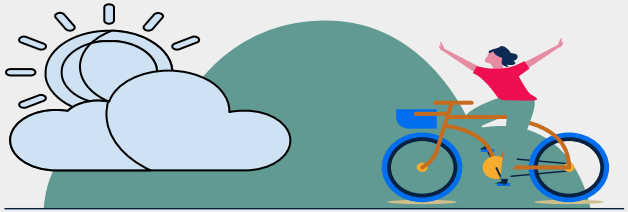
Bike Sharing Demand Prediction for Washington DC

Introduction



Our Prediction & Aim

Forecasting the amount of bike sharing according to weather patterns and calendar events can ensure a sustainable business process.

The Team

- Kristin Engelke
- Megan Leal Causton
- Salil Pachare
- Sebastian Buenker

The Data

Capital Bikeshare Usage Data

Timespan: 2011-2019

Source: Capital Bikeshare system data

All bike rentals from 2011-2019 (approx. 25 million observations).

Data for analysis is aggregated to yield total number of bike rentals per day.

Historic DC Weather Data

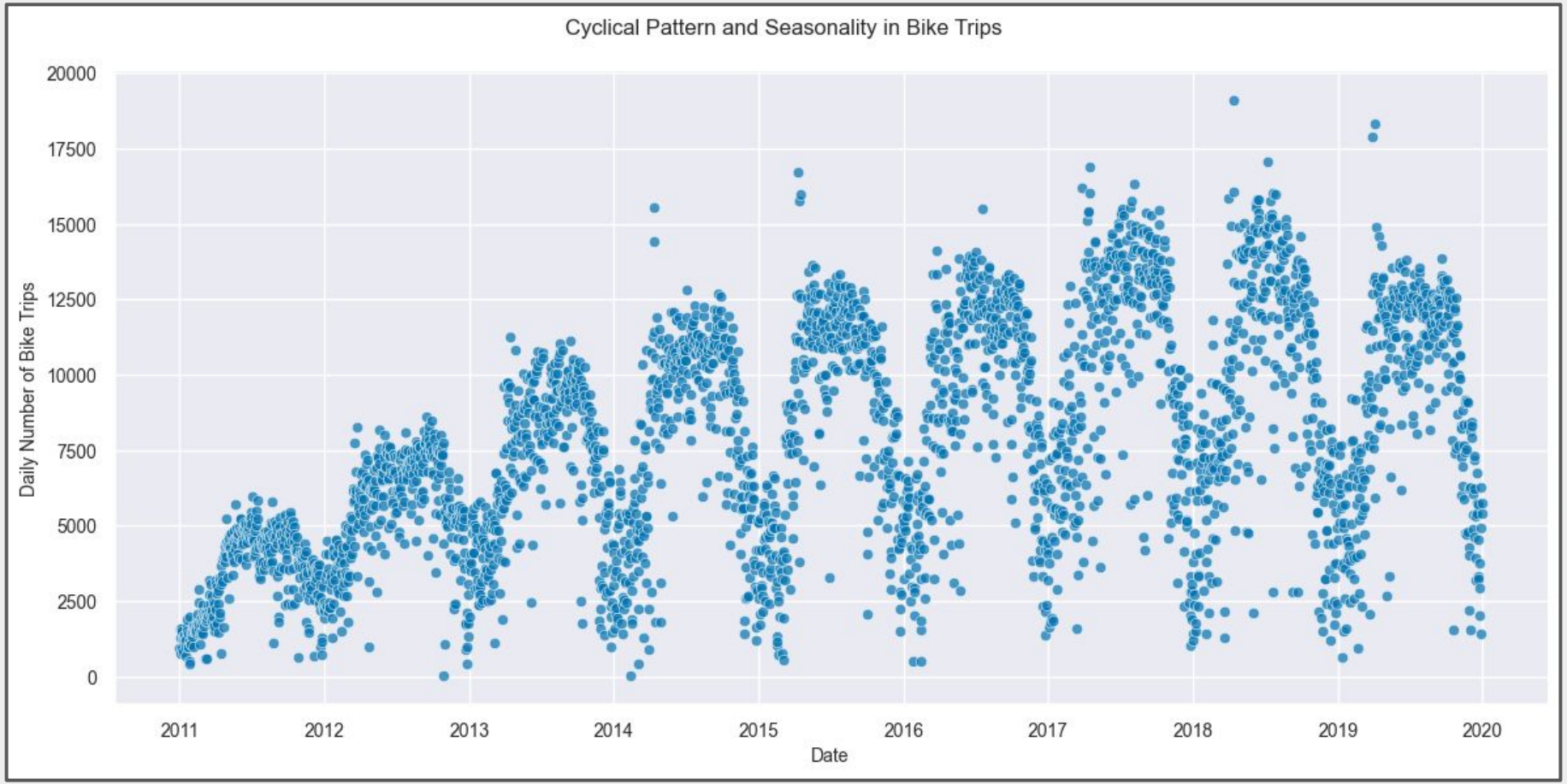
Source: National Oceanic and Atmospheric Administration

Station: Washington Reagan Airport (5 km from Downtown)

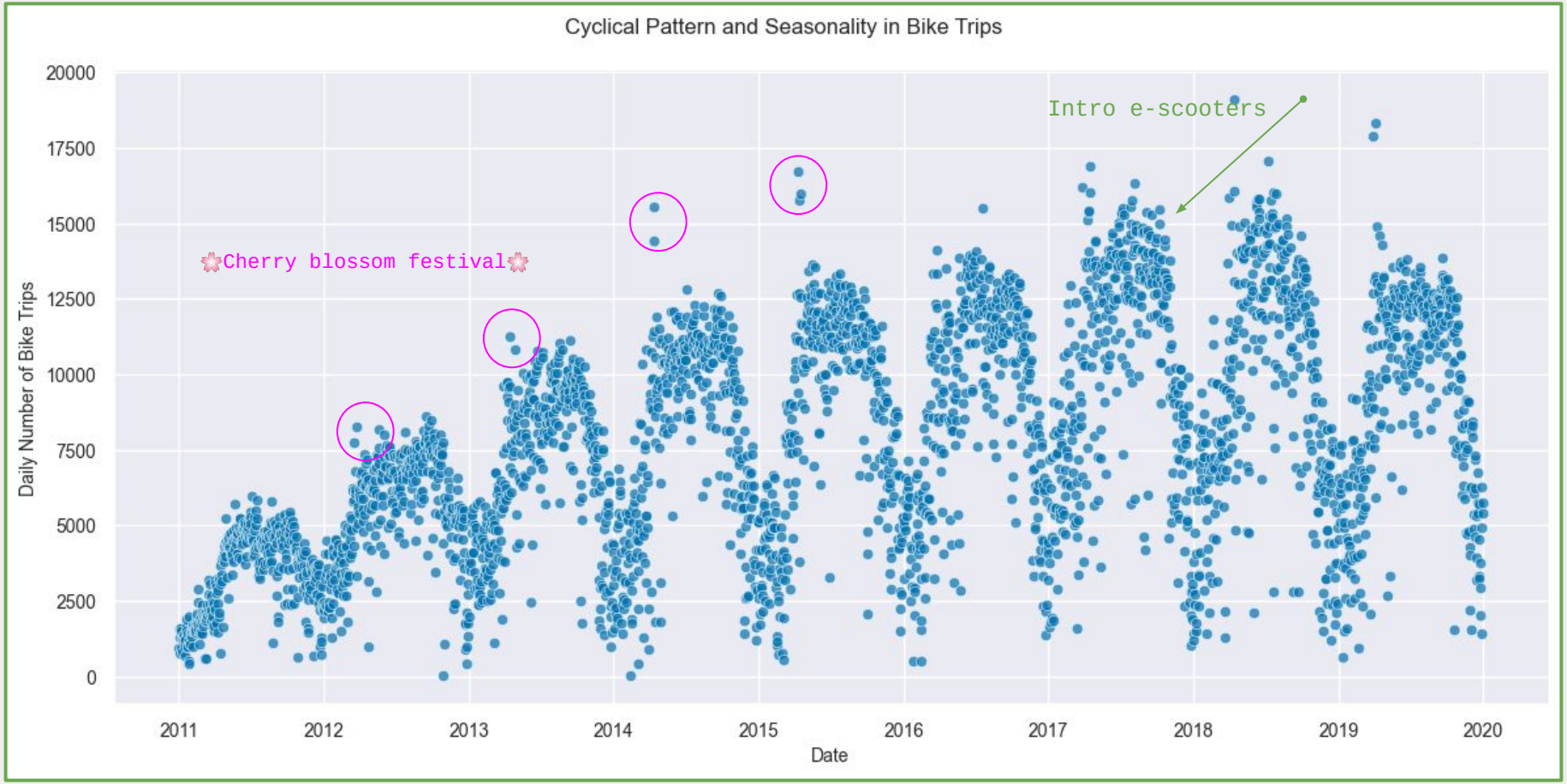
Timespan: 2011-2019

Daily data for minimum and maximum temperature, average wind speed, snow, snow depth and precipitation.

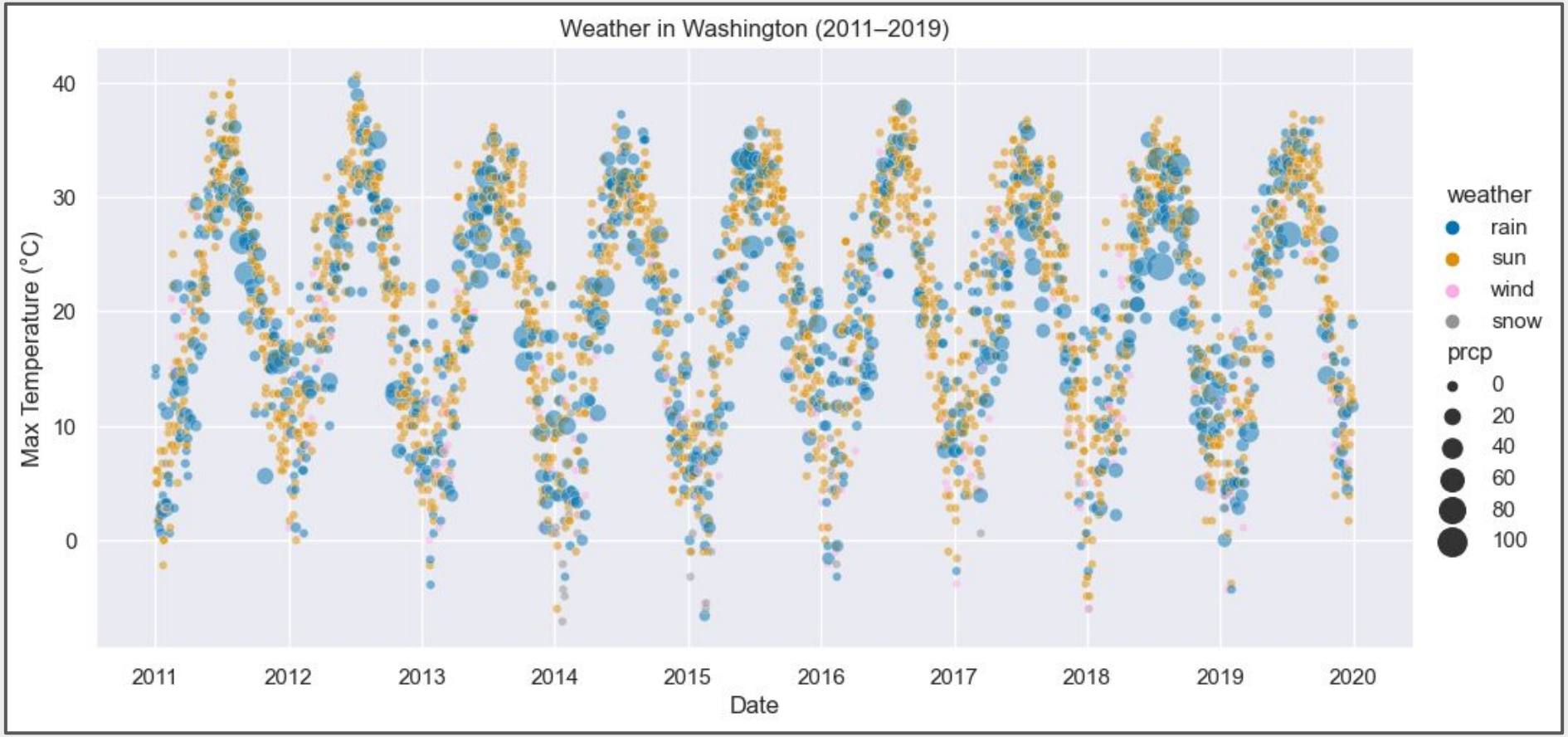
Exploratory Data Analysis



Exploratory Data Analysis

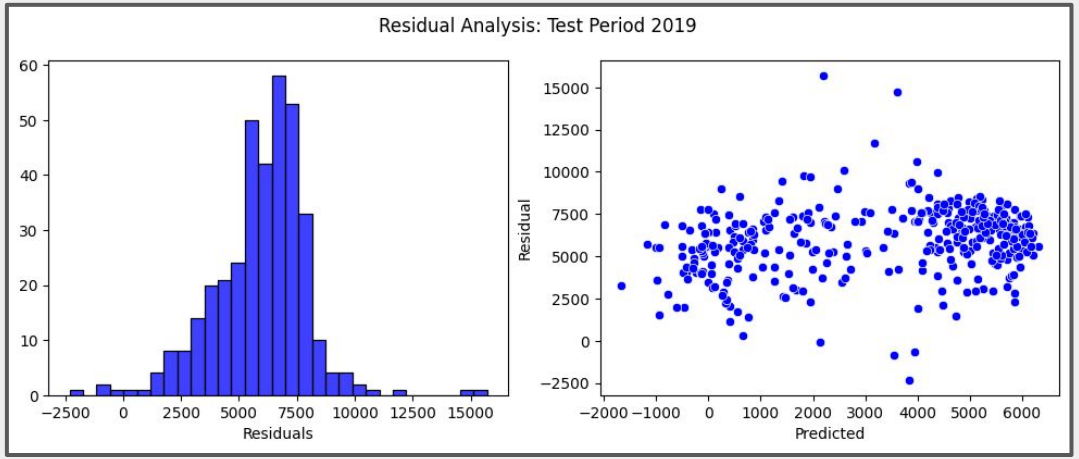


Exploratory Data Analysis



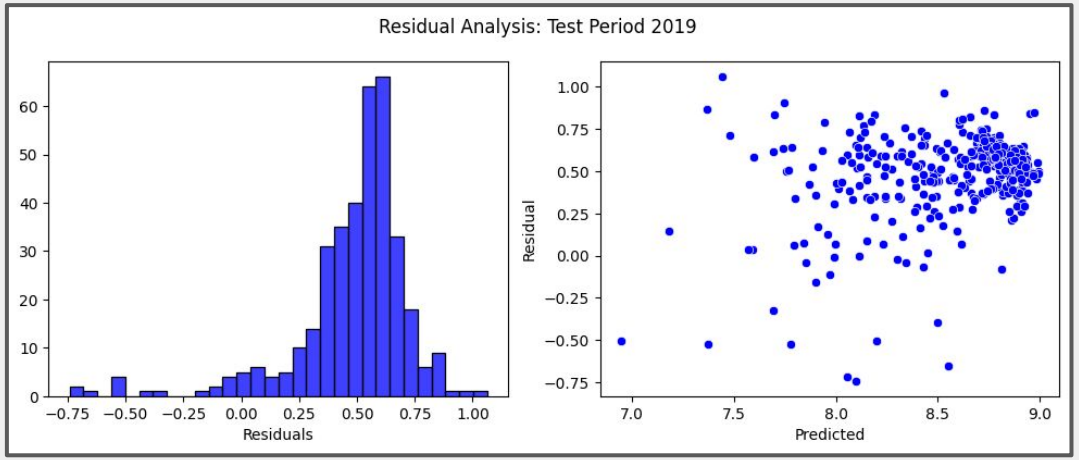
Baseline Model

Training Period	2011-2018
Test Period	2019
Train MAPE	0.474
Test MAPE	0.694



Random Forest

Training Period	2011-2018
Test Period	2019
Train MAPE	0.010
Test MAPE	0.056



Performance

Baseline error rate: **69.3%**

Avg. no of bikes/day (2019): **9311**

Forecasted avg. no of bikes/day (2019): **2858**

Variable revenue per bike: **$2.00** ($1.00 to unlock, average bike trip of 20 mins, $0.05/min)

Forecasted variable revenue loss: **$12,906/day**

Random Forest error rate: **5.6%**

Avg. no of bikes/day (2019): **9311**

Forecasted avg. no of bikes/day (2019): **8790**

Variable revenue per bike: **$2.00** ($1.00 to unlock, average bike trip of 20 mins, $0.05/min)

Forecasted variable revenue loss: **$1,042/day**

Future Work

- Forecasting demand for bike routes, stations.
- Hourly demand predictions
- Incorporate pricing data & competitor analysis

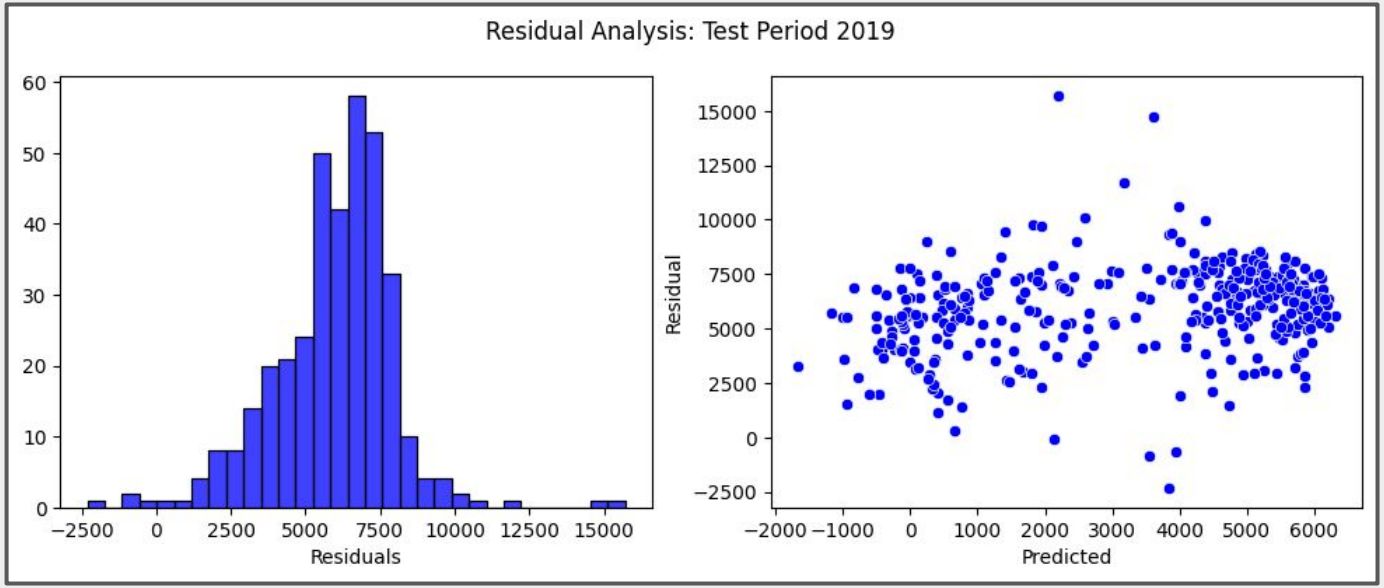
Thank You!

Do you have any questions?

Predictions

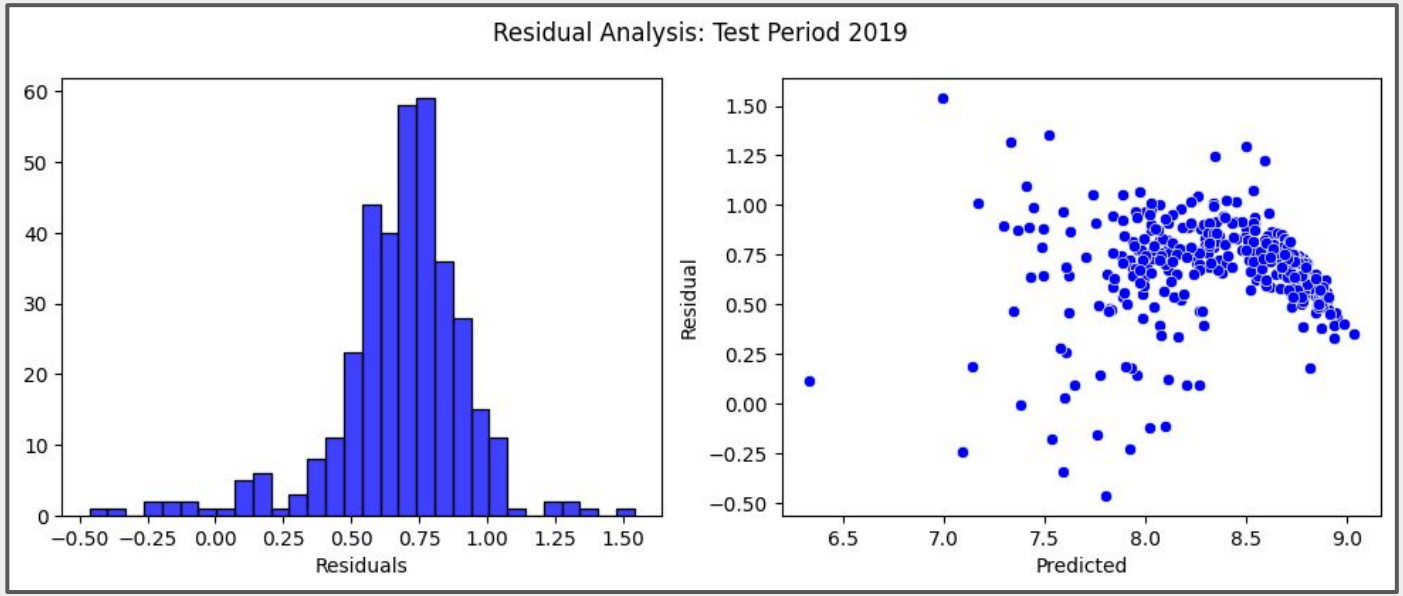
Training Period	2011-2018
Test Period	2019
Train MAPE	0.474
Test MAPE	0.693
ADF Test Residual Stationarity	-2.587786
p-value	0.095548
Critical Values	
1%	-3.449
5%	-2.870
10%	-2.571

Baseline Model - Linear Regression



Poisson Model

Training Period	2011-2018
Test Period	2019
Train MAPE	0.025
Test MAPE	0.076
ADF Test Residual Stationarity	-8.460738
p-value	0.000000
Critical Values	
1%	-3.449
5%	-2.870
10%	-2.571



Random Forest Regression

Training Period	2011-2018
Test Period	2019
Train MAPE	0.01
Test MAPE	0.056