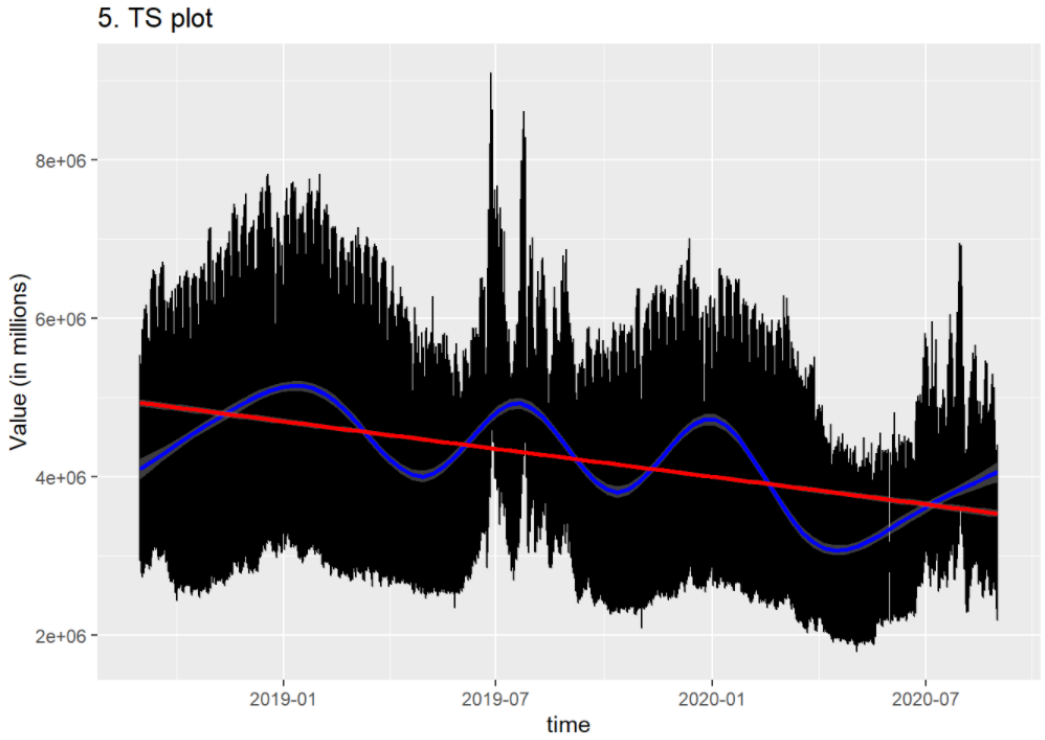
TIME SERIES ANALYSIS PROJECT

Sofia Davoli

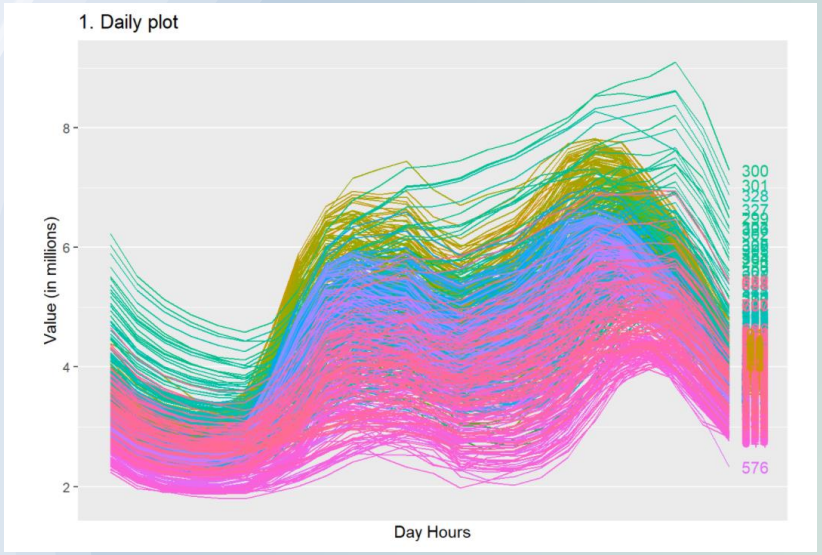
813479

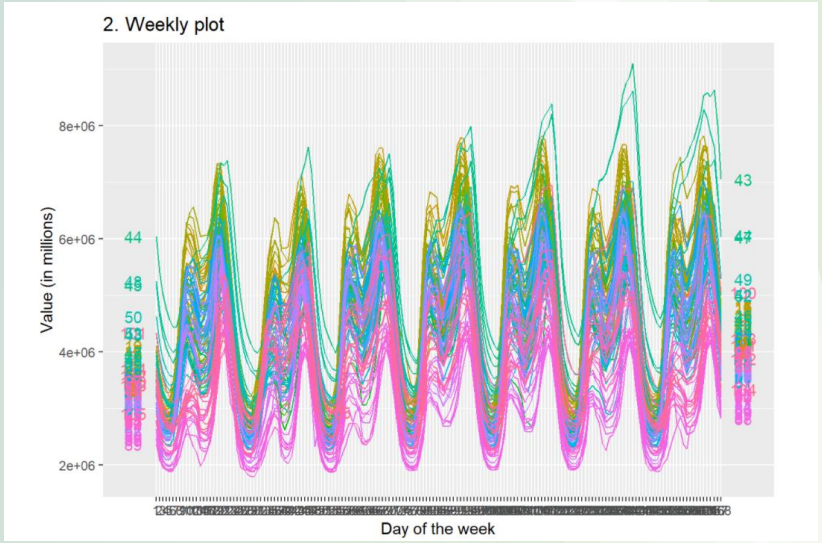
DATA PREPROCESSING AND EXPLORATION

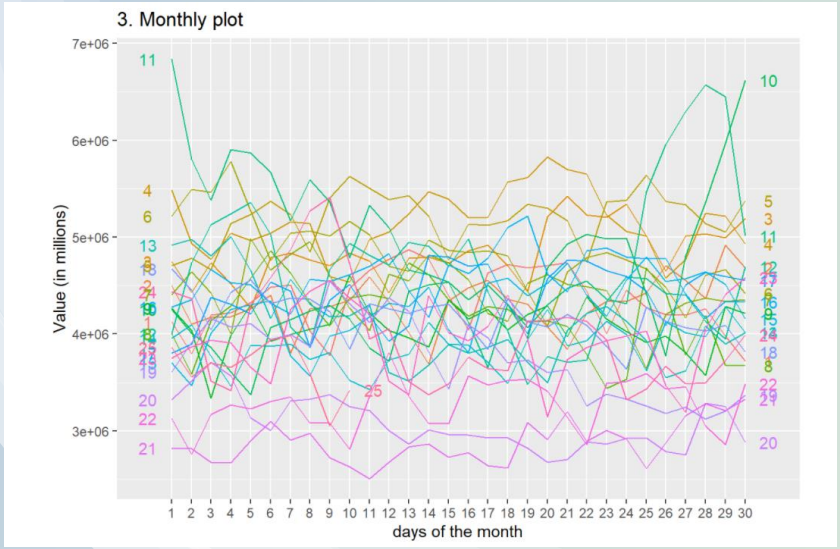


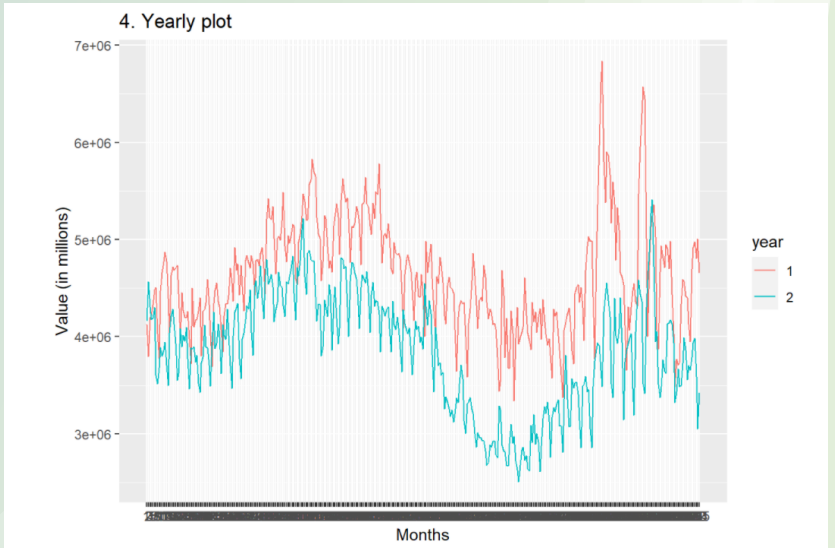
- Scaling (to optimize model training)
- Missing (solved using substitution with previous hour value)

SEASONALITY

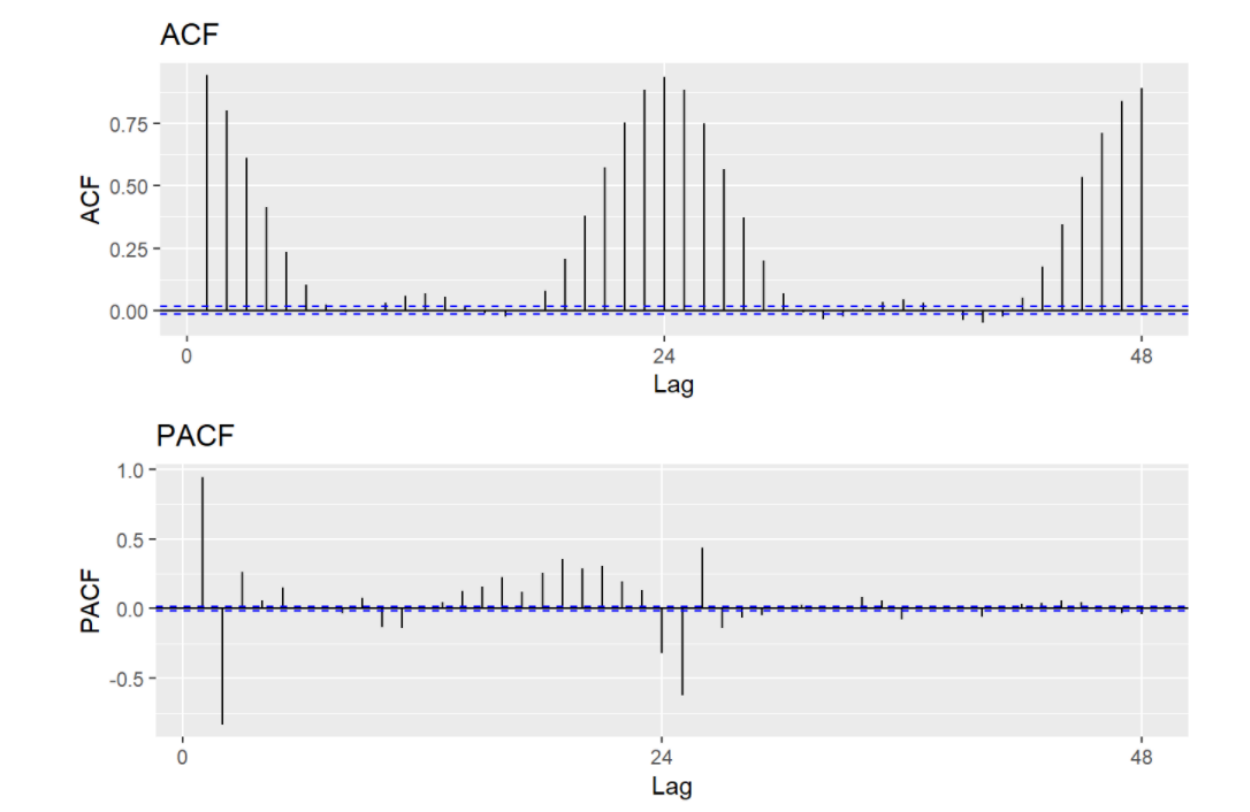








ARIMA AND NAIVE METHODS



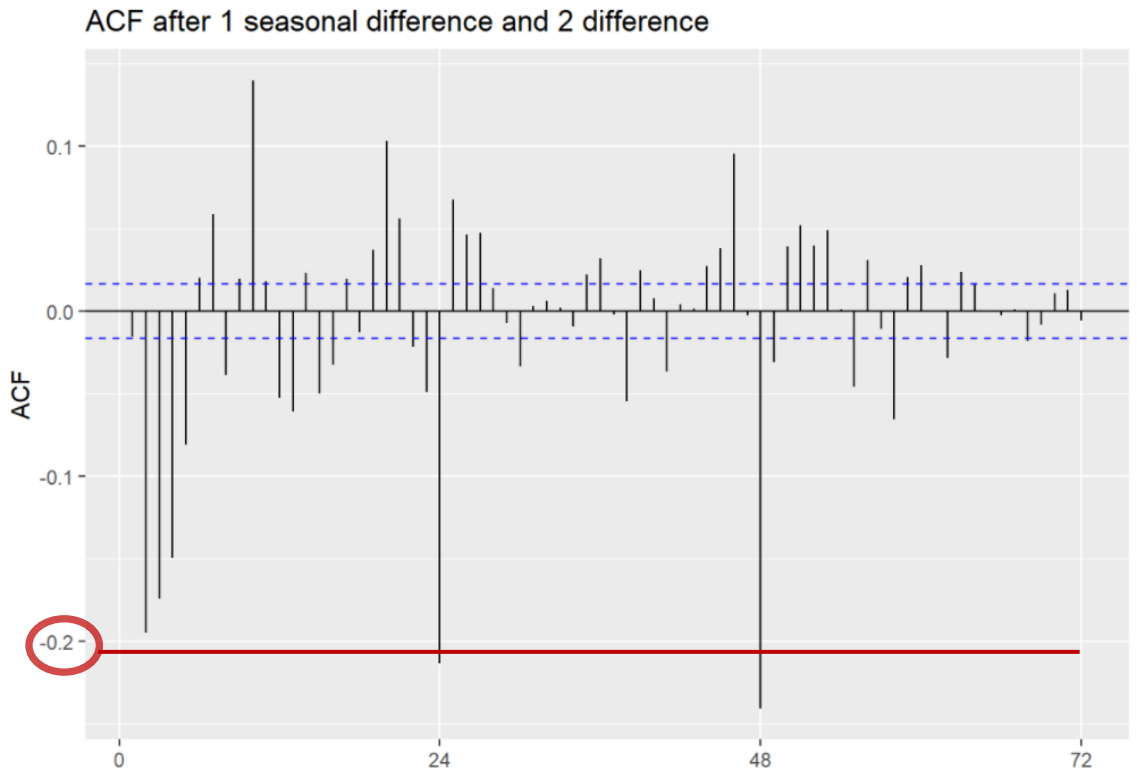
Training (80%) – test (20%)

-> seasonality 24 (daily)

-> 5 non-zero partial autocorrelations
-> AR(2),AR(3) or AR(5)

BOX-JENKINS METHODOLOGY

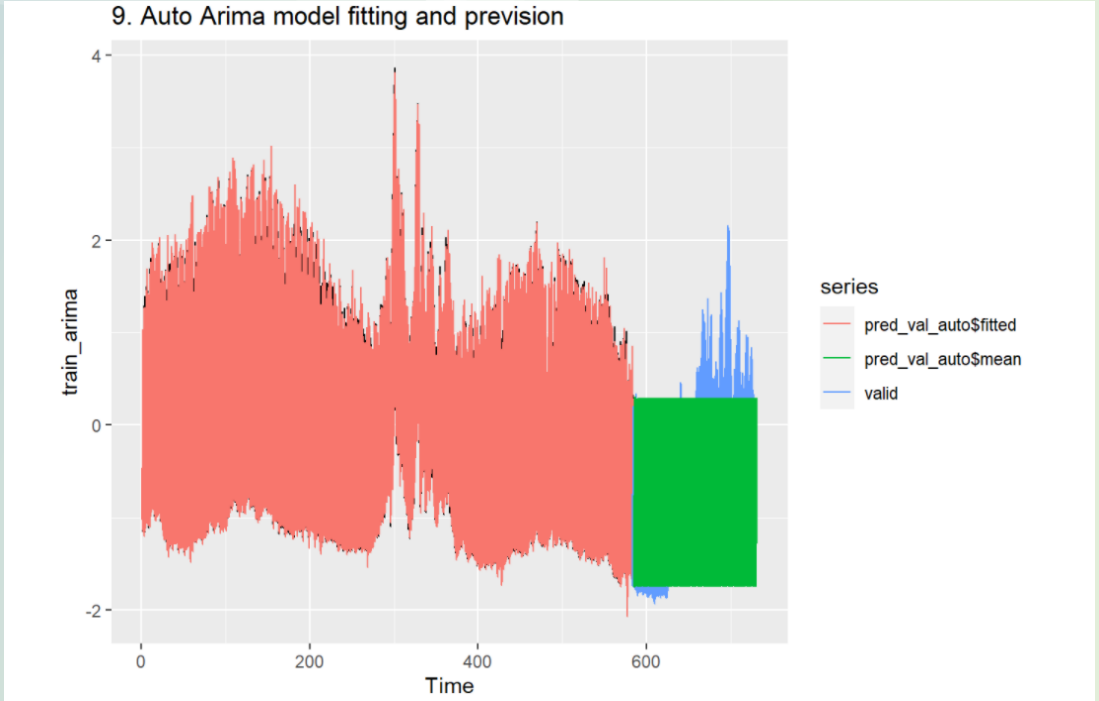
- lambda: 0.9284836
- seasonal differences at lag 24
- Box-Ljung test reject H0 of White Noise
- 2 more differences has to be taken to get to a WN process
- SARIMA model with daily seasonality, 2 integration and AR component equal to 2, 3 or 5

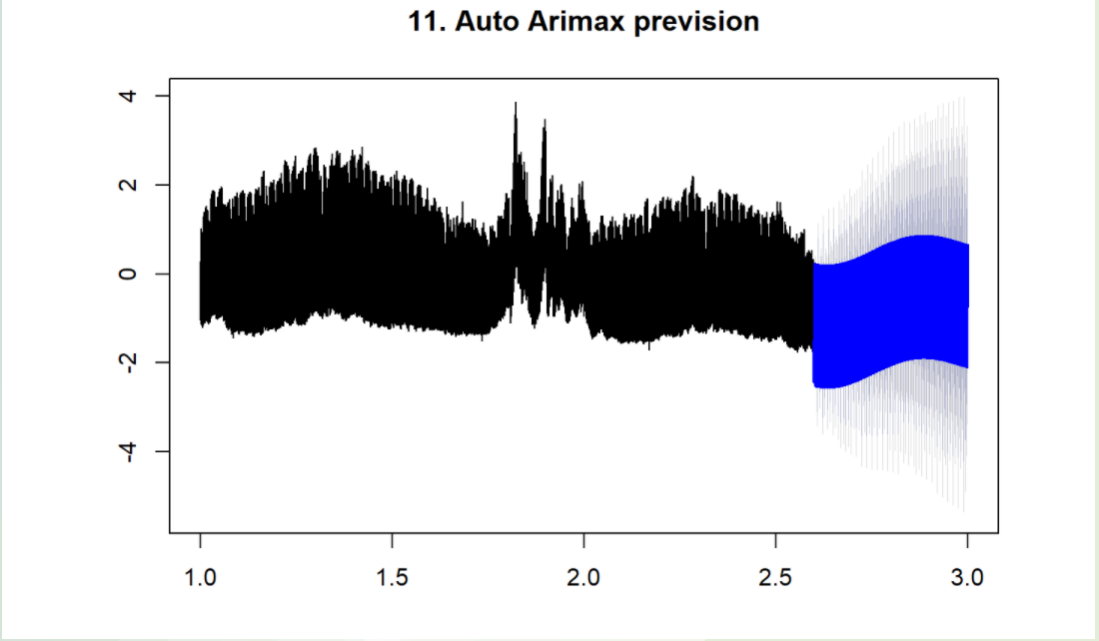



```
##  Box-Ljung test
##
## data:  train_diff_trend2
## X-squared = 3.4234, df = 1, p-value = 0.06428
```

ARIMA AND NAIVE METHODS

MODEL	MAE VALIDATION
naive	0.6563
ARIMA(2,1,1) with drift	0.5436
Arima (3,1,2)) with drift	0.5206016
Auto Arima (5,0,2)(2,1,0) [24]	0.496358
ARIMAX(5,1,2) fourier	0.4367472

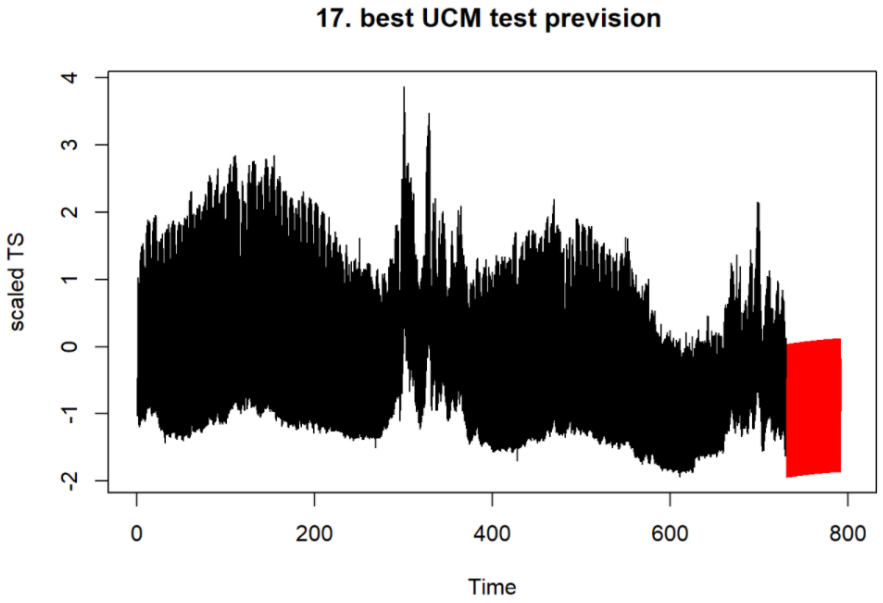


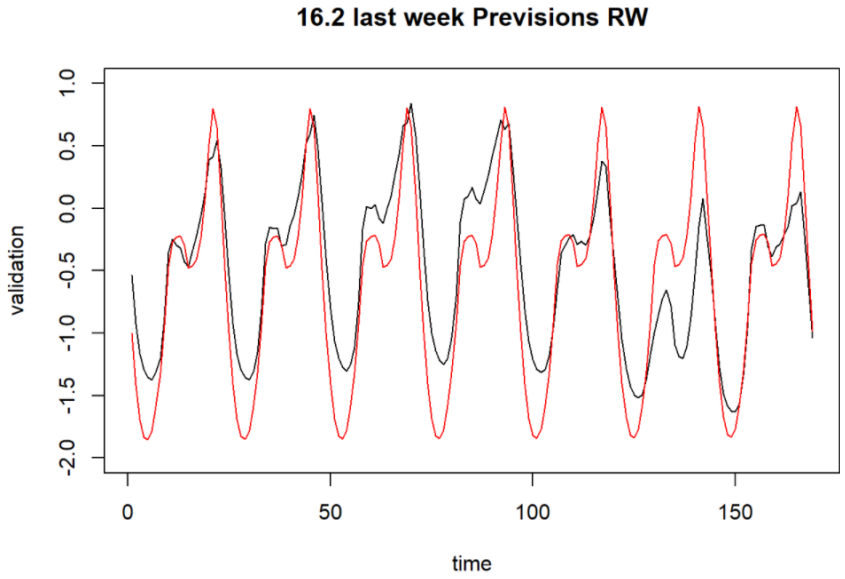


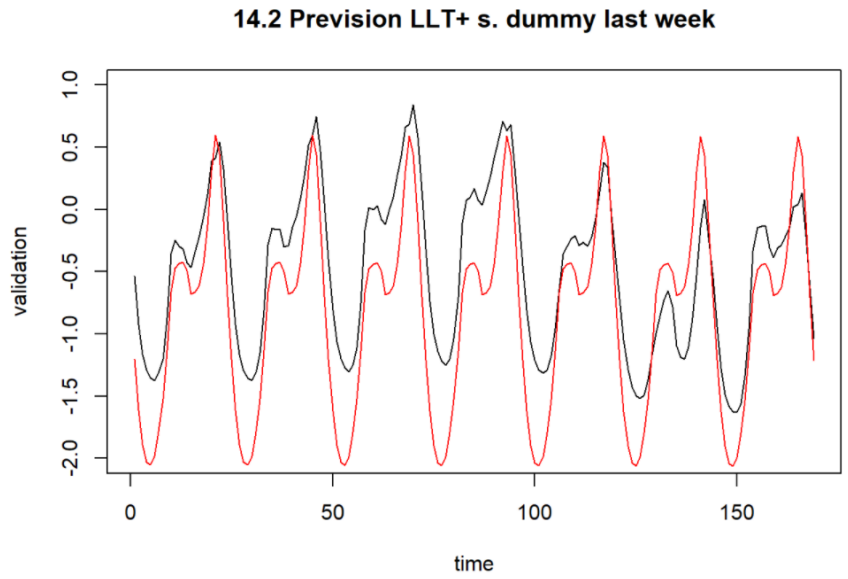
UCM

Training (80%) – test (20%)

MODEL	MAE VALIDATION
Local Linear Trend	0.6857667
LLT with daily stochastic dummy.	0.511847
LLT+seasonal dummy model+ cycle	0.5731635
trigonometric seasonality	0.6601218
Random Walk	0.4814821

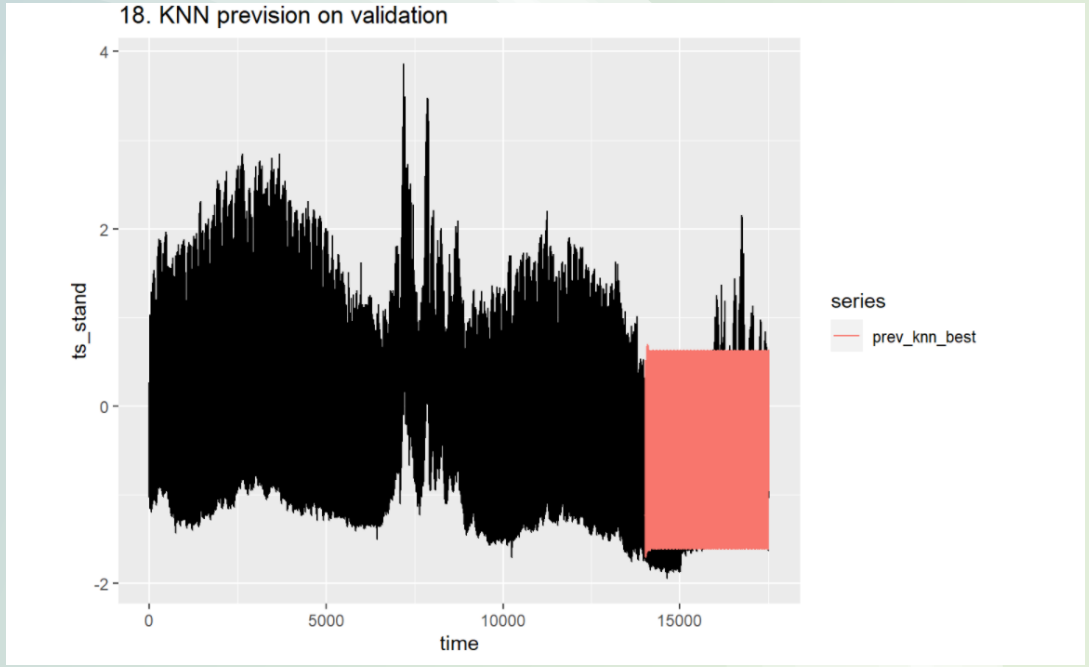


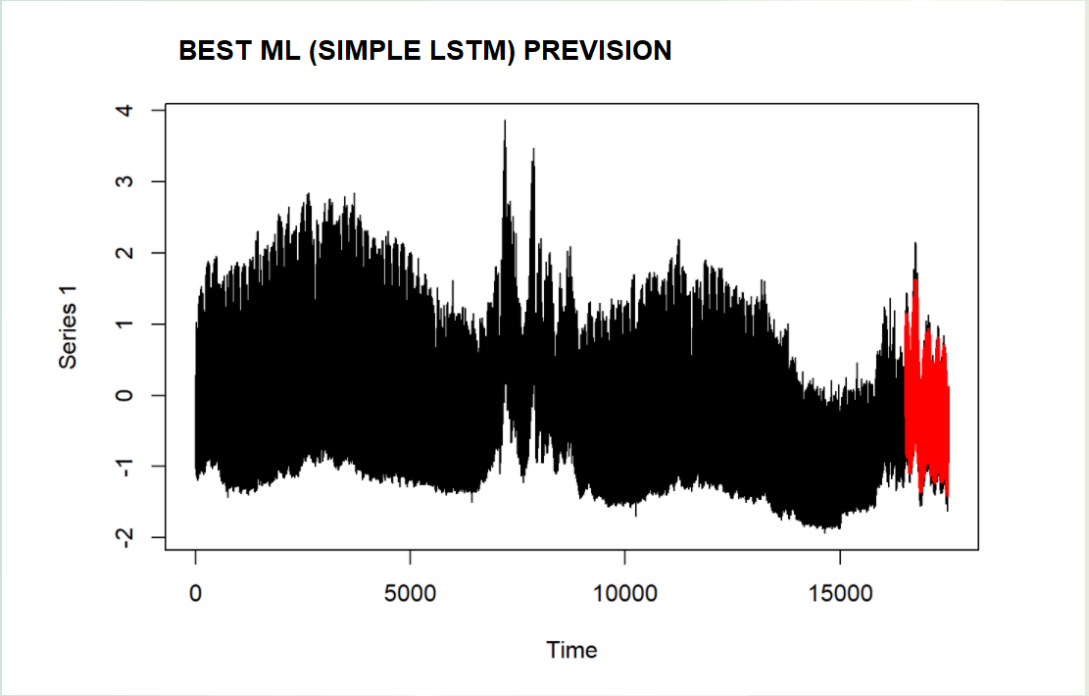




ML MODELS

MODEL	BEST MAE VALIDATION
Week KNN (k form 1 to 15)	0.4345685
Day KNN (k from 1 to 20)	0.4223891
LSTM (time step 1 week, 10 epochs) + dense	0.4086391
2 LSTM (time step 1 week, 10 epochs) +dropout + dense	0.3982217
GRU	0.3938358





THANK YOU

MODEL	MAE VALIDATION
naive	0.6563
ARIMA(2,1,1) with drift	0.5436
Arima (3,1,2)) with drift	0.5206016
Auto Arima (5,0,2)(2,1,0) [24]	0.496358
ARIMAX(5,1,2) fourier	0.4367472

MODEL	MAE VALIDATION
Local Linear Trend	0.6857667
LLT with daily stochastic dummy.	0.511847
LLT+seasonal dummy model+ cycle	0.5731635
trigonometric seasonality	0.6601218
Random Walk	0.4814821

MODEL	BEST MAE VALIDATION
Week KNN (k form 1 to 15)	0.4345685
Day KNN (k from 1 to 20)	0.4223891
LSTM (time step 1 week, 10 epochs) + dense	0.4086391
2 LSTM (time step 1 week, 10 epochs) +dropout + dense	0.3982217
GRU)	0.3938358