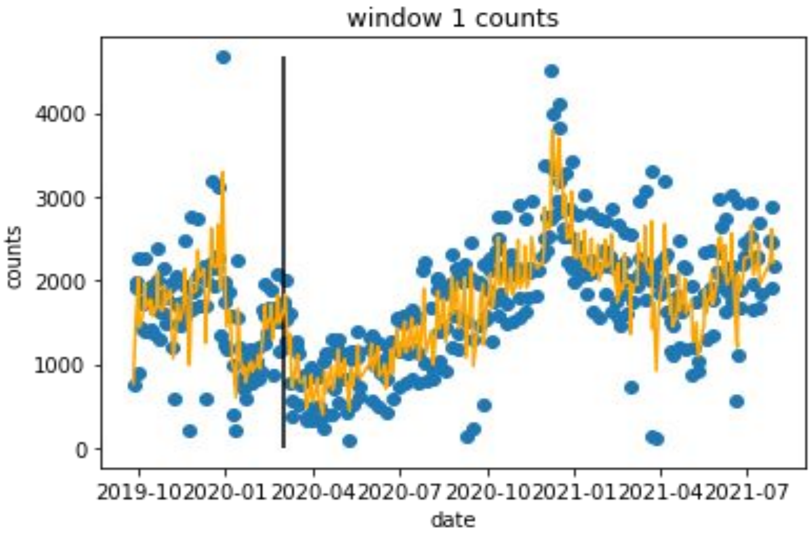
- Weekends always have 0 Traffic
- Official Holidays have 0 Traffic
- Not official Holidays?

Data Visualizations

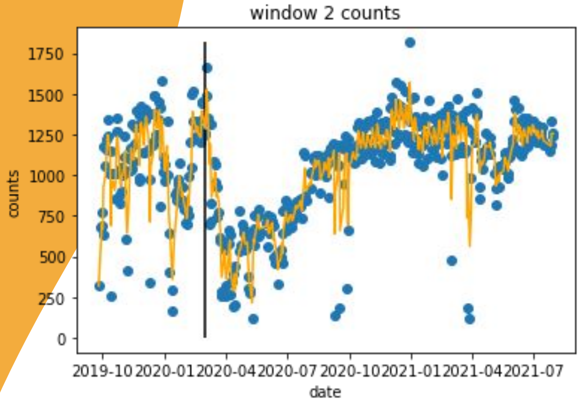
- Exponential Moving Avg Analysis

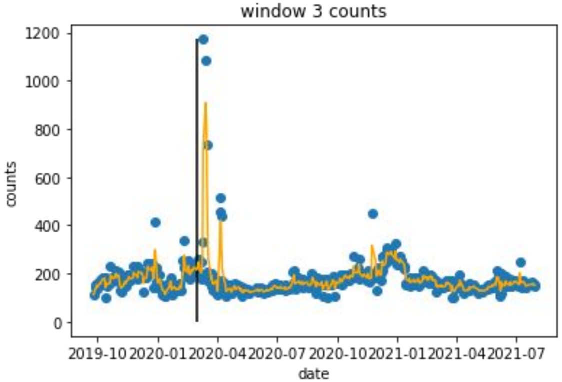
- The trend for each window was investigated. An example is shown.

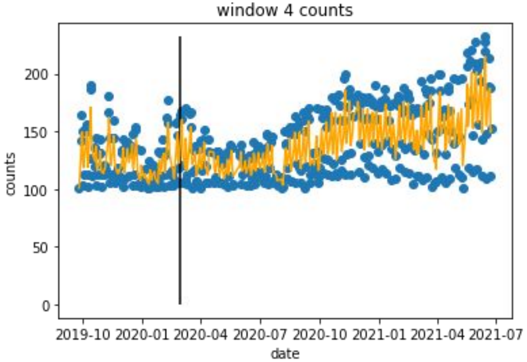
- We marked 01-03-2020 as the separator between pre-covid and covid time in Egypt.

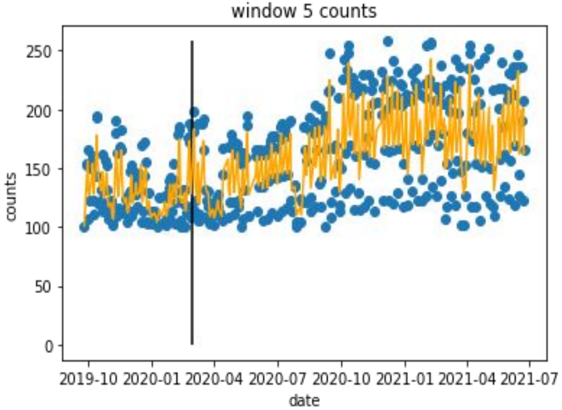


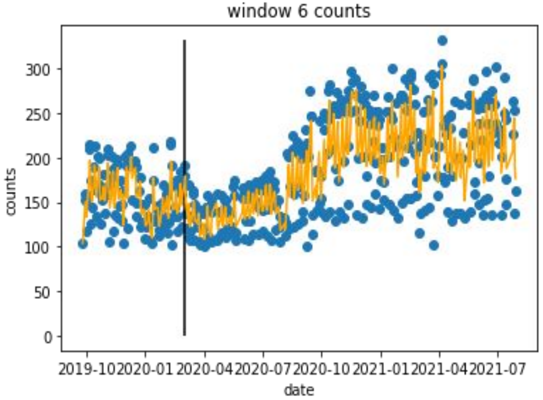
- Exponential Moving Avg Analysis

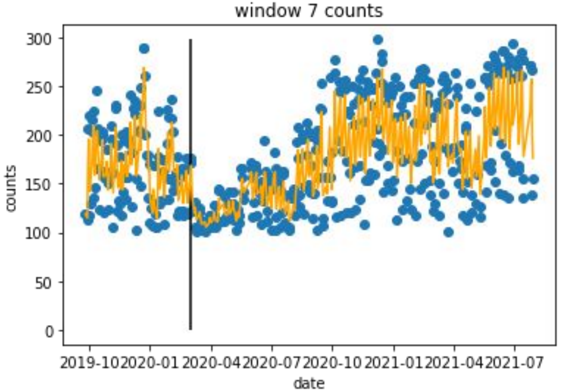




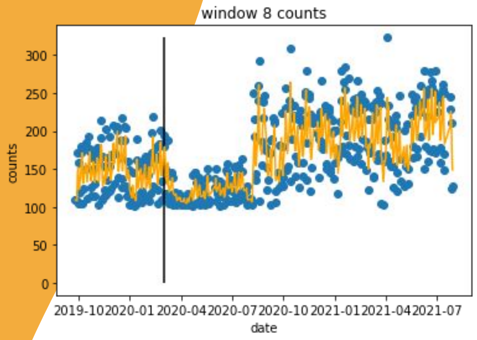


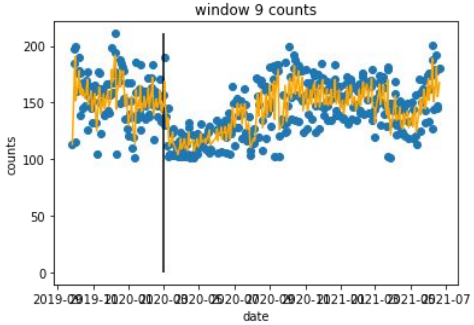


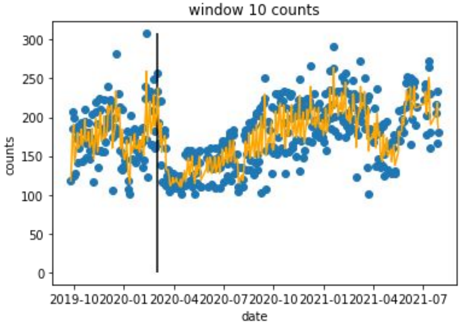


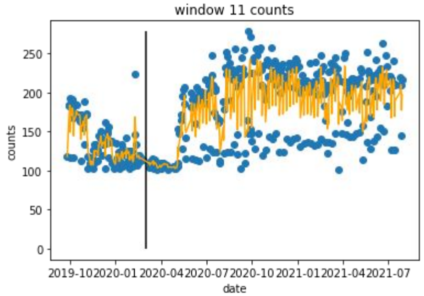


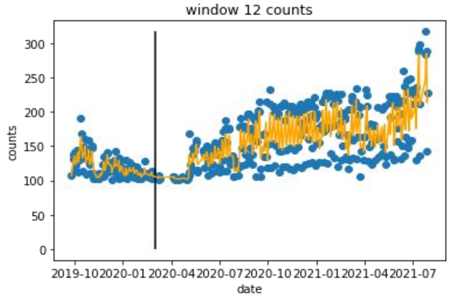
- Exponential Moving Avg Analysis

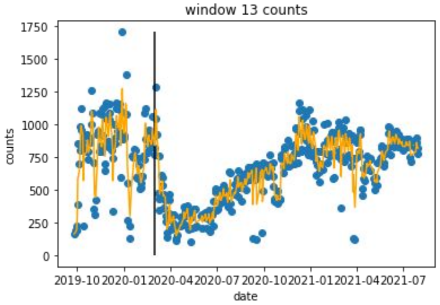


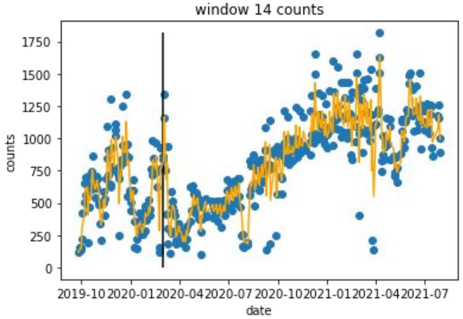


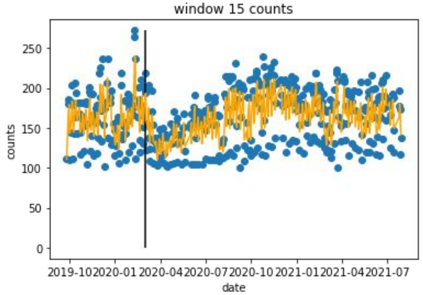






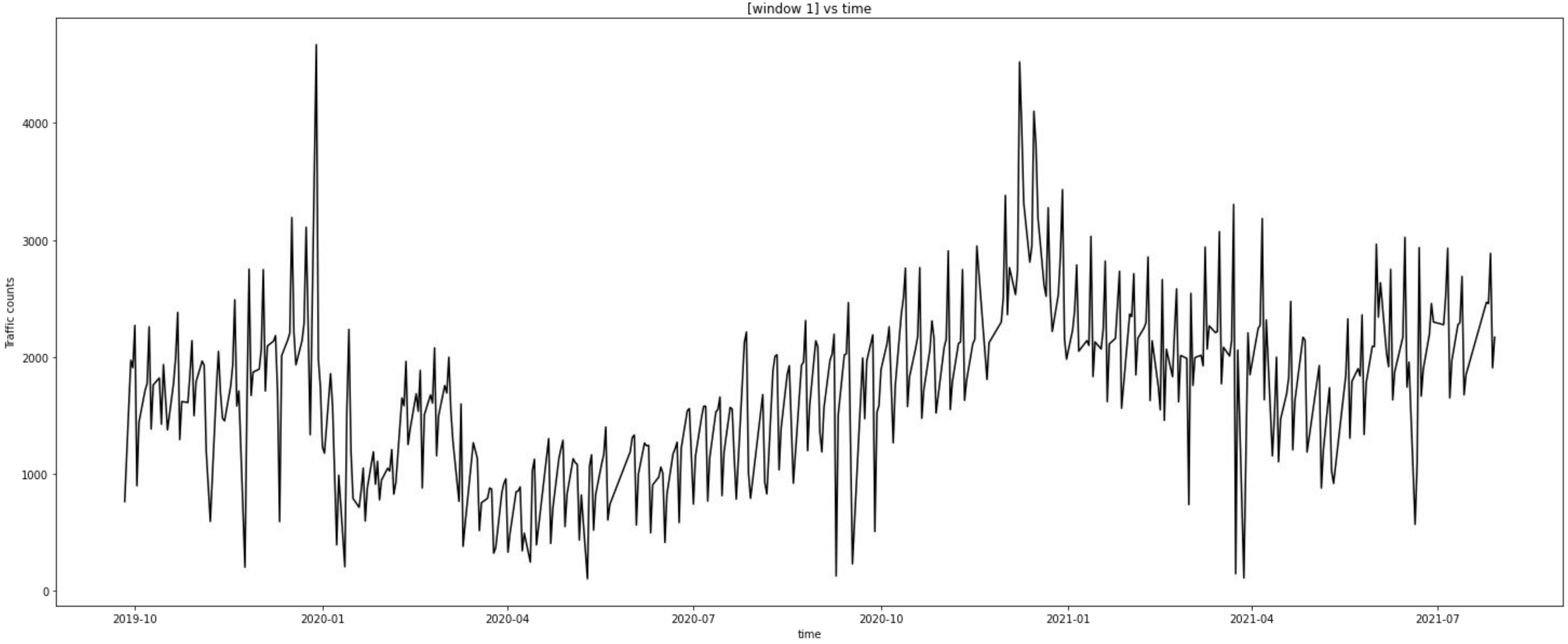




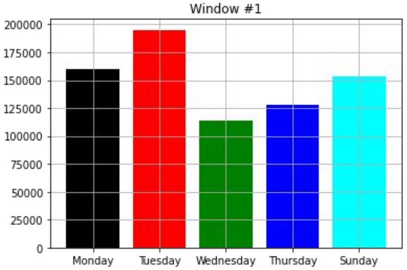


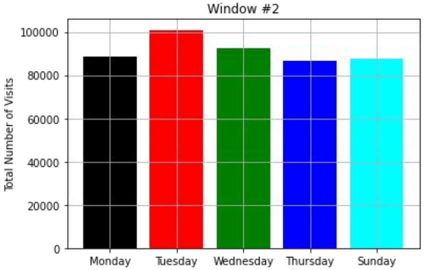
- Windows Vs Time Analysis

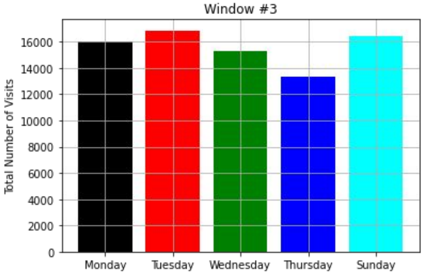
✓ **Similar trends are captured using simple line plots as well.**

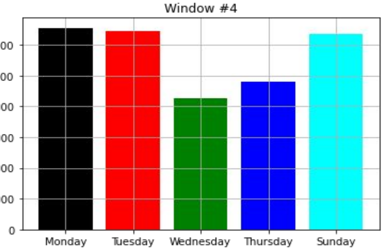


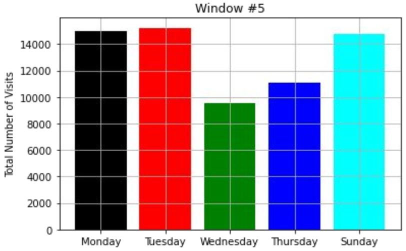
Window vs Week Day

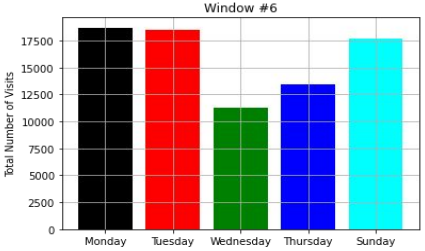


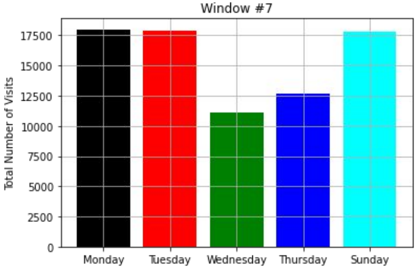


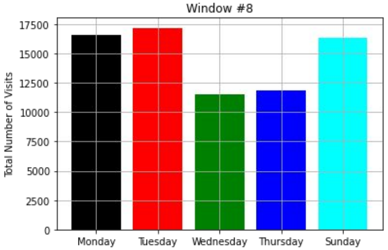


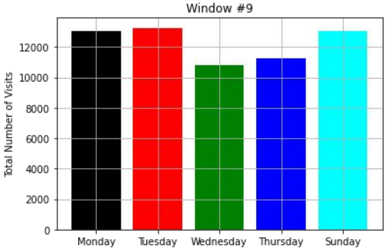




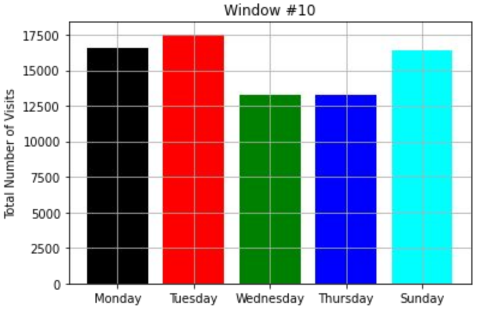


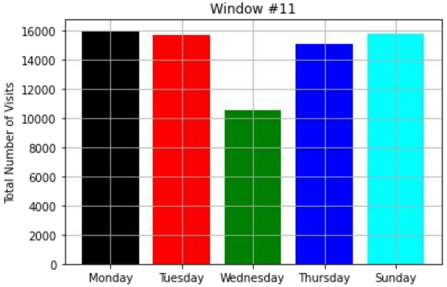


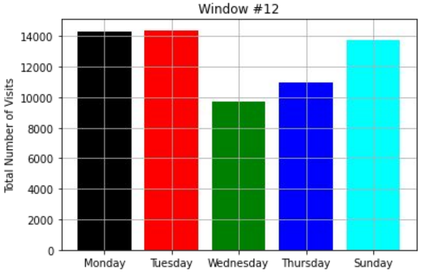


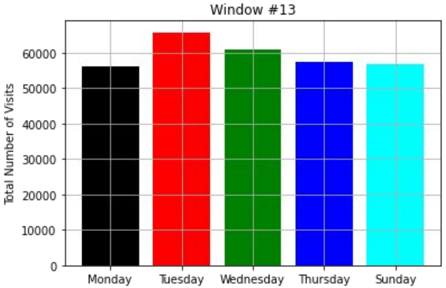


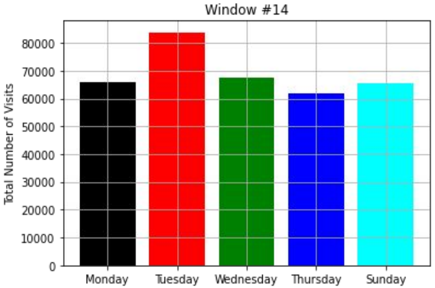
Window vs Week Day

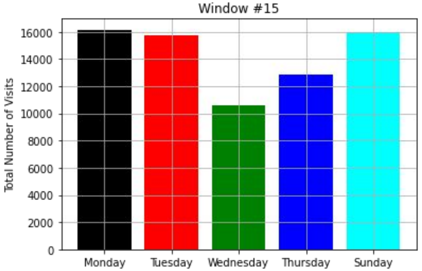








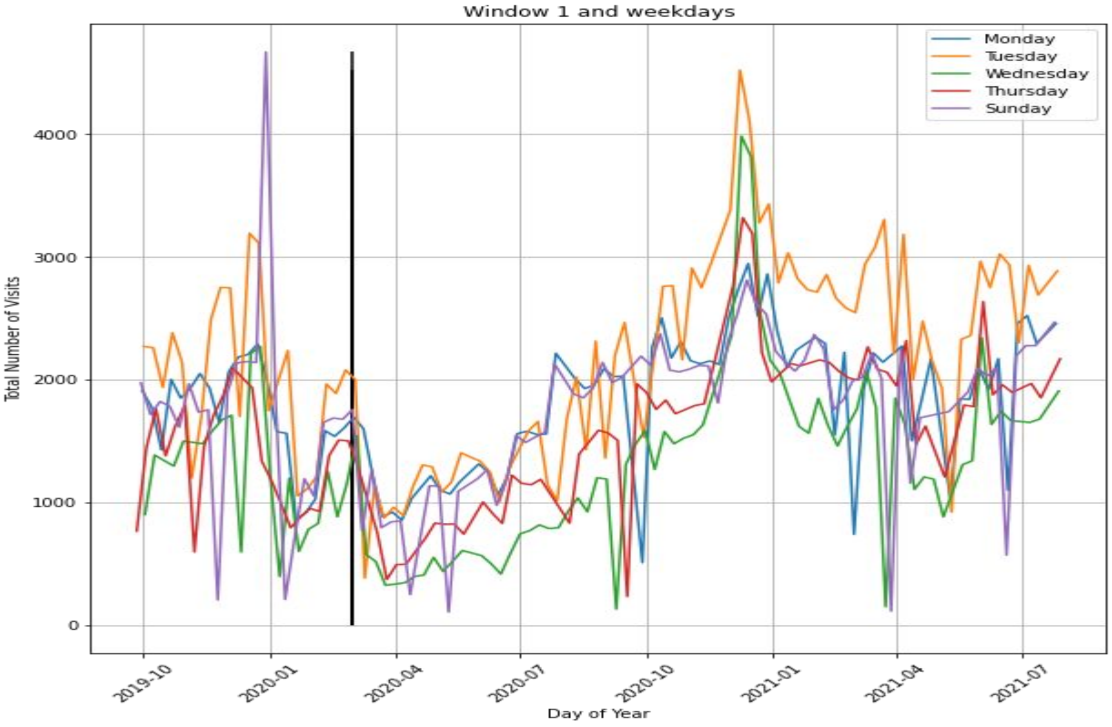




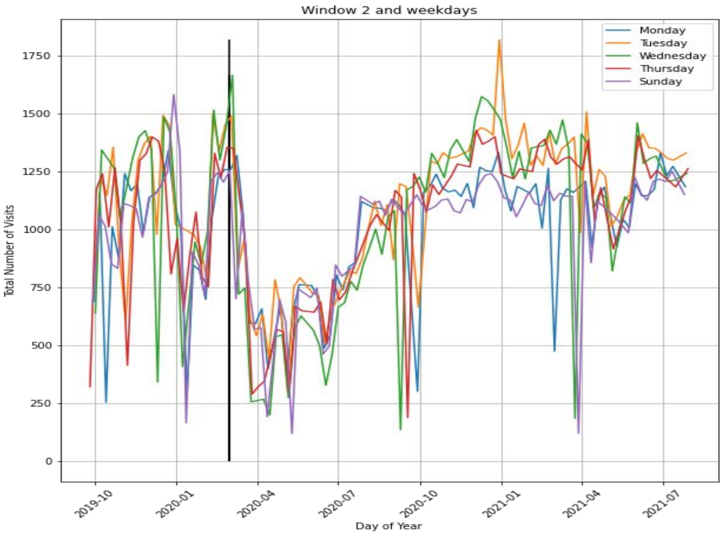
Windows vs Week day trends

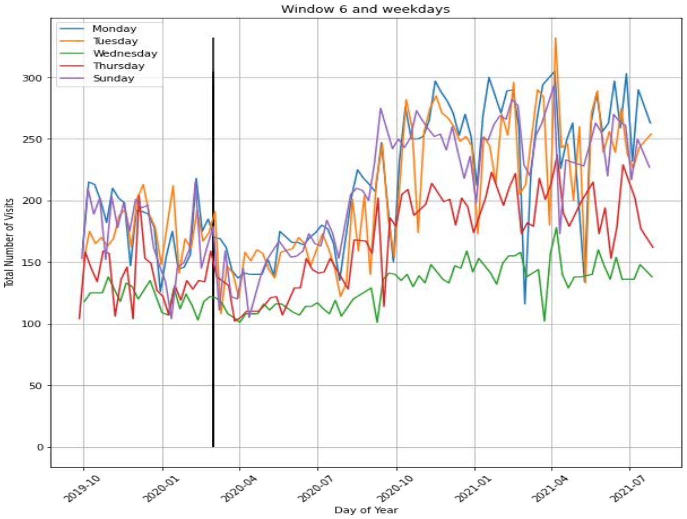
- The shown Trend Analysis was done on all windows. An example is shown.

- We marked 01-03-2020 as the separator between pre-covid and covid time in Egypt.



Windows vs Week day trends





- **Tuesdays** represent the crest of traffic
- **Wednesdays** represent the trough of traffic

03

Methodology

In our Analysis, we noticed a variation in the traffic for each Teller window. Hence, We decided to build a model to estimate the Traffic at each window. Then, through summing them we get our Traffic Estimate for the Day.

04

Predictive Analytics

Model Selection

We Trained Our Models on the Entire Series* and tested on July 2021

PROPHET

Darts

AutoTS

Simplest Model

1. Combining the results from all Windows
2. Fitting a Prophet Model
3. Predict :)

3000+ person

Mean Absolute Error :')

AutoARIMA

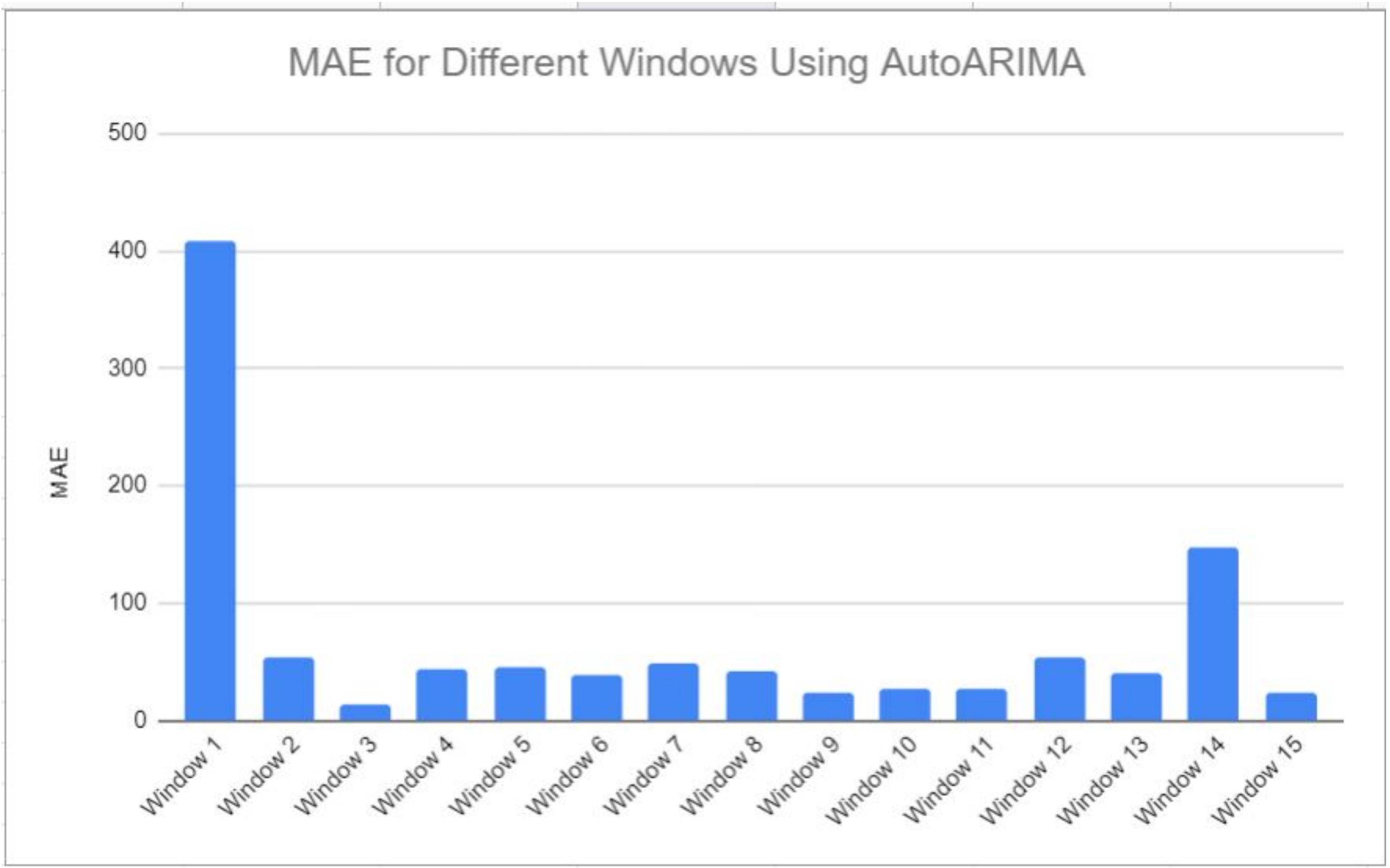
AR: Autoregression. A model that uses the dependent relationship between an observation and some number of lagged observations.

I: Integrated. The use of differencing of raw observations (e.g. subtracting an observation from an observation at the previous time step) in order to make the time series stationary.

MA: Moving Average. A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Darts

https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/


MAE for Different Windows Using AutoARIMA
MAE
500
400
300
200
100
0
Window 1
Window 2
Window 3
Window 4
Window 5
Window 6
Window 7
Window 8
Window 9
Window 10
Window 11
Window 12
Window 13
Window 14
Window 15

1042 person

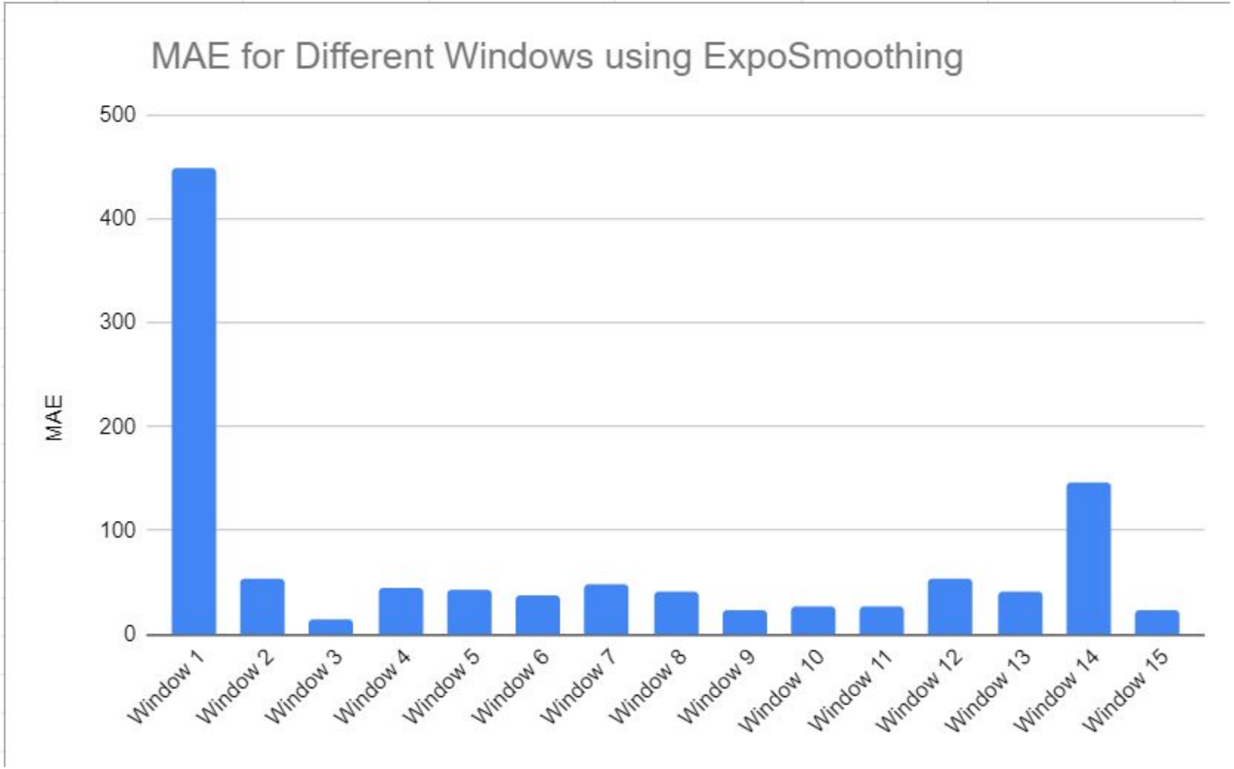
Mean Absolute Error

Exponential Smoothing

Although simple Averaging Methods assign Equal Weights for past observations, exponential functions are used to assign exponentially decreasing weights over time. *

https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/

Result


MAE for Different Windows using ExpoSmoothing
MAE
500
400
300
200
100
0
Window 1
Window 2
Window 3
Window 4
Window 5
Window 6
Window 7
Window 8
Window 9
Window 10
Window 11
Window 12
Window 13
Window 14
Window 15

1270 person

Mean Absolute Error

Prophet

Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects.

Prophet is open source software released by Facebook’s Core Data Science team.

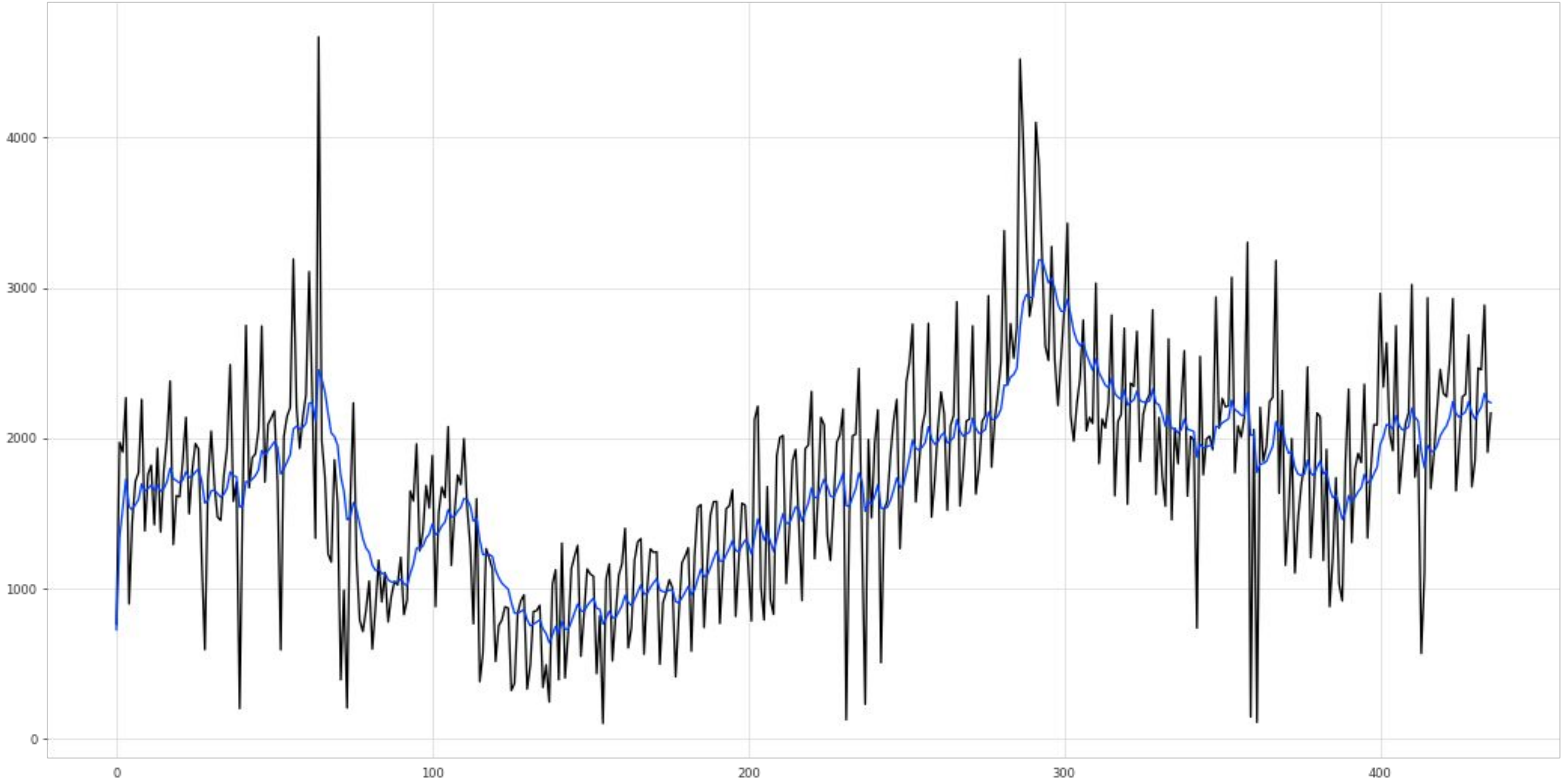
Prophet is robust to outliers, missing data, and dramatic changes in your time series.

PROPHET

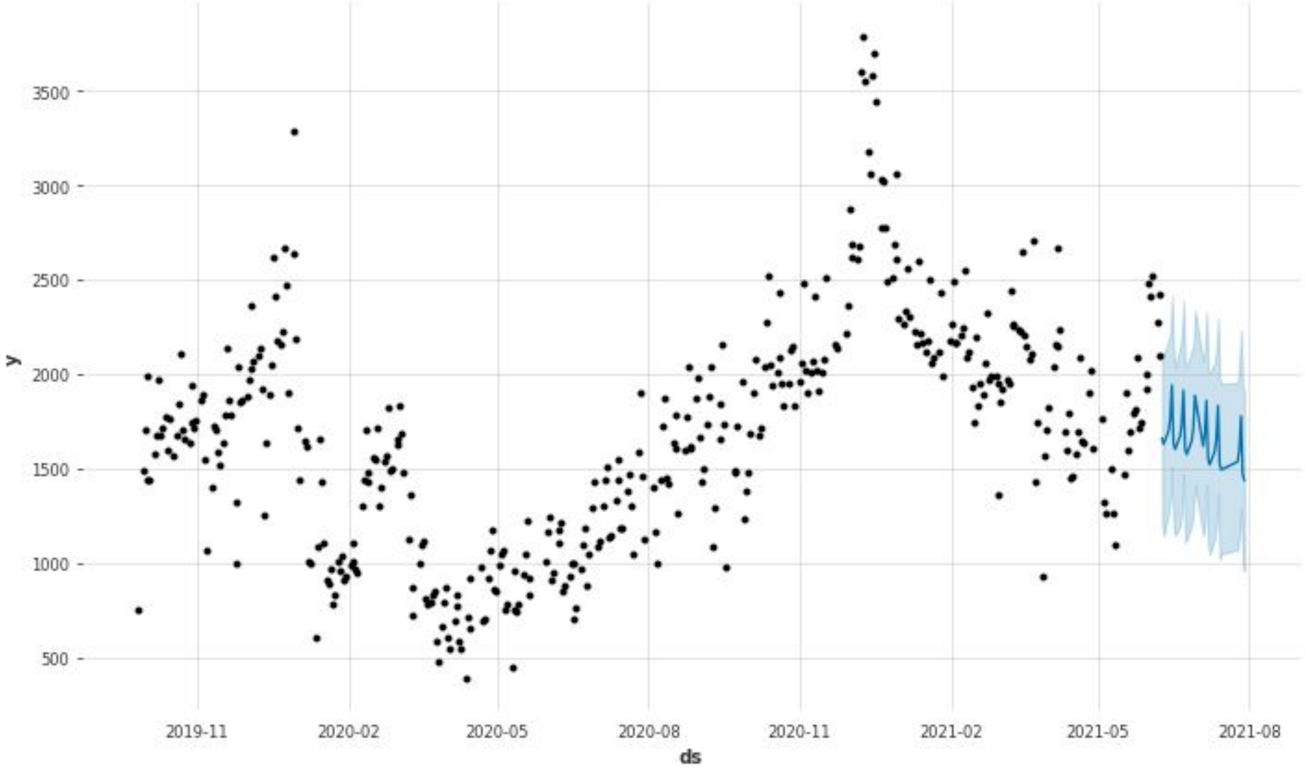
https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/

What’s Wrong with Window 1 ?

Kalman Filtering



Results

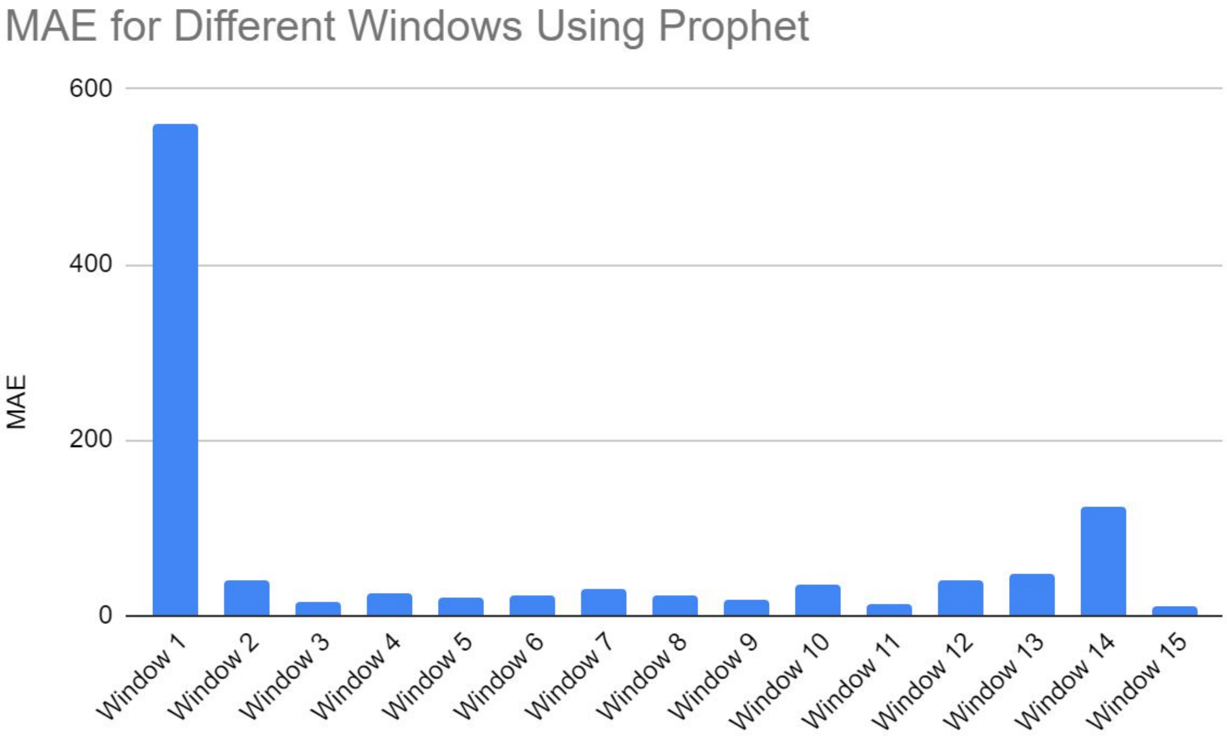


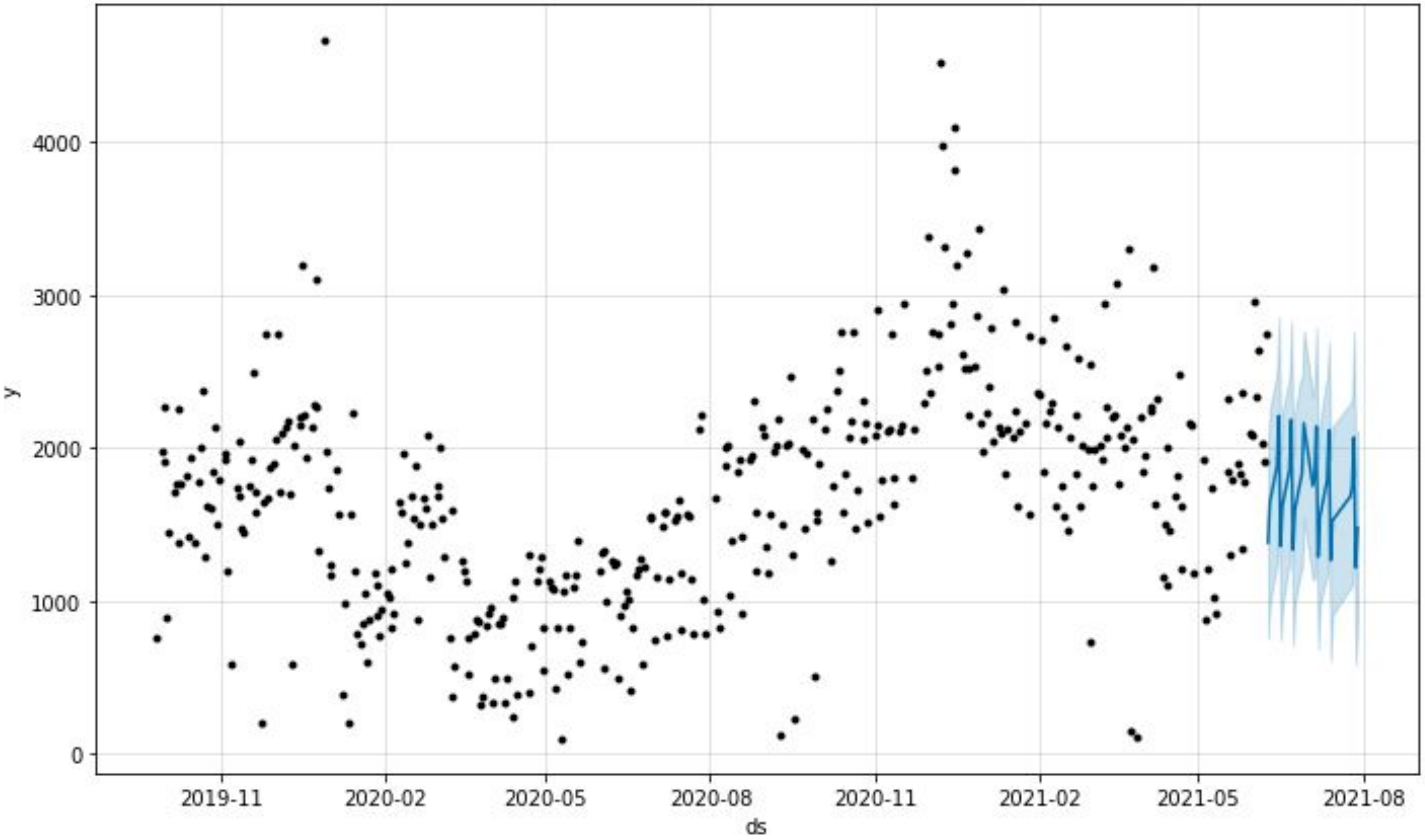
y
3500
3000
2500
2000
1500
1000
500
2019-11
2020-02
2020-05
2020-08
2020-11
2021-02
2021-05
2021-08
ds

1700+ person

Mean Absolute Error

Result

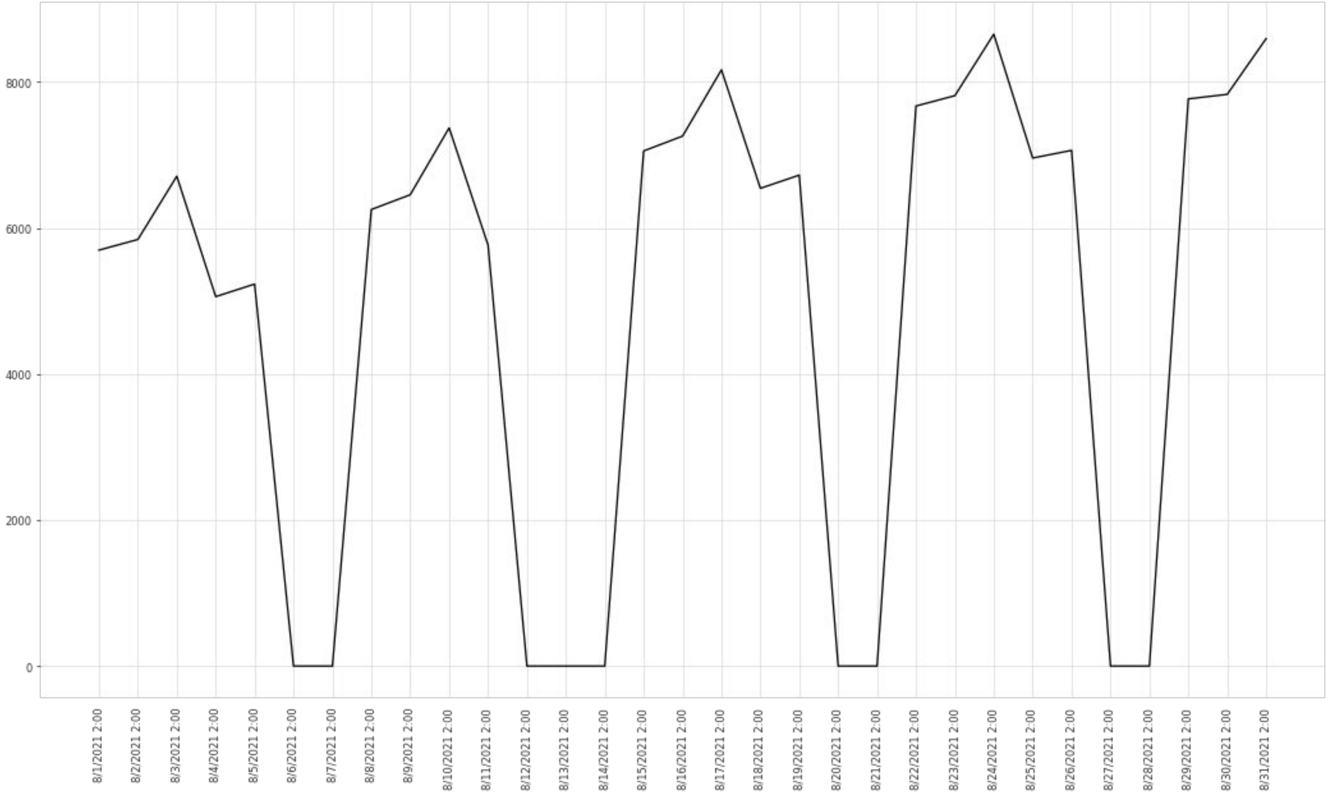

MAE for Different Windows Using Prophet
600
400
200
0
MAE
Window 1
Window 2
Window 3
Window 4
Window 5
Window 6
Window 7
Window 8
Window 9
Window 10
Window 11
Window 12
Window 13
Window 14
Window 15


y
4000
3000
2000
1000
0
2019-11
2020-02
2020-05
2020-08
2020-11
2021-02
2021-05
2021-08
ds

930 person

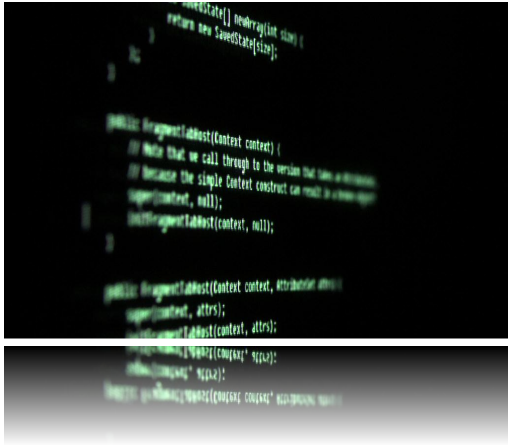
Mean Absolute Error

Predicted Traffic



Script Link:

Synapse Hackathon.ipynb - Colaboratory (google.com)



Thanks!