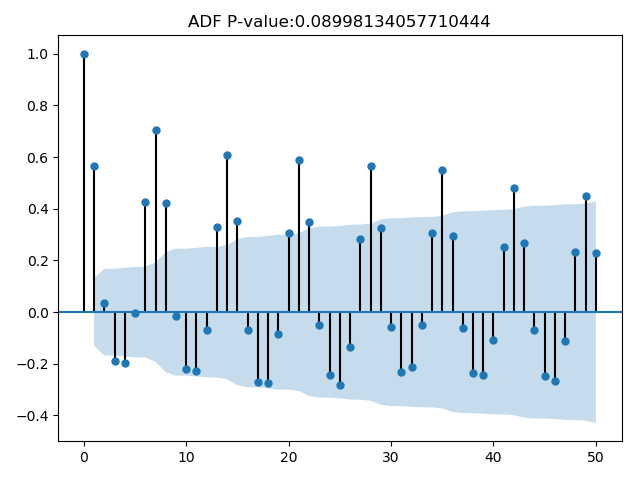
# Time series analysis sample project

## Task formulation

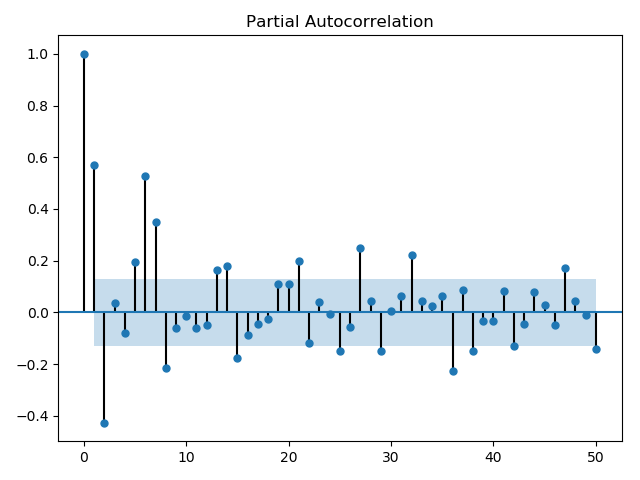
Predict metric (number of drives) for one fleet (BER) for 1 week ahead.

## Predictions via time series

Initial data analysis:

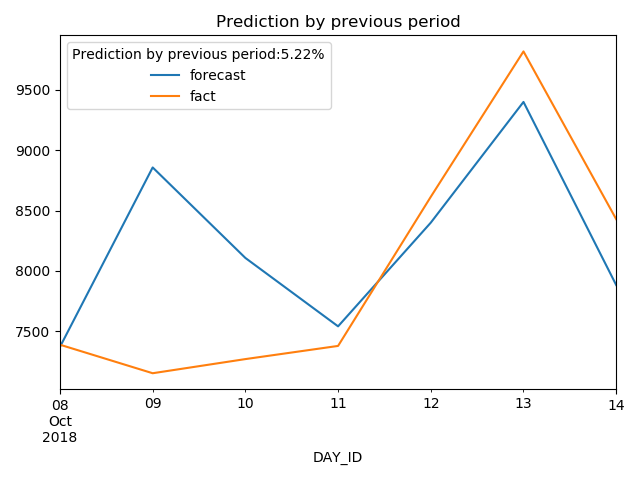


ADF P-value >0.05 which means we can't reject hypothesis that the process is a random walk, although near to it.

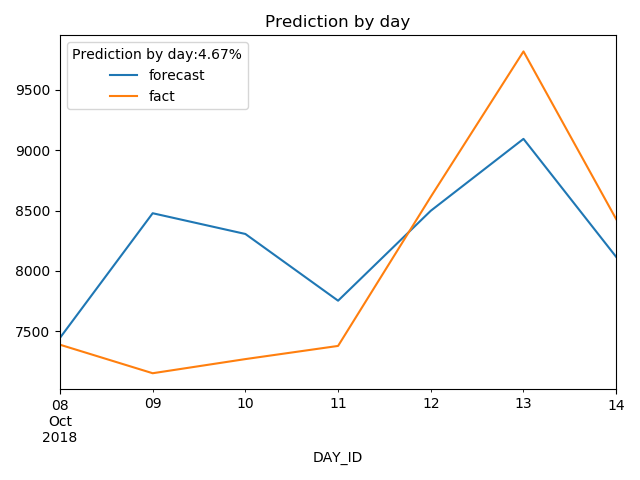


Prediction error formula used below is median absolute error divided by mean fact value.

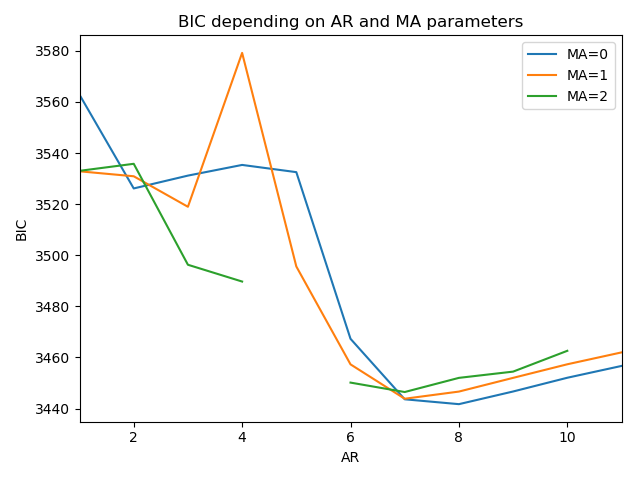
Simple prediction by using previous week data (predicted value for Monday is previous Monday, for Tuesday is previous Tuesday, etc):



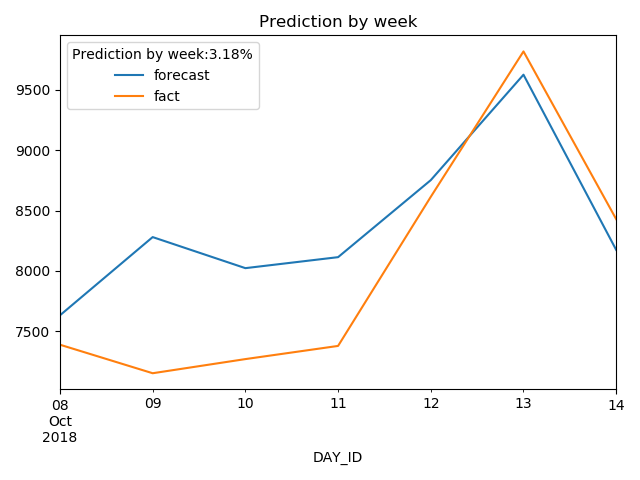
Prediction from daily data with (AR, MA) = (7, 2):



Tuning daily parameters result:

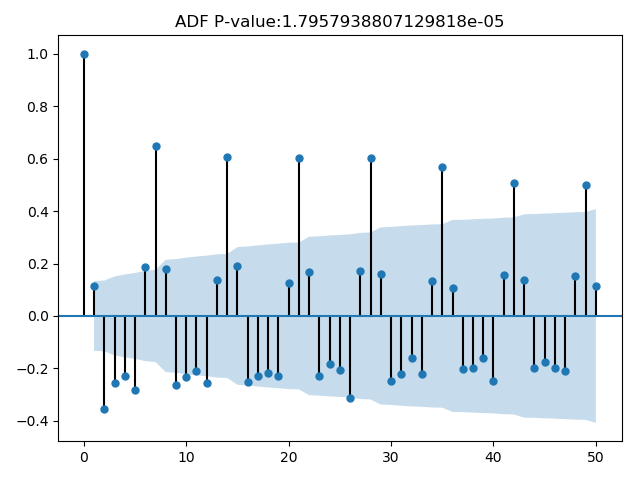


Prediction from weekly data with (AR, MA) = (2, 0) (data is resampled by week and each week day is modeled separately):



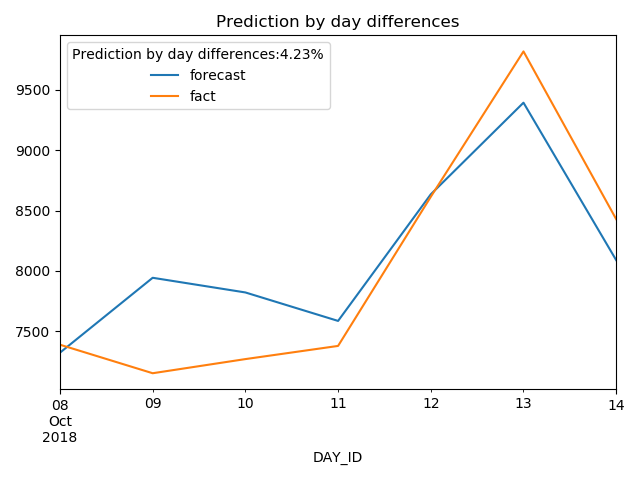
## Analysis of daily data via differences

Transformation: difference between current and previous value



ADF P-value <0.05 which means we can reject hypothesis that the process is a random walk

Prediction with (AR, MA) = (14, 1)



Tuning parameters results

