IBM Coursera Advanced Data Science Capstone

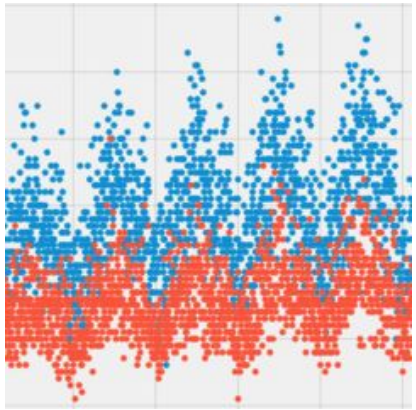
Patrick Schneider

Date:09.09.2018

https://github.com/patrick-s-upc/sales-forecast

Data set - demand prediction

- Time series data
- Kaggle - Store Item Demand Forecasting Challenge
- Task: Predict 3 months of item sales

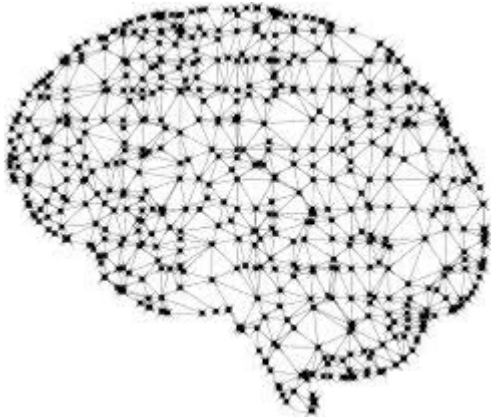


Use case - time series forecast

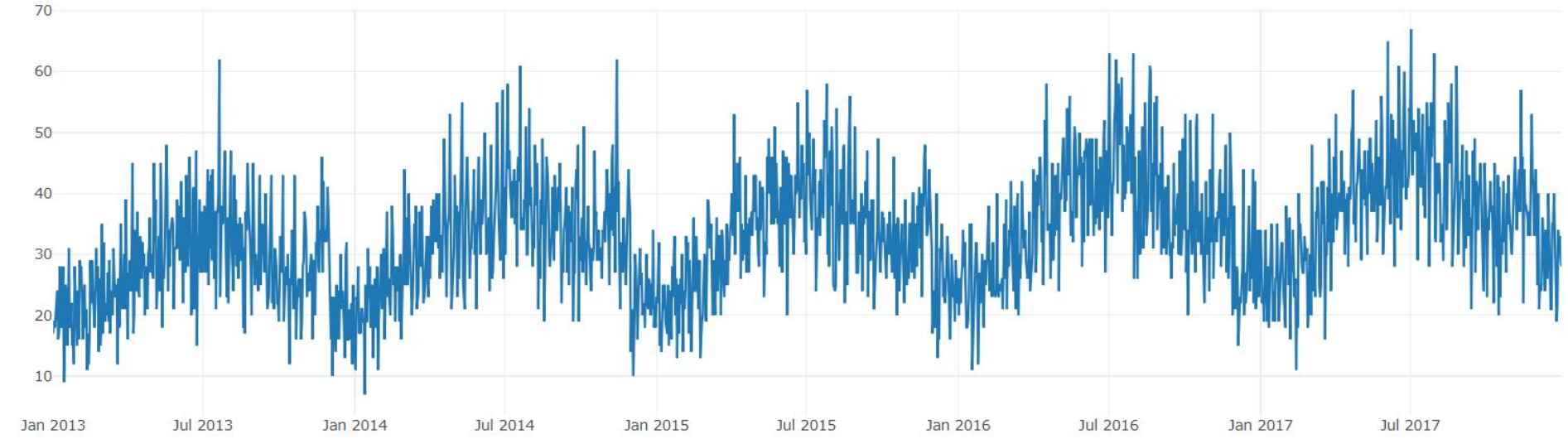
- short-term success of companies, products, and services
- Business forecasting for operational business - production plan, financing

Solution algorithm

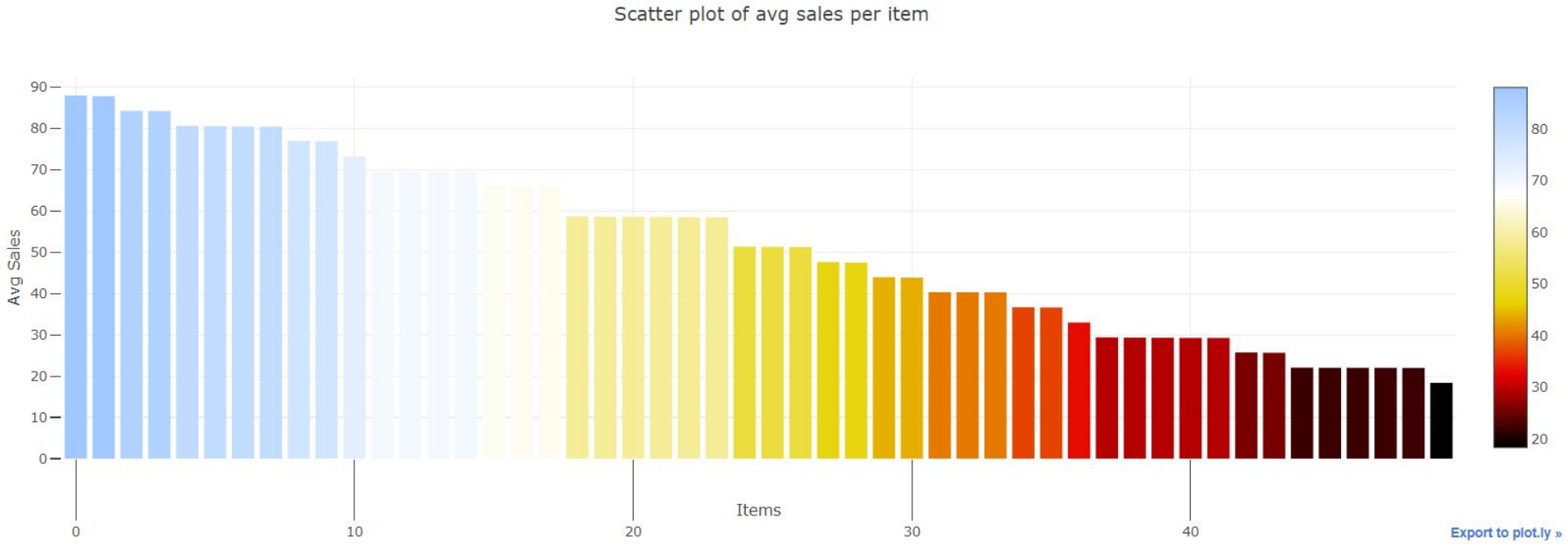
- ARIMA
- Neural Network
- Long short-term memory - NN
- XGBoost
- **Prophet (presentation topic)**



Data assessment

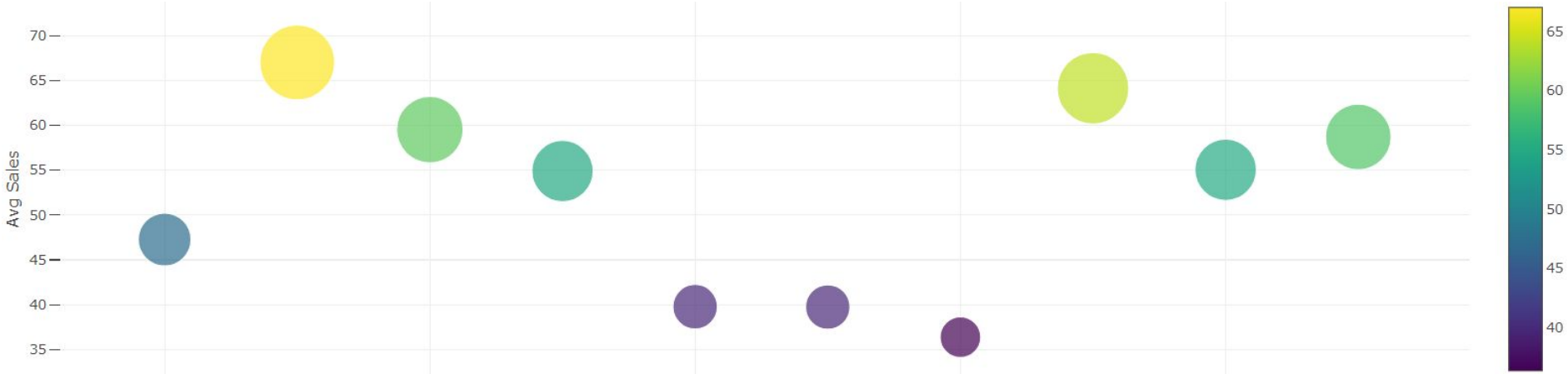


Data assessment

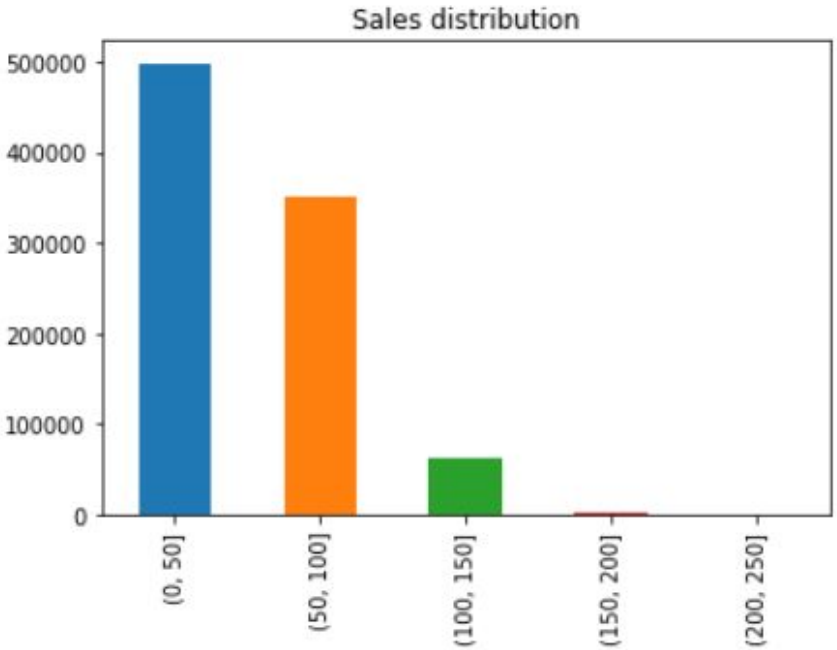


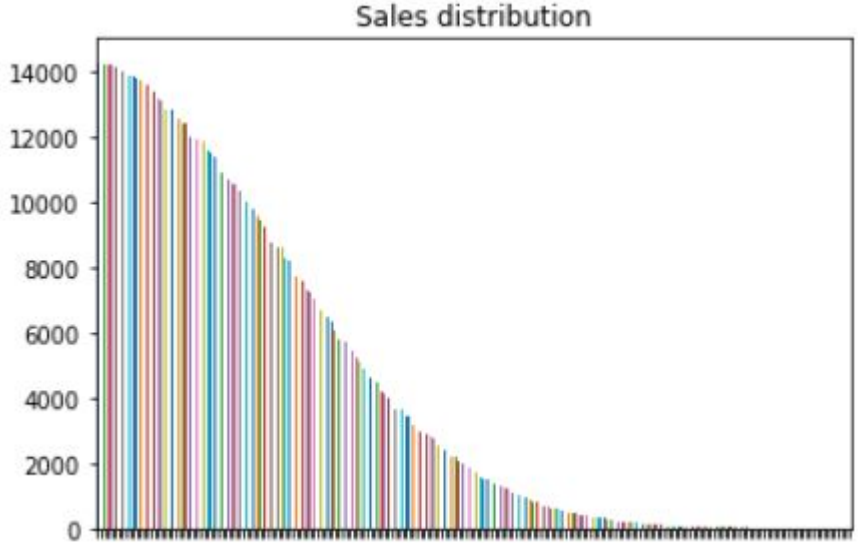
Data assessment

Scatterplot of individual store sales



Data assessment





Prophet algorithm components

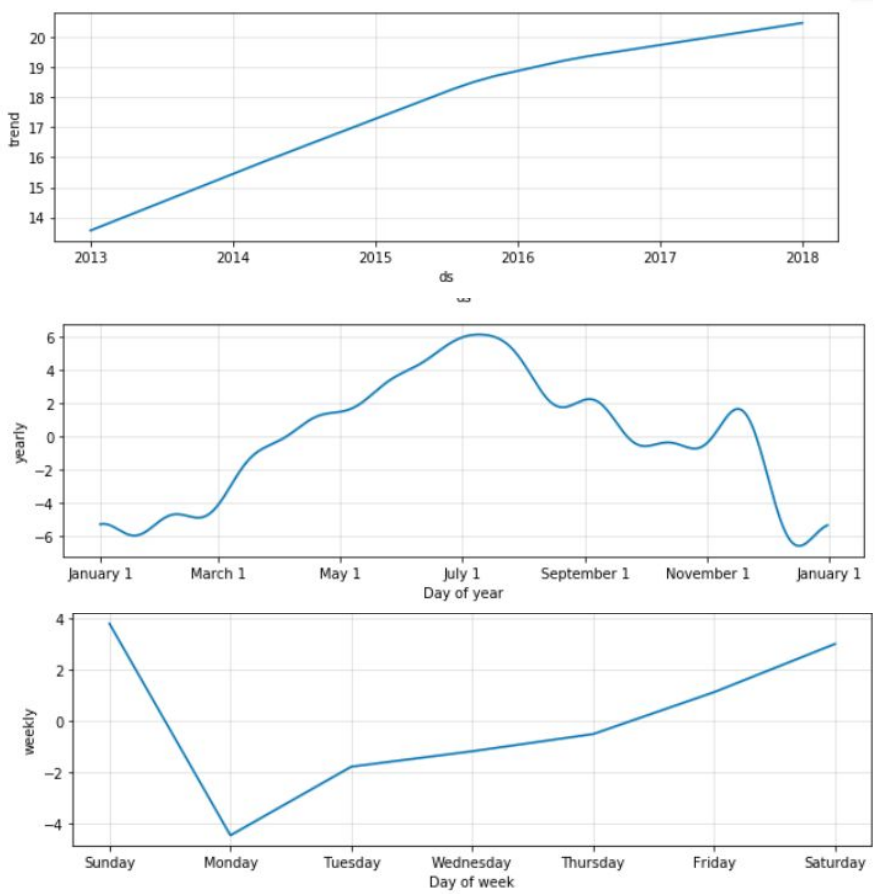
$$y(t) = g(t) + s(t) + h(t) + t$$

g (t) = Trends

s (t) = Seasons

h (t) = Cycles (holiday effects)

t = Irregularities (error term)


trend
20
19
18
17
16
15
14
2013
2014
2015
2016
2017
2018
ds
yearly
6
4
2
0
−2
−4
−6
January 1
March 1
May 1
July 1
September 1
November 1
January 1
Day of year
weekly
4
2
0
−2
−4
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week

Model performance

MAE	3.490
MAPE	26.26

Further Resources

Facebooks Prophet paper - https://peerj.com/preprints/3190.pdf

Jason Brownlee's Timeseries introduction blog - https://machinelearningmastery.com/