



Traffic Forecast Prediction

ADVANCED DATA SCIENCE CAPSTONE PROJECT

Contents

- ▶ Introduction (Use Case | Business Value)
- ▶ Data Exploration
- ▶ SARIMA(X)
- ▶ LSTM
- ▶ Results / Outlook

Use Case | Business Value

▶ Use Case

- ▶ Visits Traffic Forecast for a News Media Website
- ▶ Multi-Step Time Series Prediction required
- ▶ Challenges for predicting Visits of a New Media Website

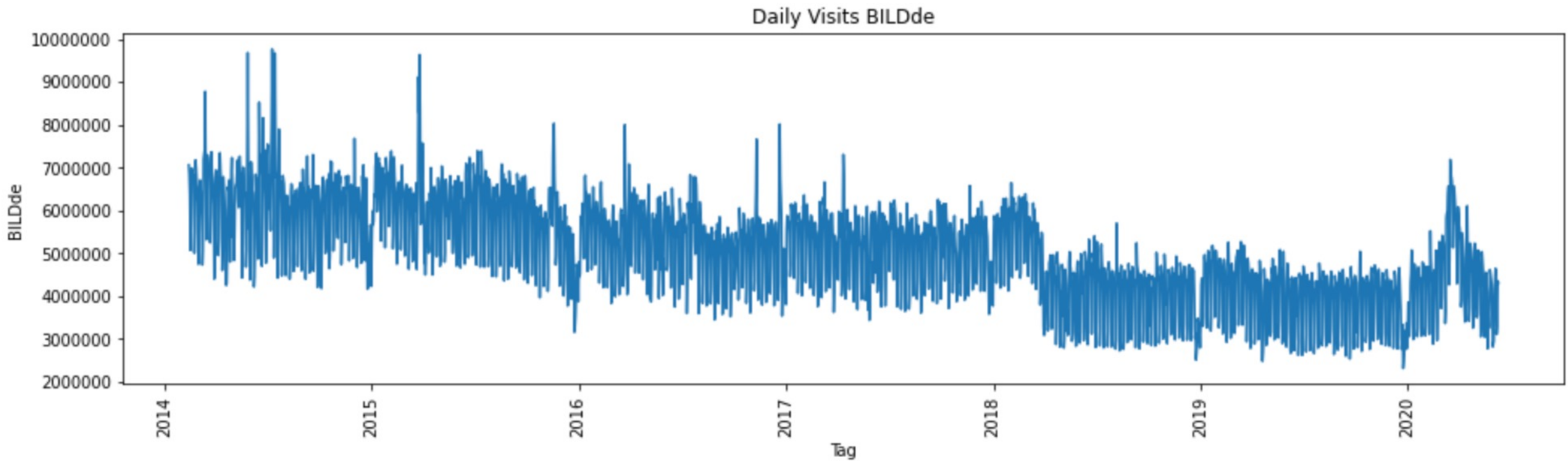
▶ Business Value

- ▶ Forecast to receive estimated revenue from the Marketer
- ▶ Automate the Traffic Forecast process to be more efficient

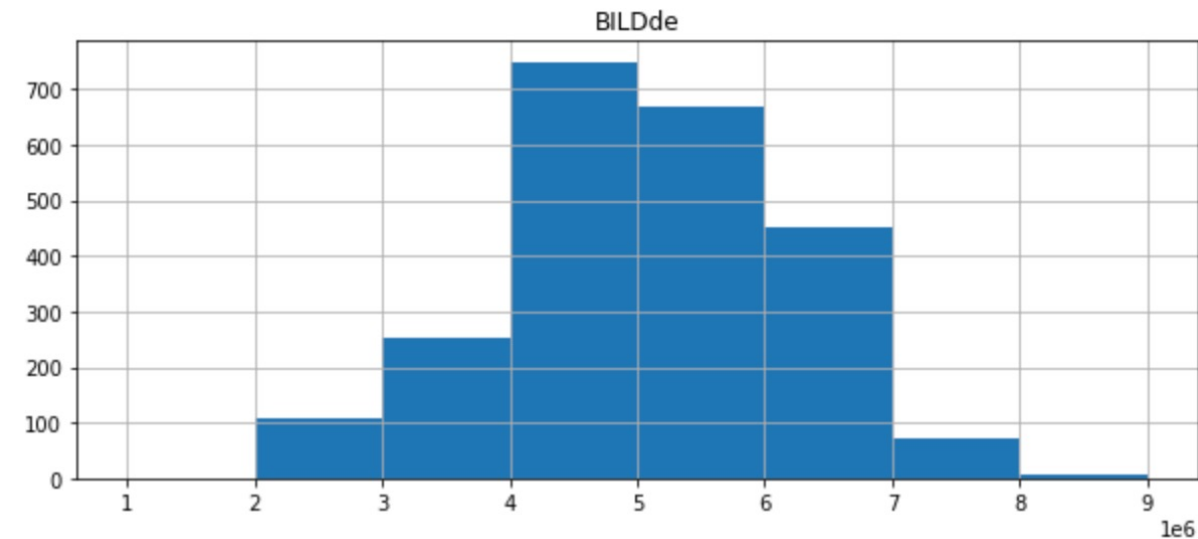
Data Exploration

- ▶ Checking for Missing Values
- ▶ Checking if values range within one/two Std. Deviations
- ▶ Checking for Outliers
- ▶ Checking for Stationary, Trend and Seasonality

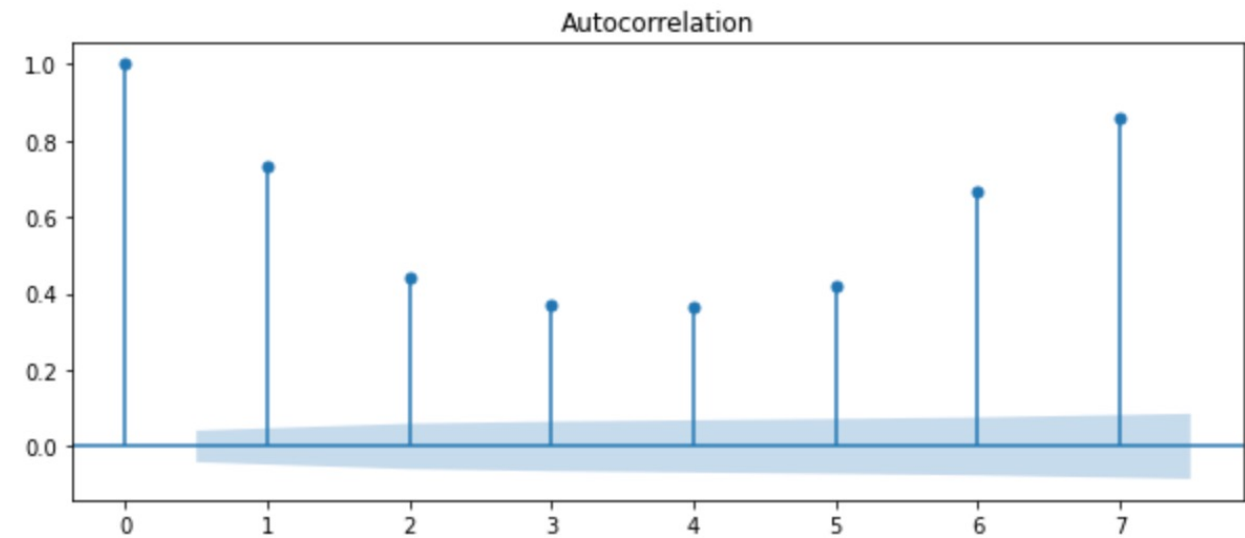
Data Exploration - Overview



Data Exploration – Outlier, Stationary

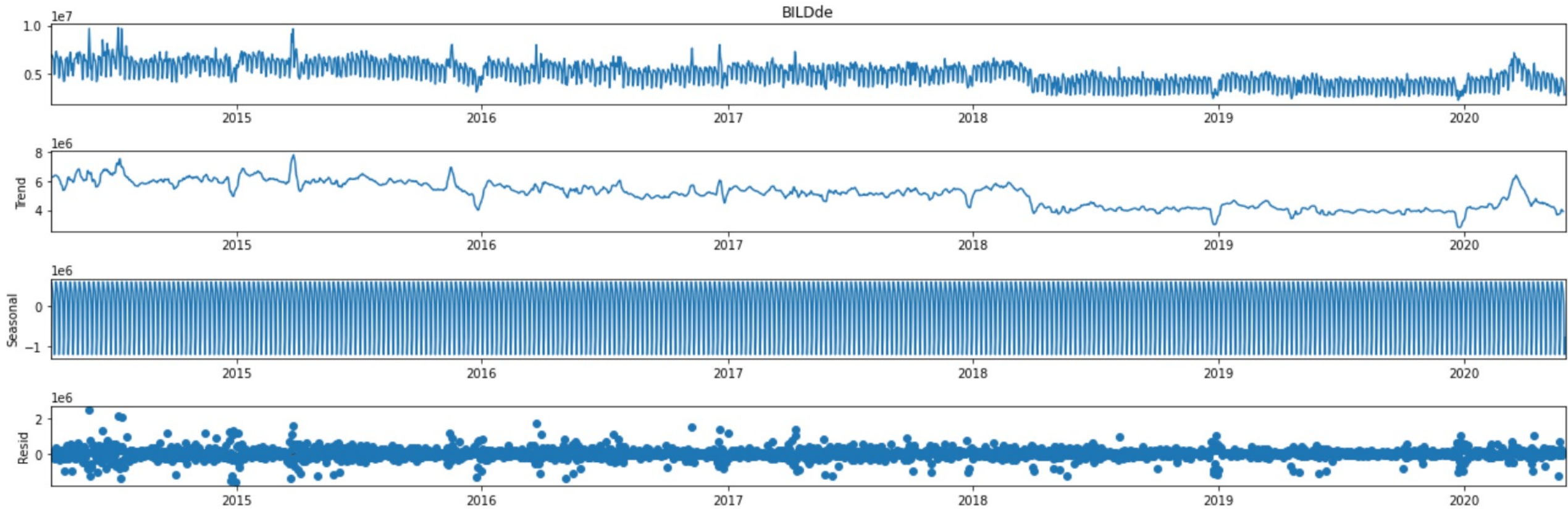


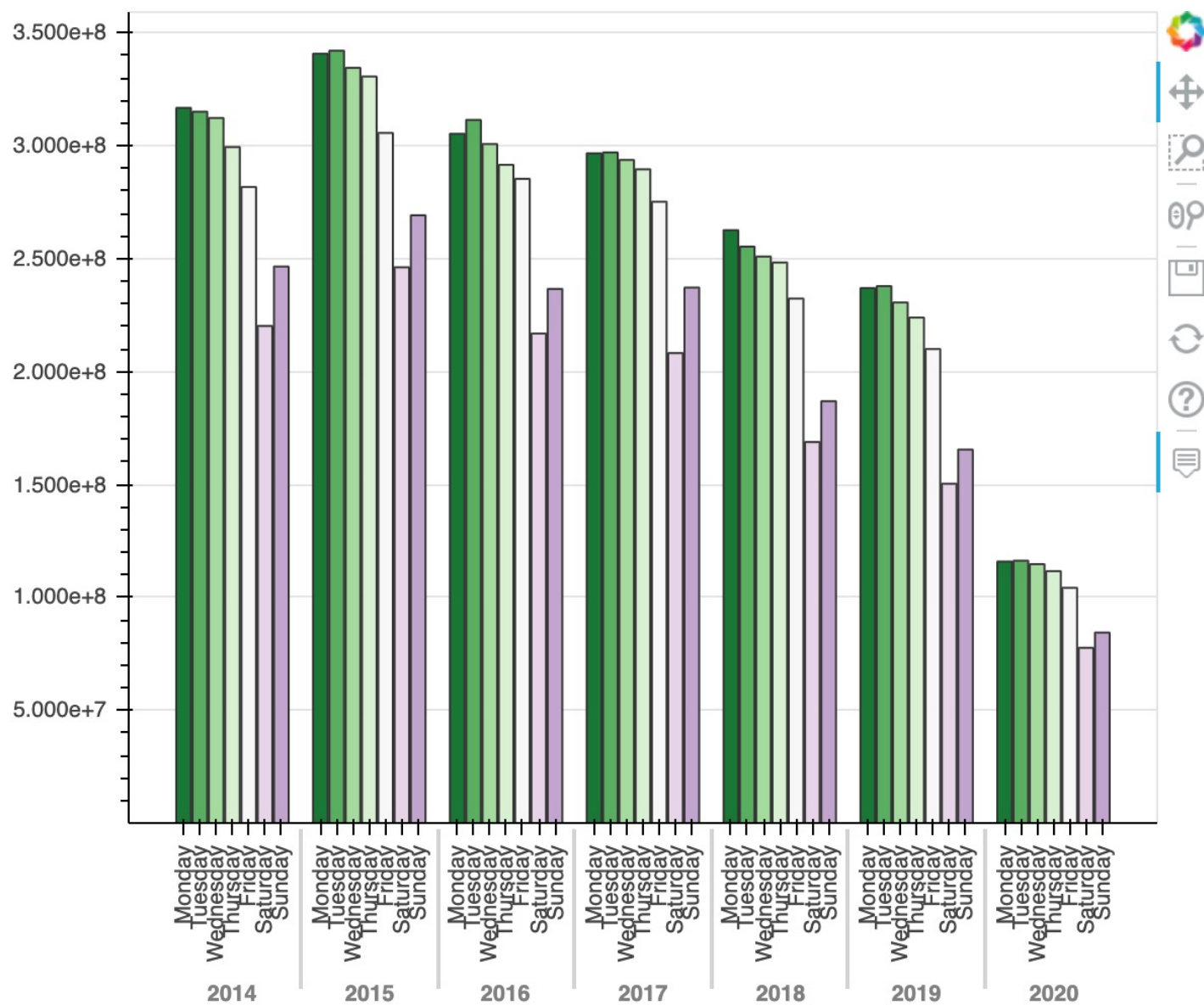
- Outlier in range: Christmas, 5% Traffic low bucket (2-3 Mio.), only a few with (7-9 Mio.) Visits



Adfuller Test: $p\text{-value} > 0.01$ ($p=0.27$)
Reject Null Hypothesis that Time Series is stationary

Data Exploration – Trend, Seasonality,





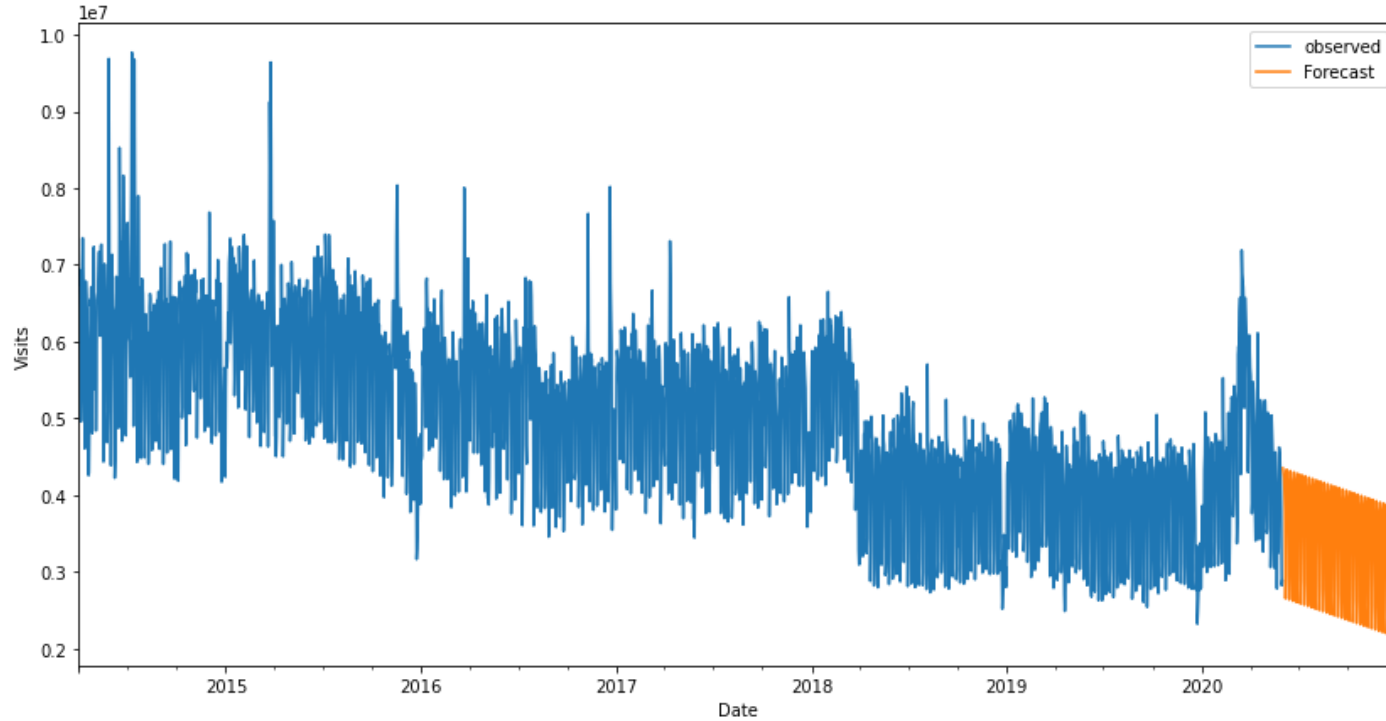
Data Exploration – Weekday Seasonality,

Model: SARIMA(X)

► SARIMA

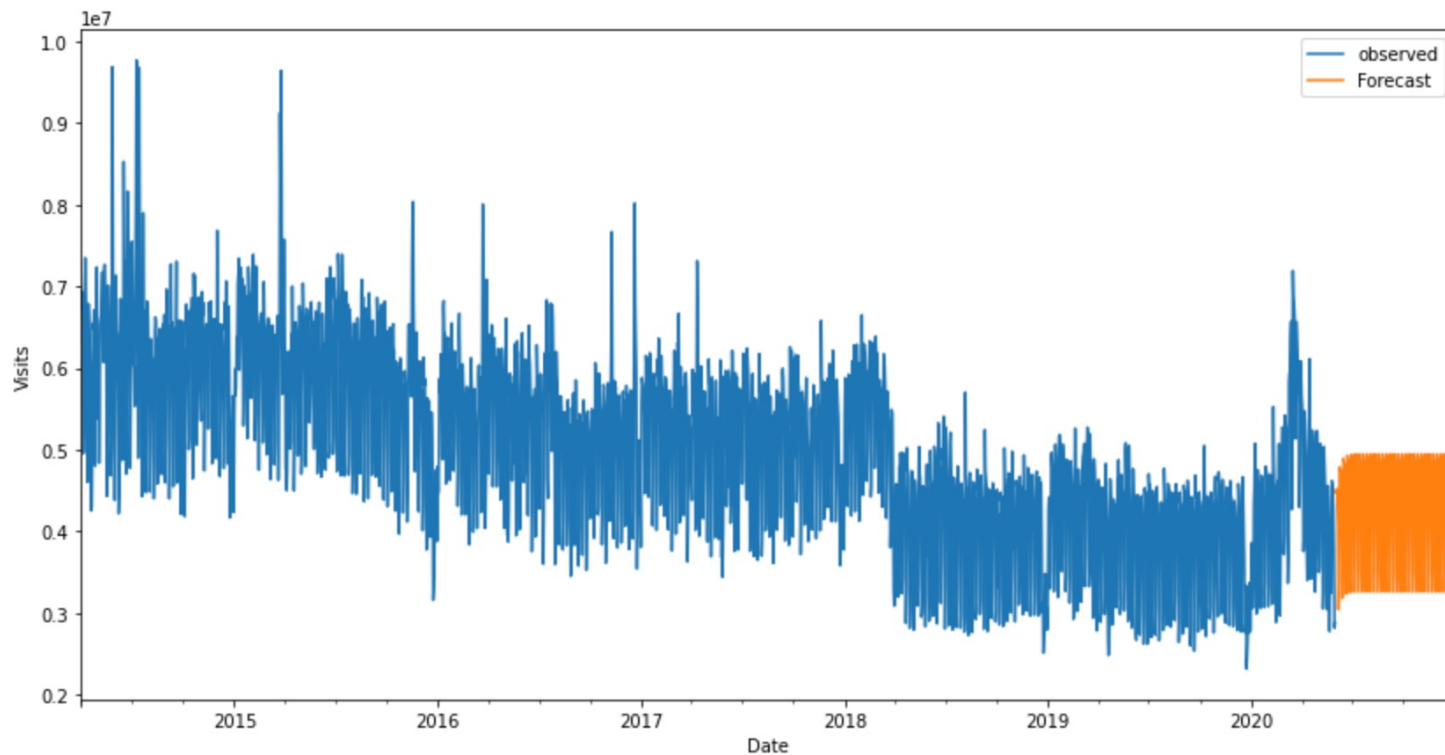
- Seasonal ARIMA Model (Autoregressive integrated moving average)
 - Seasonal → seasonal effect (year, month, week, weekday, etc)
 - AR (autoregressive) → refers to using lagged values of our target variable to make our prediction
 - I (integrated) → differencing to make Times Series stationary
 - MA (moving average) → based on the error of the lagged forecast
- Hyperparameter: $(p,d,q) \times (P,D,Q)S$

Model: SARIMA(X) - Own



- ▶ SARIMA
 - ▶ $(p, d, q) \times (P, D, Q)_S$
 - ▶ $(0, 1, 1) \times (0, 1, 1)_7$
- ▶ R-squared = 0.87

Model: SARIMA(X) – AIC (Grid Search)



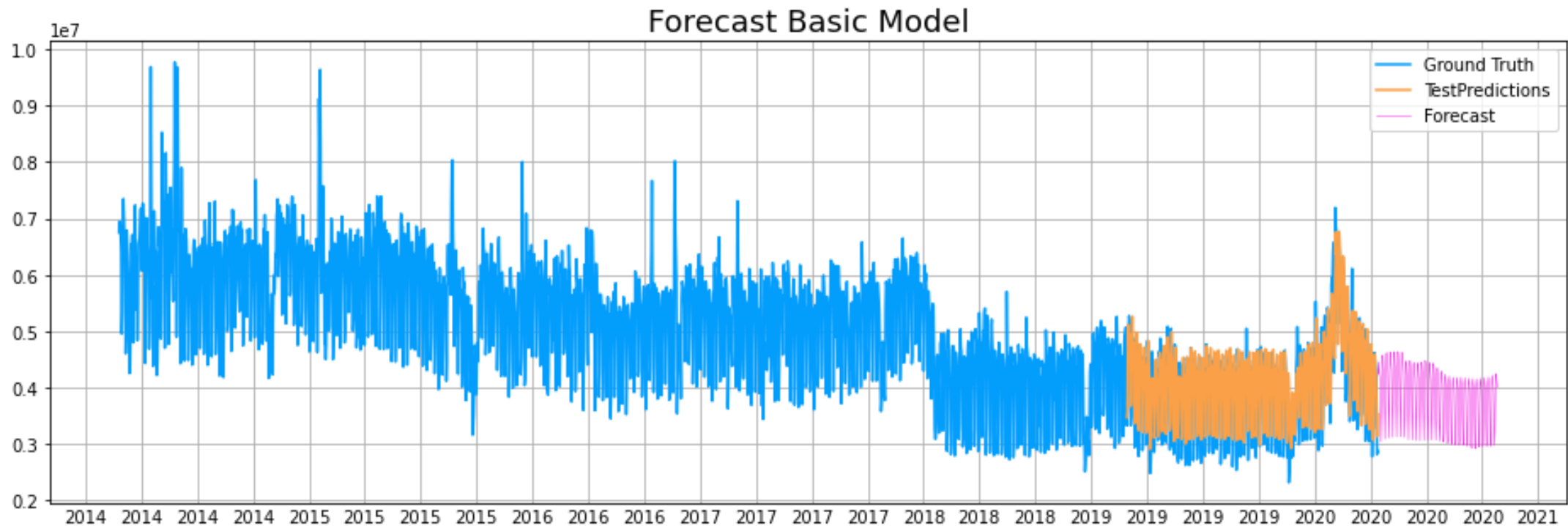
- ▶ SARIMA (AIC)
 - ▶ $(p, d, q) \times (P, D, Q)_S$
 - ▶ $(1, 0, 1) \times (1, 1, 1)_7$
- ▶ R-squared = 0.86

Model: LSTM

► LSTM

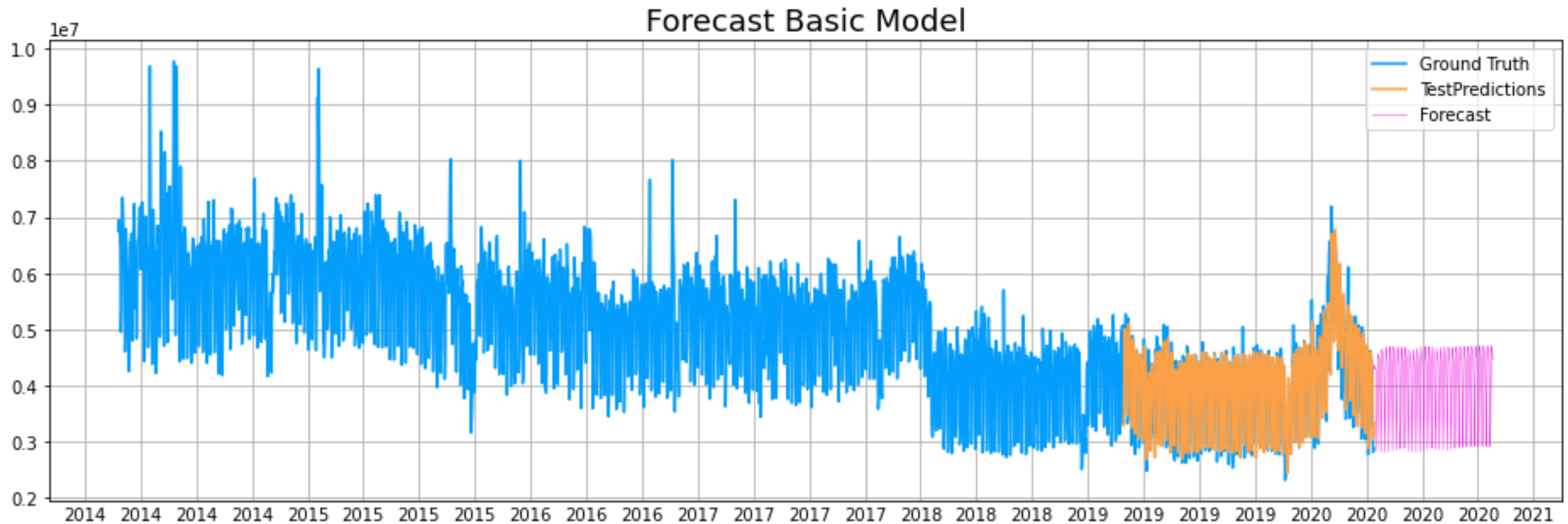
- LSTM stands for Long Short Term Memory. It is a special form of an RNN
- capability to take past events into consideration and only store the relevant information over time
- Hyperparameter:
 - Batch Size (sequences looked at at the same time during learning. It's the limitation of the batches shown to the network before a weight update can be performed)
 - Epochs (number of iterations the algorithm goes through the training dataset for learning purposes)
 - Neurons (capability of the network to learn)

Model: LSTM



$R^2 = 0.83$. This seems to be a good combination of hyper params.

Model: LSTM – Batch Size



$R^2 = 0.86$. This means that the LSTM performed far better than before with this combination of hyper params.

Findings | Evaluation | Outlook

- ▶ Findings
 - ▶ No outliers, decreasing trend, technical change, Seasonality, non-Stationary Series
- ▶ Evaluation
 - ▶ Based on R-squared: SARIMA(X) scored a little better with R-squared → looking at prediction either SARIMA(X) with R-squared = 0.86 or LSTM as recommendation
- ▶ Outlook
 - ▶ Further Testing Sarima(x) with BIC optimisation
 - ▶ stacked LSTM with return_sequence = true
 - ▶ Monthly prediction
 - ▶ Different performance measure
 - ▶ Multivariate models such as a Vector Auto Regression Model