



<https://www.theatlantic.com/photo/2018/03/bike-share-oversupply-in-china-huge-piles-of-abandoned-and-broken-bicycles/556268/>

# Time-series forecasting

Trip count prediction from bike sharing data



# Business Goal

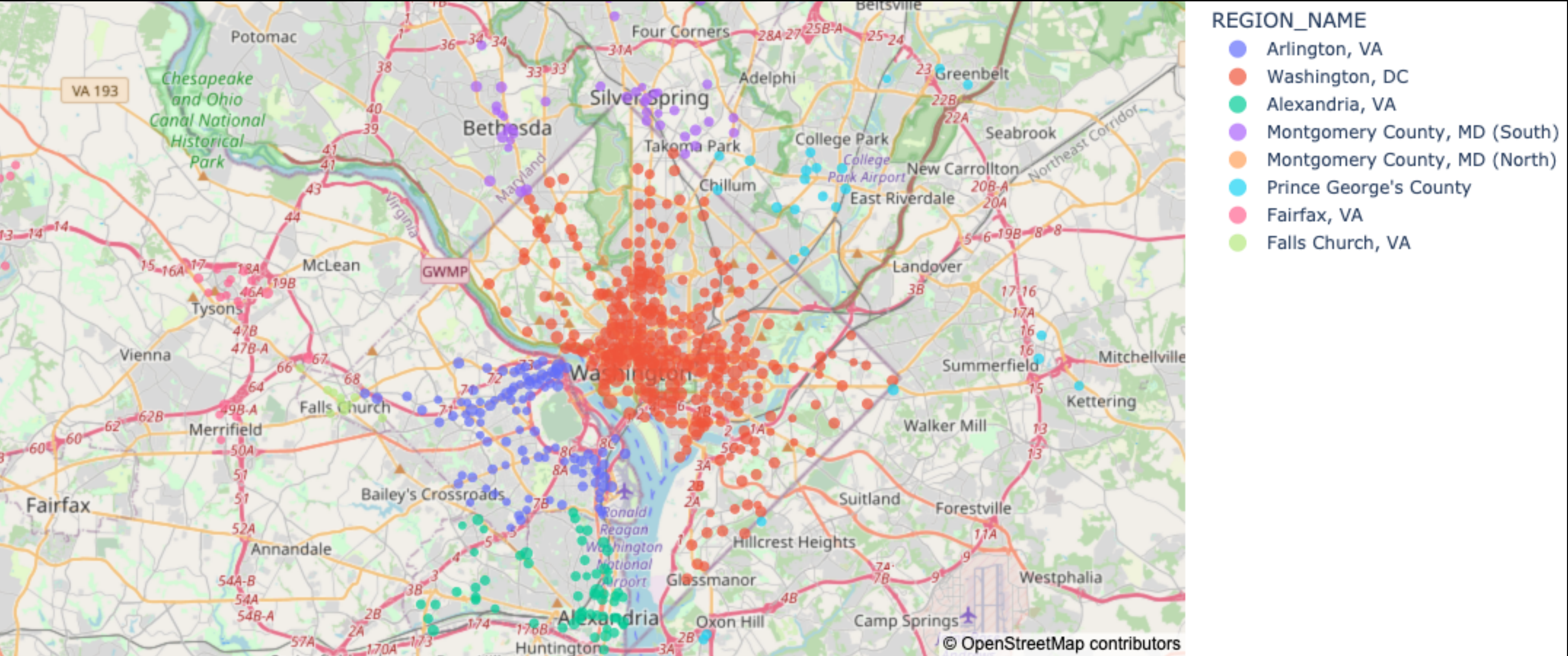
## Predict hourly trip count for a specific date

- Data :
  - Source: Public S3 bucket of Capital Bike Share
  - Duration: Sep 2010 - Feb 2022
  - ~ 300 M data points for each event
  - 8 regions of Washington DC and nearby
  - 100K data points from 335 Stations of Washington DC



# Business Goal

## Distribution of 1,200 stations





# Modelling

- NeuralProphet, heavily inspired by Facebook Prophet and backed by PyTorch
- Contains components, trend, seasonality, auto-regression, special events, future regressors and lagged regressors
- Scalable to add any additional components
- Applicable to global modelling

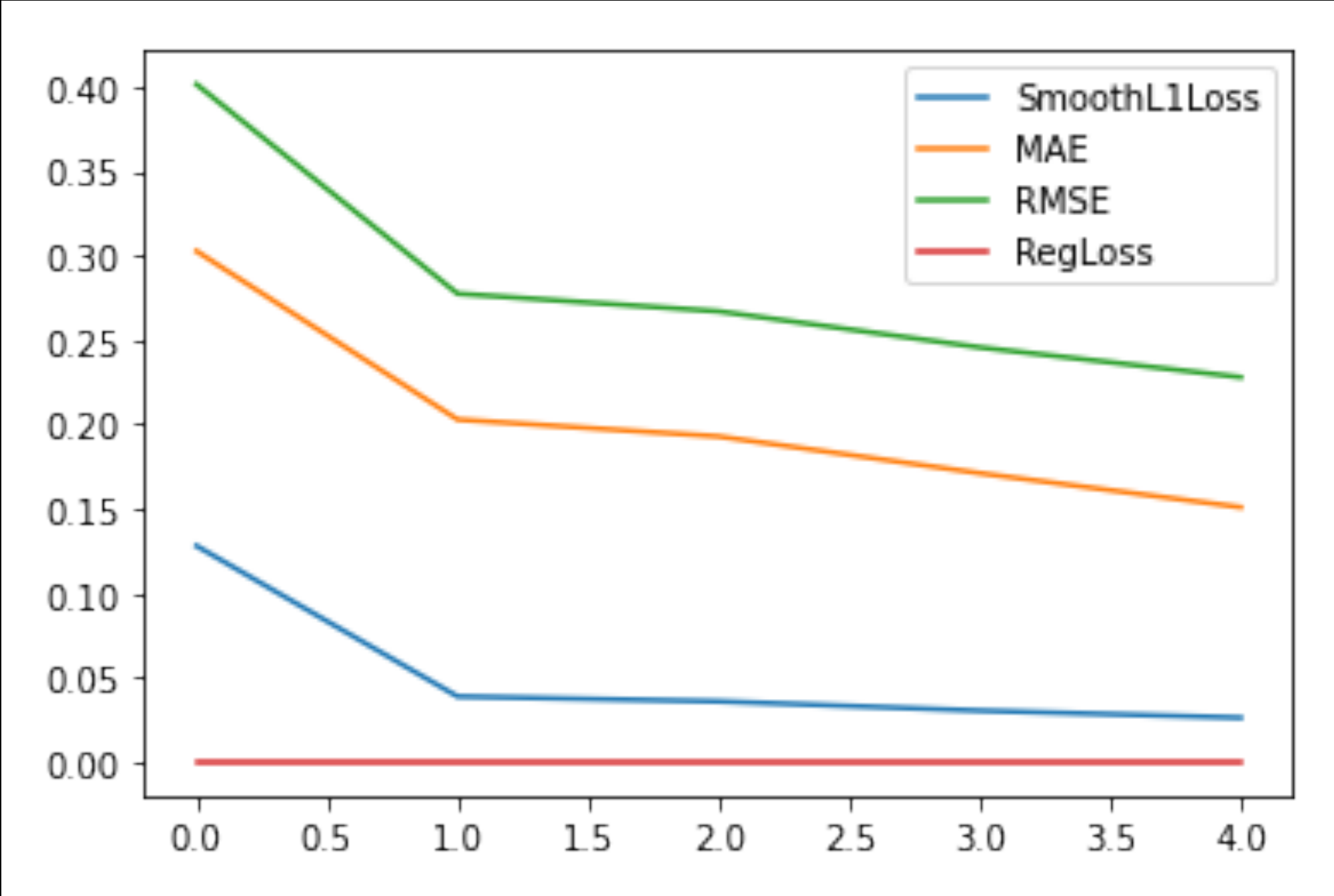
# Global Modelling:

```
m = NeuralProphet(n_lags = 24, epochs=5)
df_train_dict, df_test_dict = m.split_df(
    df_dict, valid_p = 0.10,
    local_split = True)

metrics = m.fit(df_train_dict, freq = 'H')

future = m.make_future_dataframe(df_test_dict,
    n_historic_predictions = True)

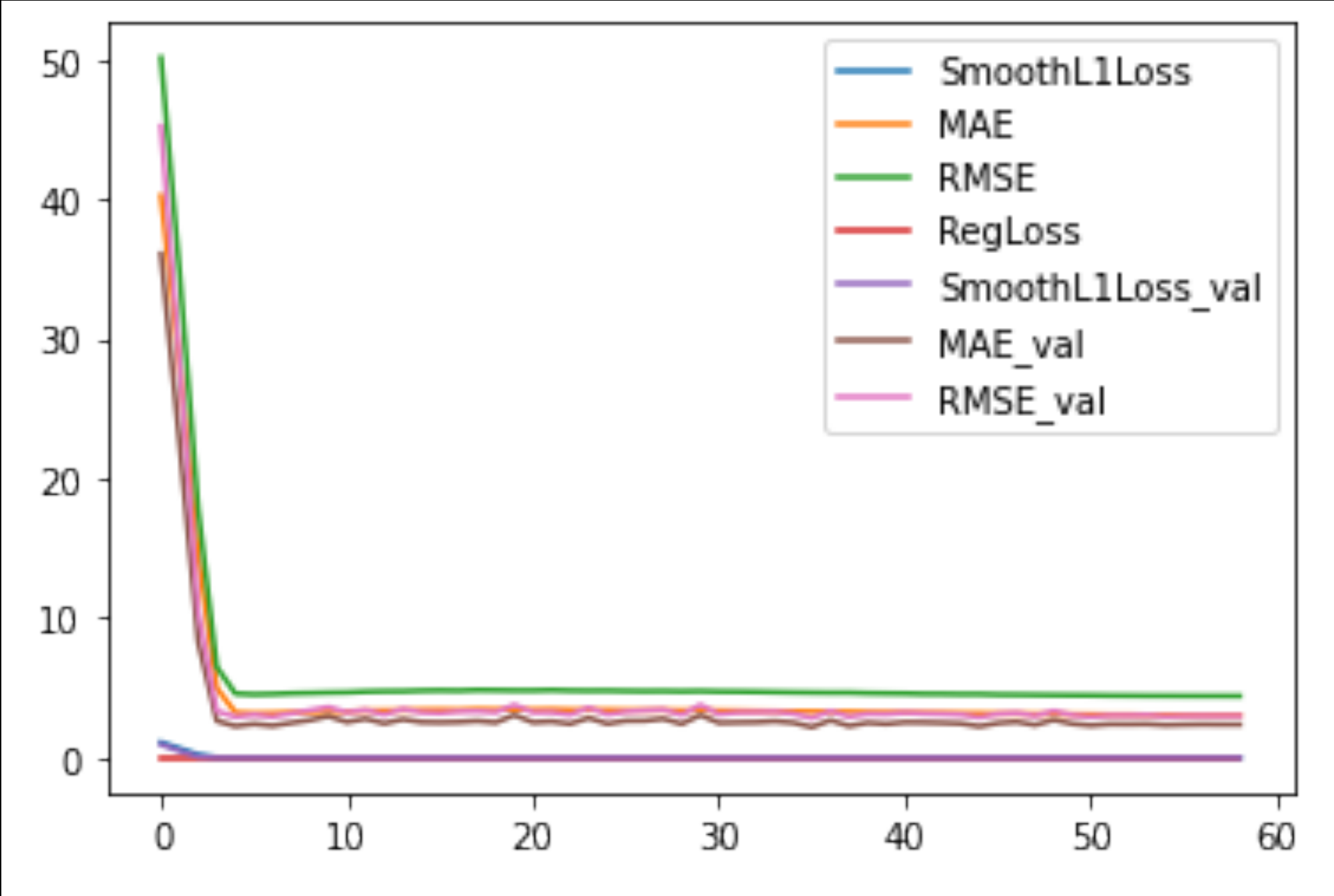
forecast = m.predict(future)
```



	SmoothL1Loss	MAE	RMSE	RegLoss
Train	0.026165	0.150889	0.227946	0.0

# Local Modelling:

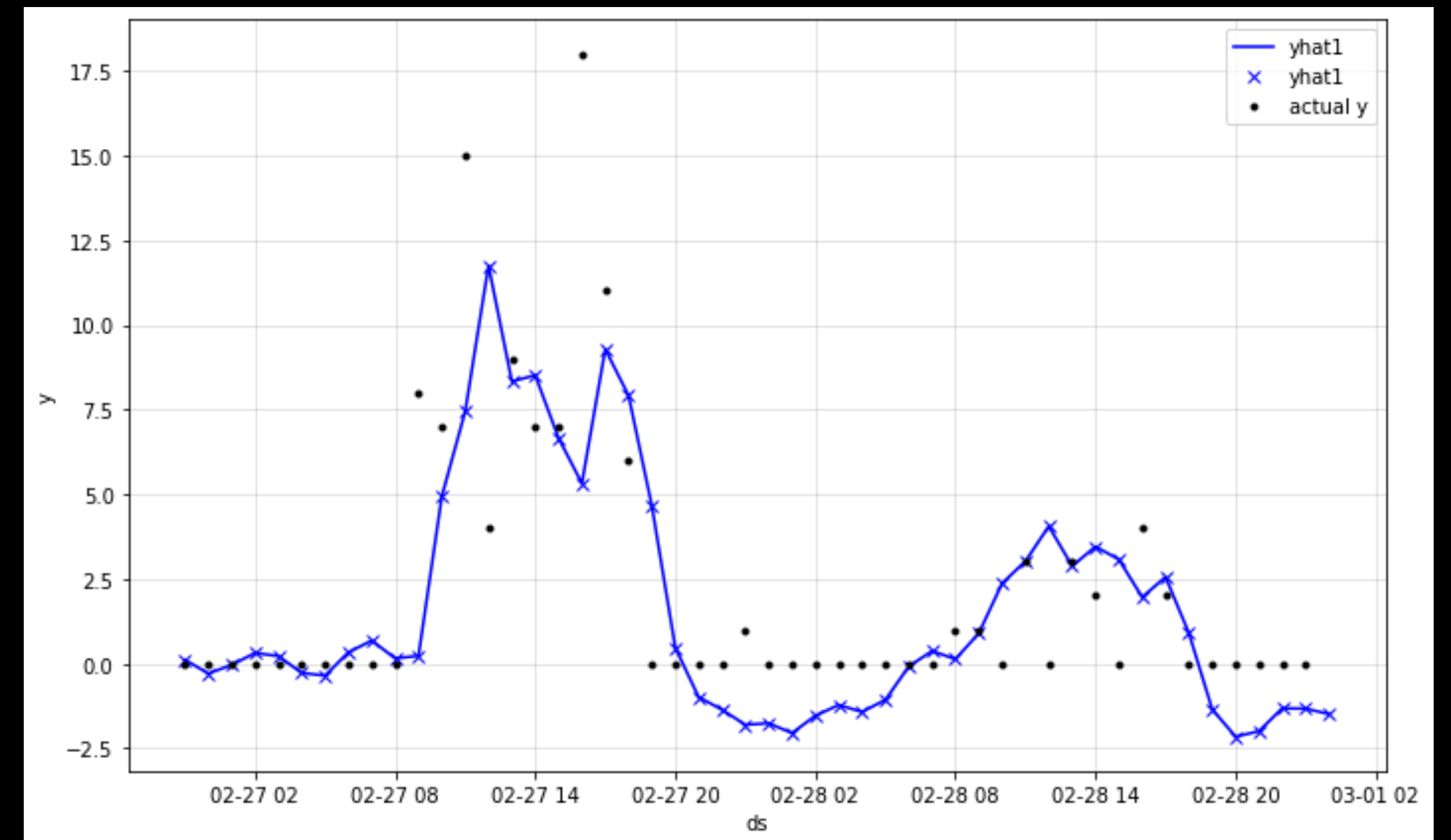
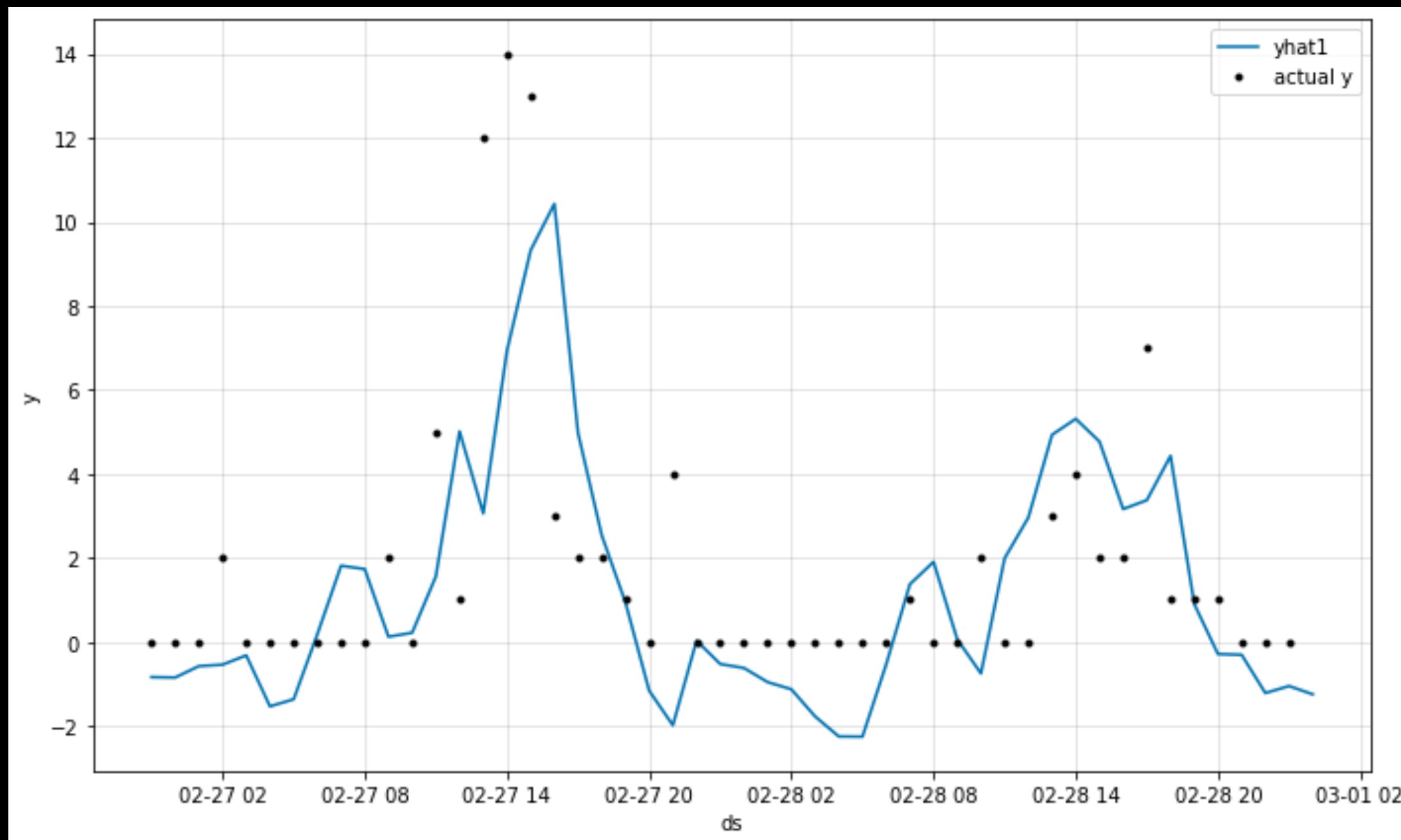
```
m = NeuralProphet(  
    growth='linear',  
    yearly_seasonality=True,  
    weekly_seasonality=True,  
    daily_seasonality=True,  
    n_lags=3*24,  
    learning_rate=0.01,  
)  
  
m.add_country_holidays('US')  
  
df_train, df_test = m.split_df(df, freq='H',  
                               valid_p = 0.10)  
  
metrics = m.fit(df_train, freq='H',  
                validation_df=df_test, progress='plot')
```



	SmoothL1Loss	MAE	RMSE	RegLoss
Train	0.014861	3.069174	4.427516	0.0
Validation	0.006771	2.387943	3.012018	0.0

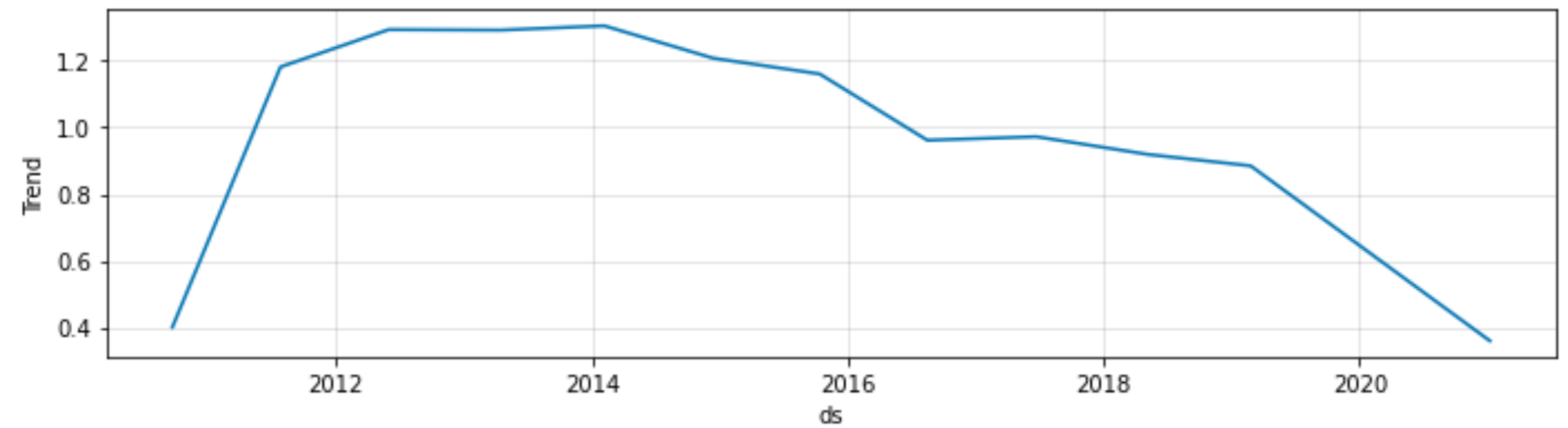
# Test comparison:

## Global vs Local

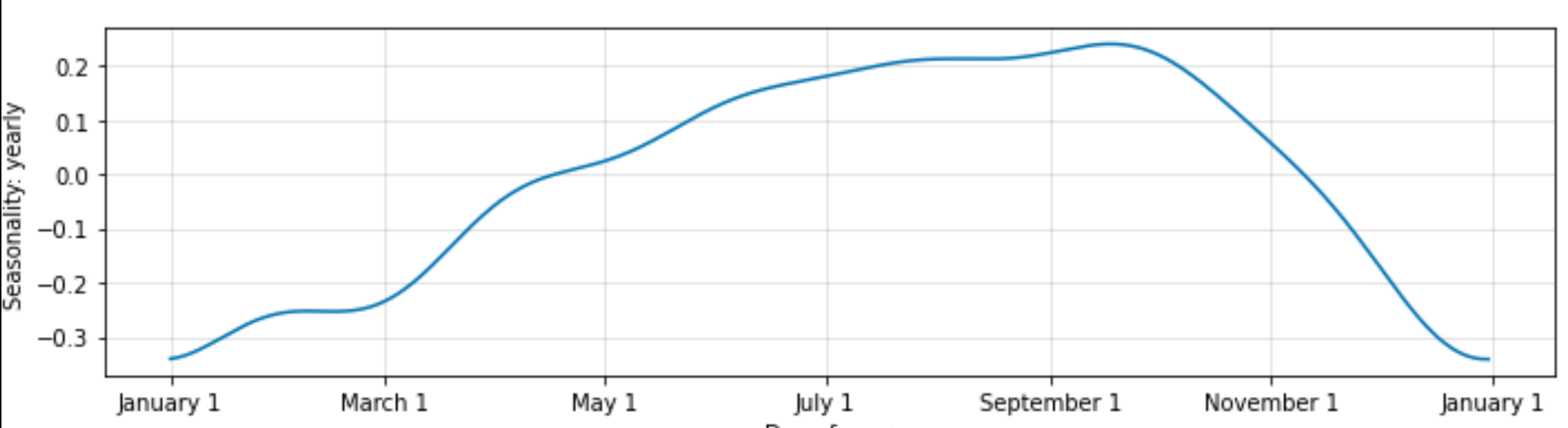


# Components:

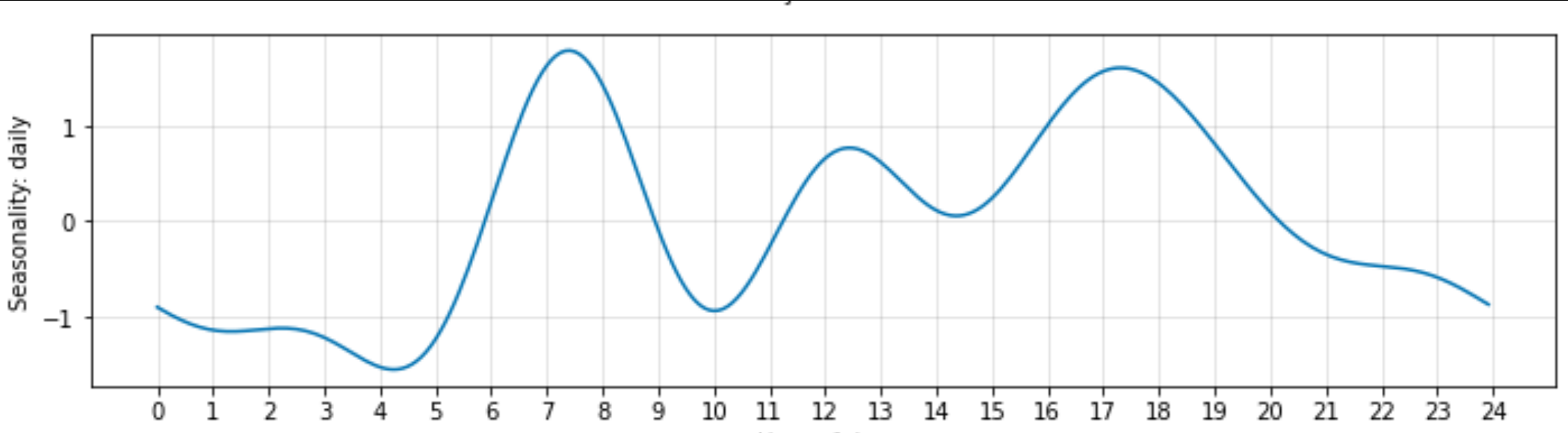
Trend



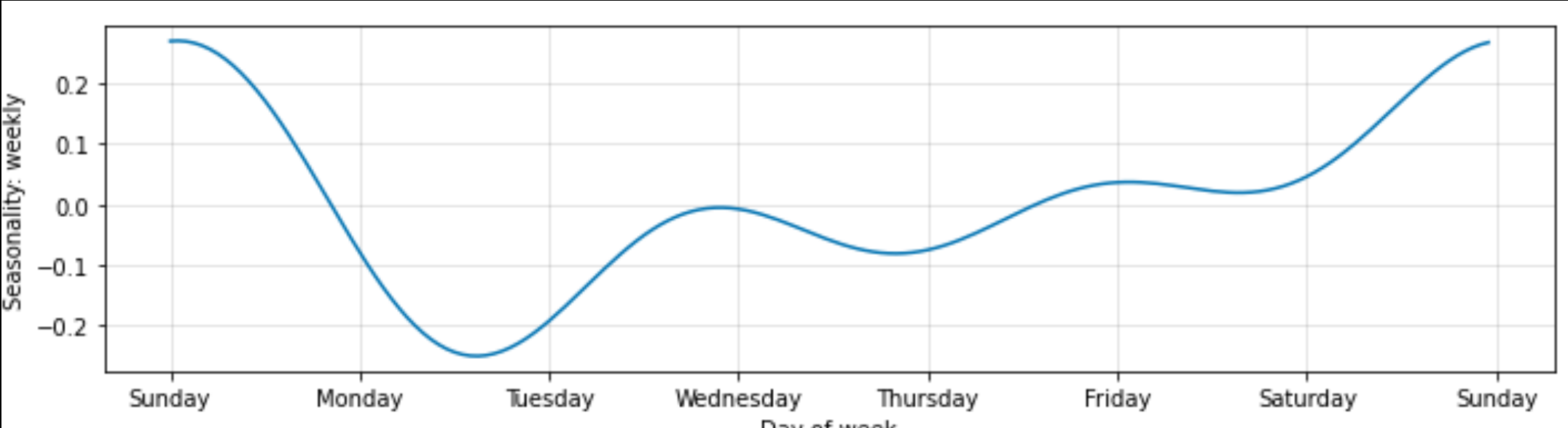
Monthly seasonality



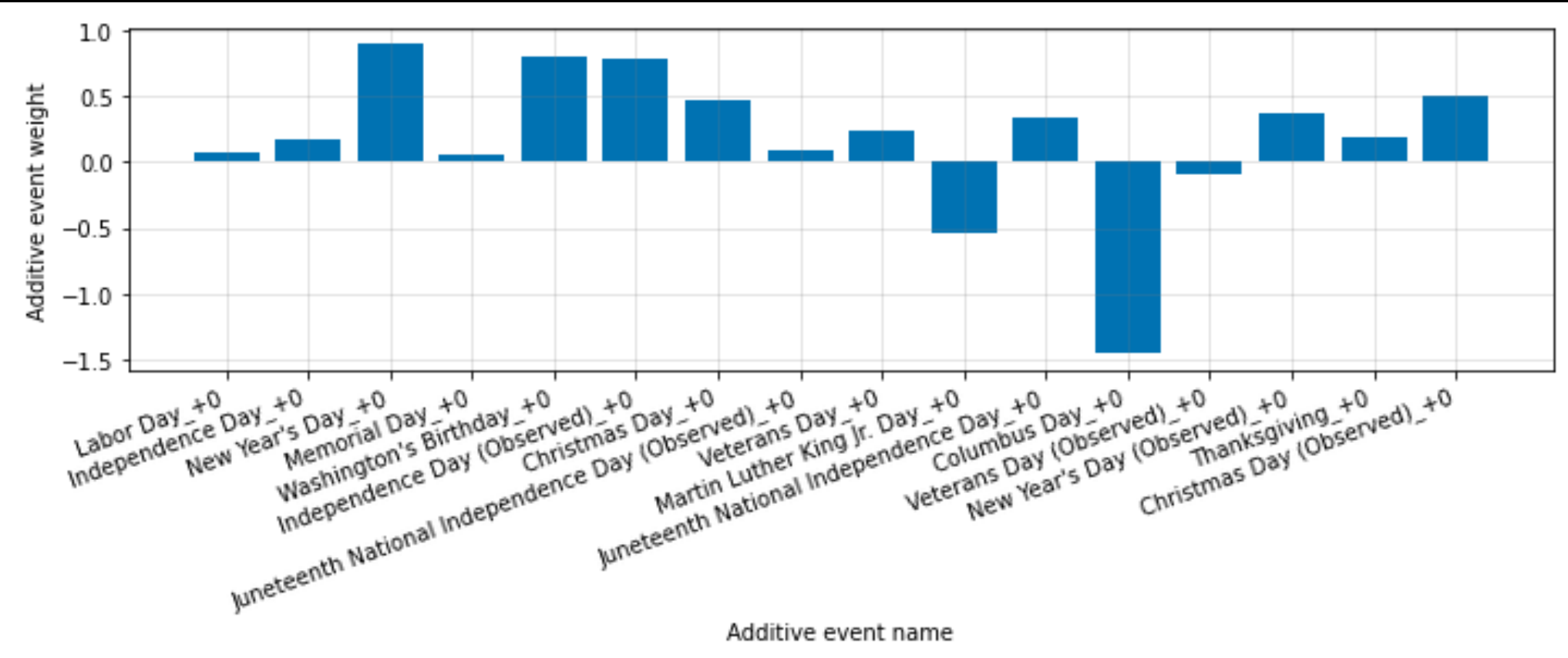
Daily seasonality



Weekly seasonality



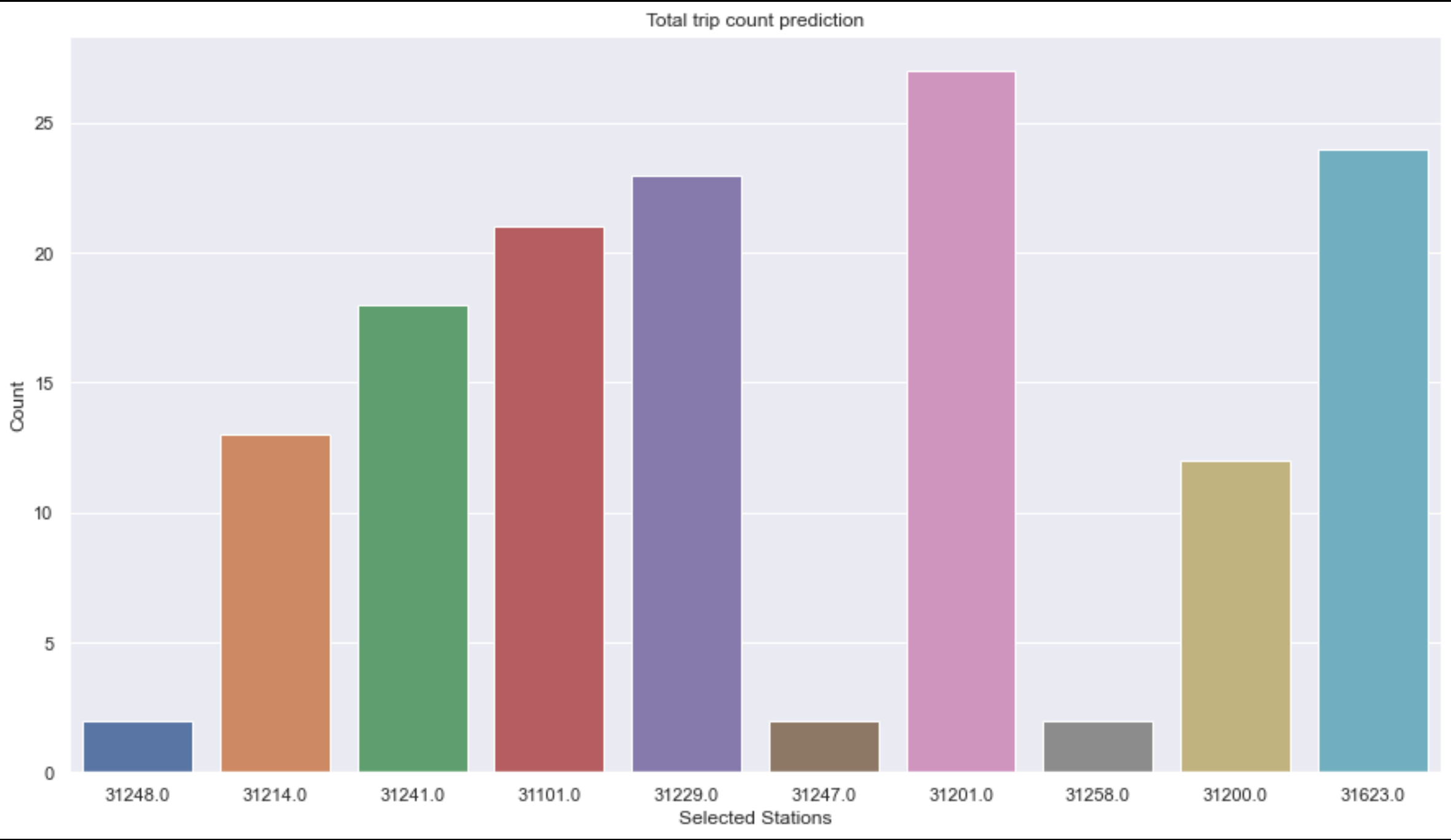
Weight from Holidays



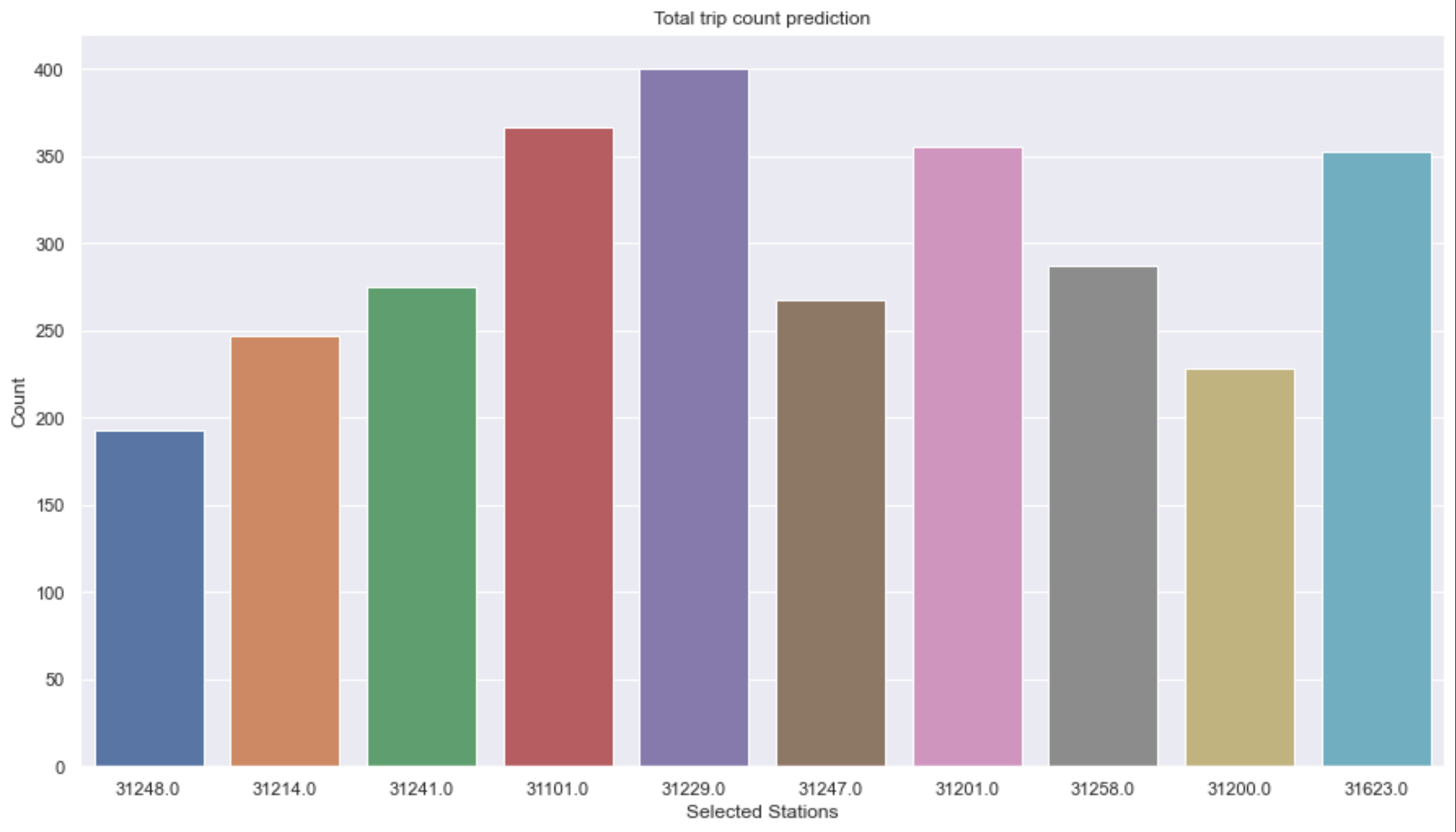


# Prediction:

Actual: Feb 4, 2022



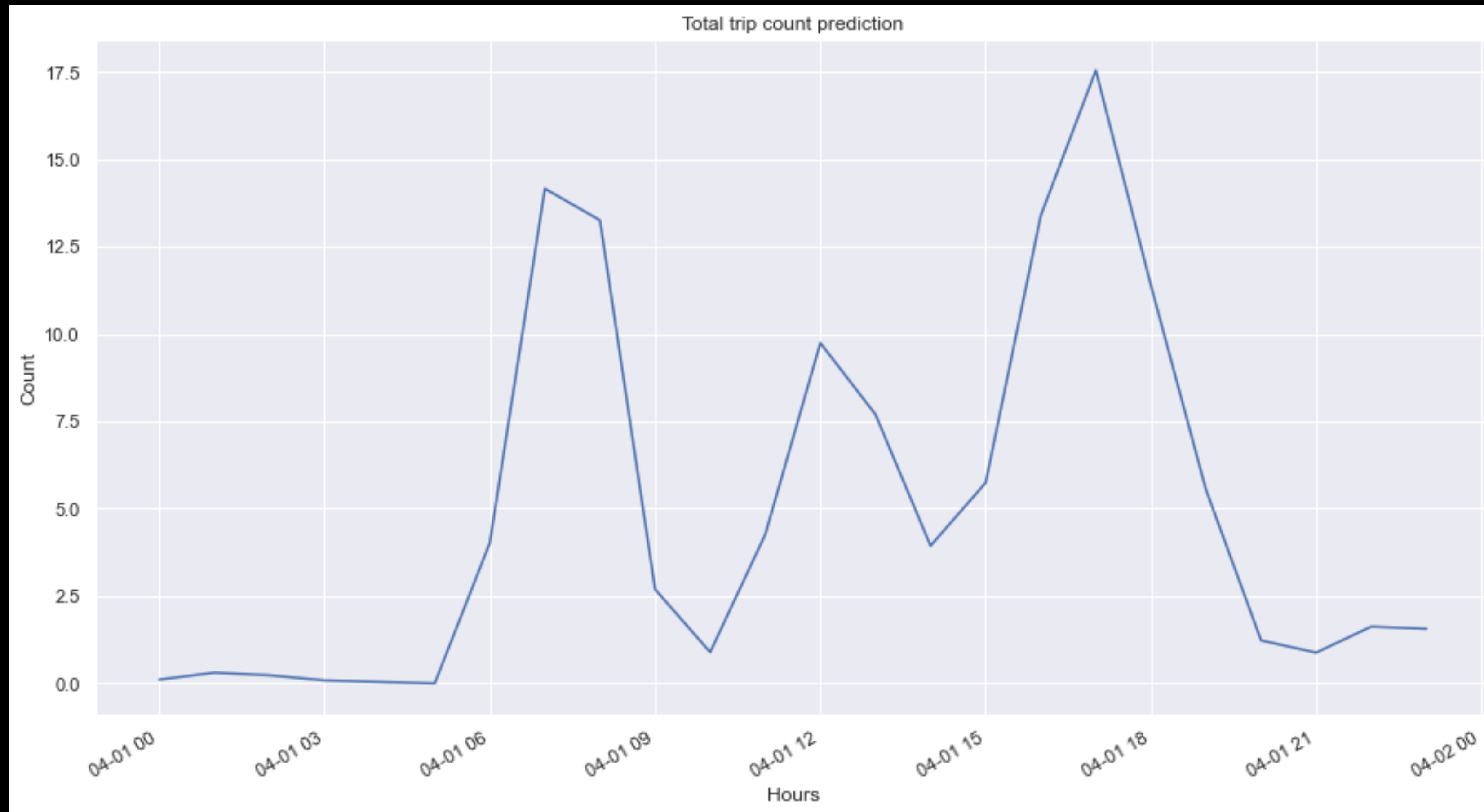
Prediction: Apr 1, 2022 (today)





# Prediction:

Prediction: Apr 1, 2022 (today)





# Way forward:

- Feature improvement of NeuralProphet (such: early stopping)
- Global model for regions could be applied with support of GPU
- An application will be supportive in decision making

## Thanks you all

The logo for SPICED, featuring the word "SPICED" in a bold, purple, sans-serif font.The logo for Neural Prophet, with the words "Neural Prophet" in a blue, sans-serif font, where "Neural" is enclosed in a blue rectangular box.The logo for pandas, featuring a stylized icon of vertical bars in blue, orange, and yellow, followed by the word "pandas" in a blue, sans-serif font.The logo for NumPy, featuring a stylized icon of a cube with blue and orange faces, followed by the text "NumPy" in a blue, sans-serif font.The logo for pydeck, featuring a stylized icon of three overlapping blue and green squares, followed by the text "pydeck" in a blue, sans-serif font.The logo for PyData, featuring a stylized icon of a cube with blue and orange faces, followed by the text "PyData" in a blue, sans-serif font.The logo for plotly, featuring a stylized icon of a bar chart with blue bars, followed by the text "plotly" in a blue, sans-serif font.