Healthcare Supply Chain Demand Forecasting

Laura Eshee

Capstone Project 2

Background

- “The supply chain generally refers to the resources needed to deliver goods or services to a consumer. In healthcare, managing the supply chain is typically a very complex and fragmented process.
- Healthcare supply chain management involves obtaining resources, managing supplies, and delivering goods and services to providers and patients. To complete the process, physical goods and information about medical products and services usually go through a number of independent stakeholders, including manufacturers, insurance companies, hospitals, providers, group purchasing organizations, and several regulatory agencies.”[1]

1. ‘Exploring the Role of Supply Chain Management in Healthcare’, *Recycle Intelligence*, August 5, 2016, https://revcycleintelligence.com/news/exploring-the-role-of-supply-chain-management-in-healthcare (accessed June 16, 2020)

Problem Statement

- The client is the largest healthcare system in its state.
  - 11 hospitals
  - 5 health parks
  - > 300 medical offices
  - 9 cancer centers
  - 55 rehabilitation centers
  - 3 hospice facilities
  - 21 imaging centers
  - 15 urgent care locations
  - > 24,000 employees.

Problem Statement

- In 2018:
  - 114,750 hospital admissions across the system
  - $4.1 million in assets
  - $3.2 million in revenue.
- Supply Chain department:
  - 1 distribution center
  - Approximately 3,600 items that are stocked at the distribution center
  - Multiple storerooms at its facilities
  - Forecasting for how much of each item should be stocked is done by hand by the manager.

Data Set

- Drug Abuse Warning Network (DAWN)
  - A public health surveillance system that monitors drug abuse related visits to emergency departments in hospitals in large metro areas across the US
  - Monitors trends in drug misuse and abuse
  - Identifies the emergence of new substances and drug combinations
  - Estimates the impact of drug misuse and abuse on the Nation’s health care system

Data Set

- The data will be obtained from the client's inventory database using Microsoft SQL Server.

- A two-year history of the products that have been issued from the distribution center to the facilities will be obtained

Approach

- Time series forecasting model
  - Baseline models will be built using algorithms and features to be defined
  - Once the performance characteristics of these models are established, other models and/or tuning approaches will also be attempted
  - All models built will be compared with respect to performance metrics that align with the business problem.

Data Set Preparation (Data Wrangling)

Data Acquisition

- Data was obtained from the client's inventory database using Microsoft SQL Server

- A two-year history of the products that have been issued from the distribution center to the facilities was obtained

Data Preparation

- The date column (TRANS_DATE) was imported as an object data type. Therefore, the first action was to convert it to a datetime data type

- The client requested that the unit of measure (UOM) used for forecasting be the tracked UOM and not the purchase UOM. Therefore, the quantity was converted

Data Story

Metrics

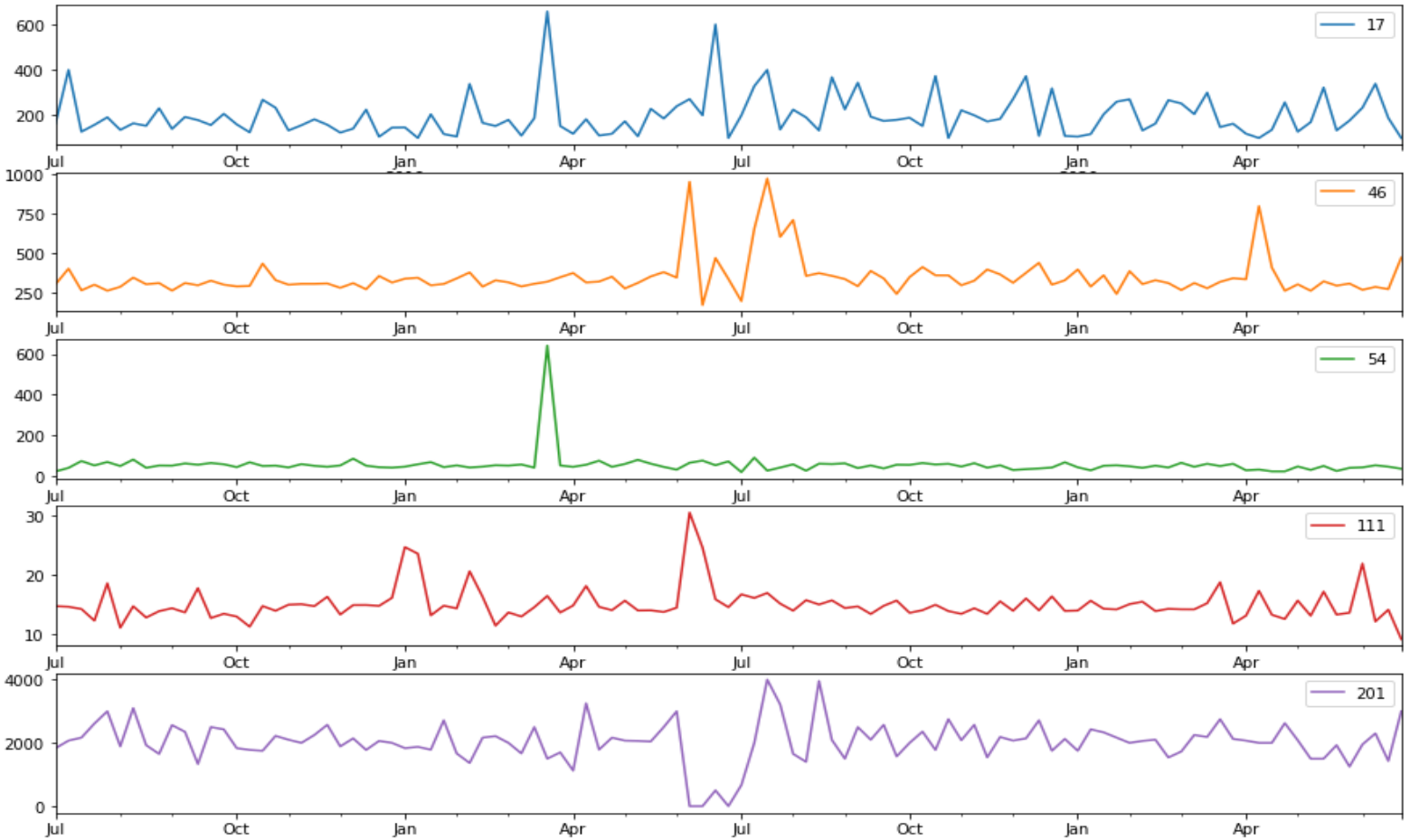
- Weekly forecast: the client asked for the forecasting to be done weekly. Therefore, the week starting date for each Trans Date was calculated. Then, the data set is checked to determine if there are any missing values.

- Missing values: it was found that several of the conversion columns were missing values, since not every item has the same number of UOMs. This finding was expected.

- Number of orders for each item: the number of orders for each item was determined. The item with the highest number of orders was Item #932, UNDERPAD INCONT30X36IN MOD ABS, with 24,728 orders, and the item with the least number of orders was Item #89, STENT PANCR 5-5 W INTNL, with 89 orders.

- Missing dates: whether there are missing dates for each item was determined. It was found that forty-five of the items were not ordered during one or more weeks in the two-year time period. It was decided to put zero for the quantity during the weeks where the items were not ordered.

Time Plots


17
600
400
200
Jul
Oct
Jan
Apr
Jul
Oct
Jan
Apr
46
1000
750
500
250
54
600
400
200
0
111
30
20
10
201
4000
2000
0

Data Story Summary

- 2 years’ worth of orders for 92 items in the warehouse.
- Sorted by item and week starting date.
- Inspected for missing dates.
- Reshaped so that it could be plotted with time plots and box plots
- Values for the missing dates were filled in with zeros

Inferential Statistics

The difference in the means of the quantities of 2 different catheters were analyzed

- Ho: mu1 = mu2
- Ha: mu1 != mu2
- Outliers were removed from both items
- Results:
  - Statistic: -37.572
  - P-value: 0.0
  - Reject Ho. The difference in the means is significant.

The difference in the means of the quantities of 2 different gowns were analyzed

- Ho: mu1 = mu2
- Ha: mu1 != mu2
- Outliers were removed from both items
- Results:
  - Statistic: -19.149
  - P-value: 0.0
  - Reject Ho. The difference in the means is significant.

The difference in the means of the quantities of 2 different hypodermic syringes were analyzed

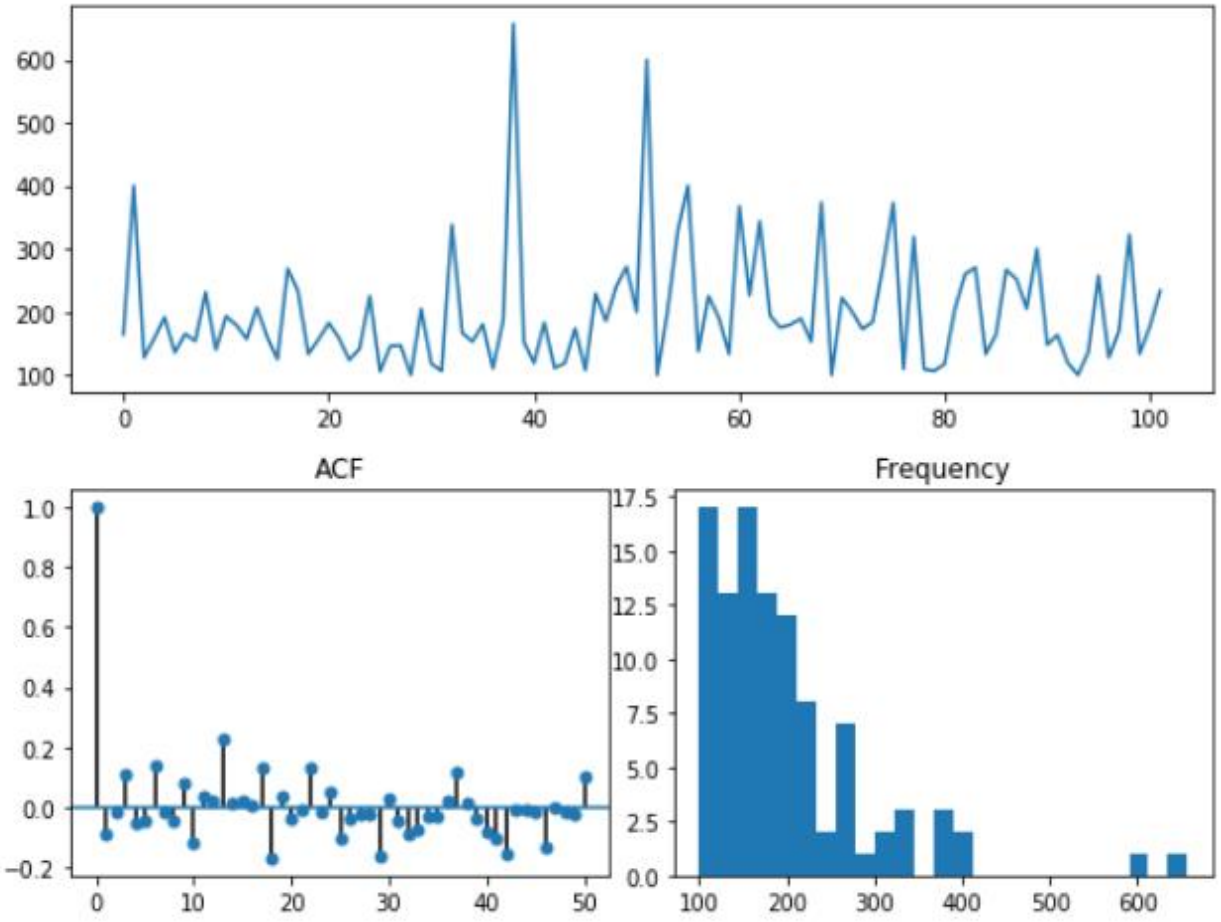
- Ho: mu1 = mu2
- Ha: mu1 != mu2
- No outliers were found for these items
- Results:
  - Statistic: -16.241
  - P-value: 0.0
  - Reject Ho. The difference in the means is significant.

Modeling

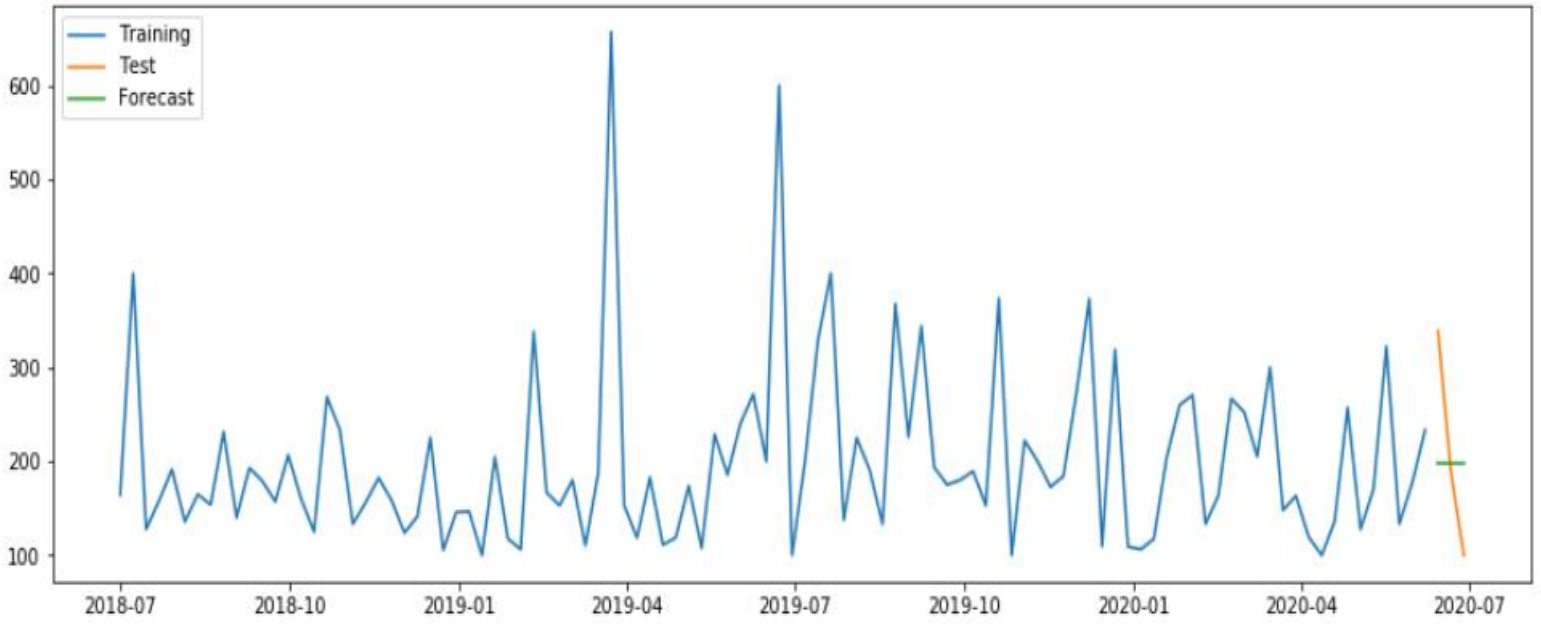
Auto-ARIMA

Item #17

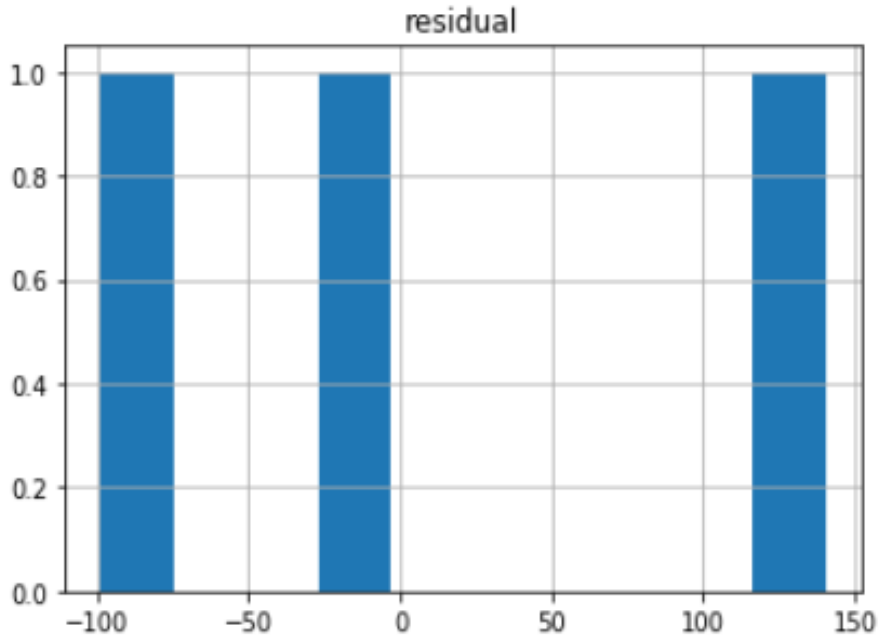
Time Plot, ACF Plot, Frequency Plot of Raw Data



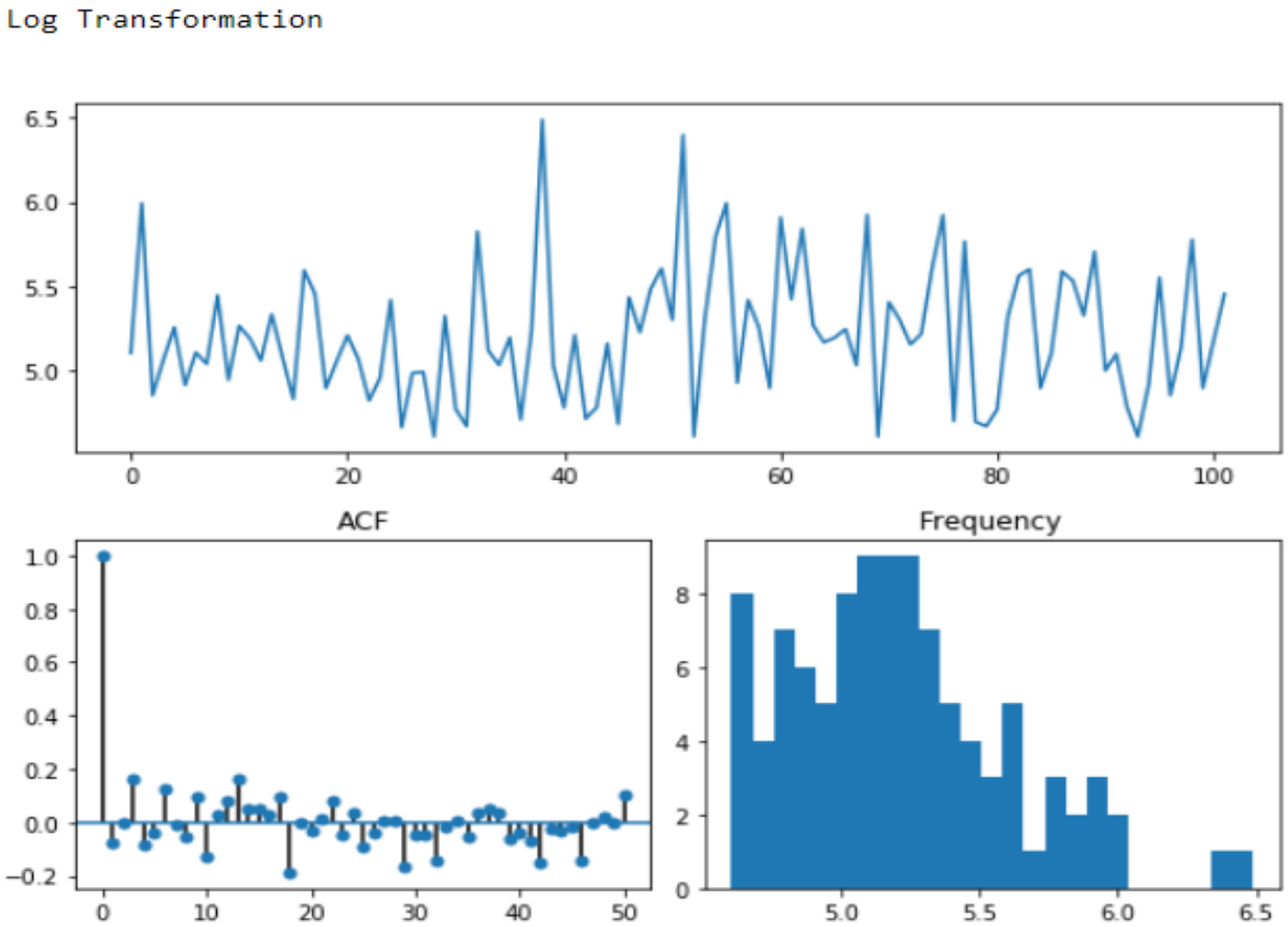
Plot of Raw Data Training Set, Test Set and Forecast



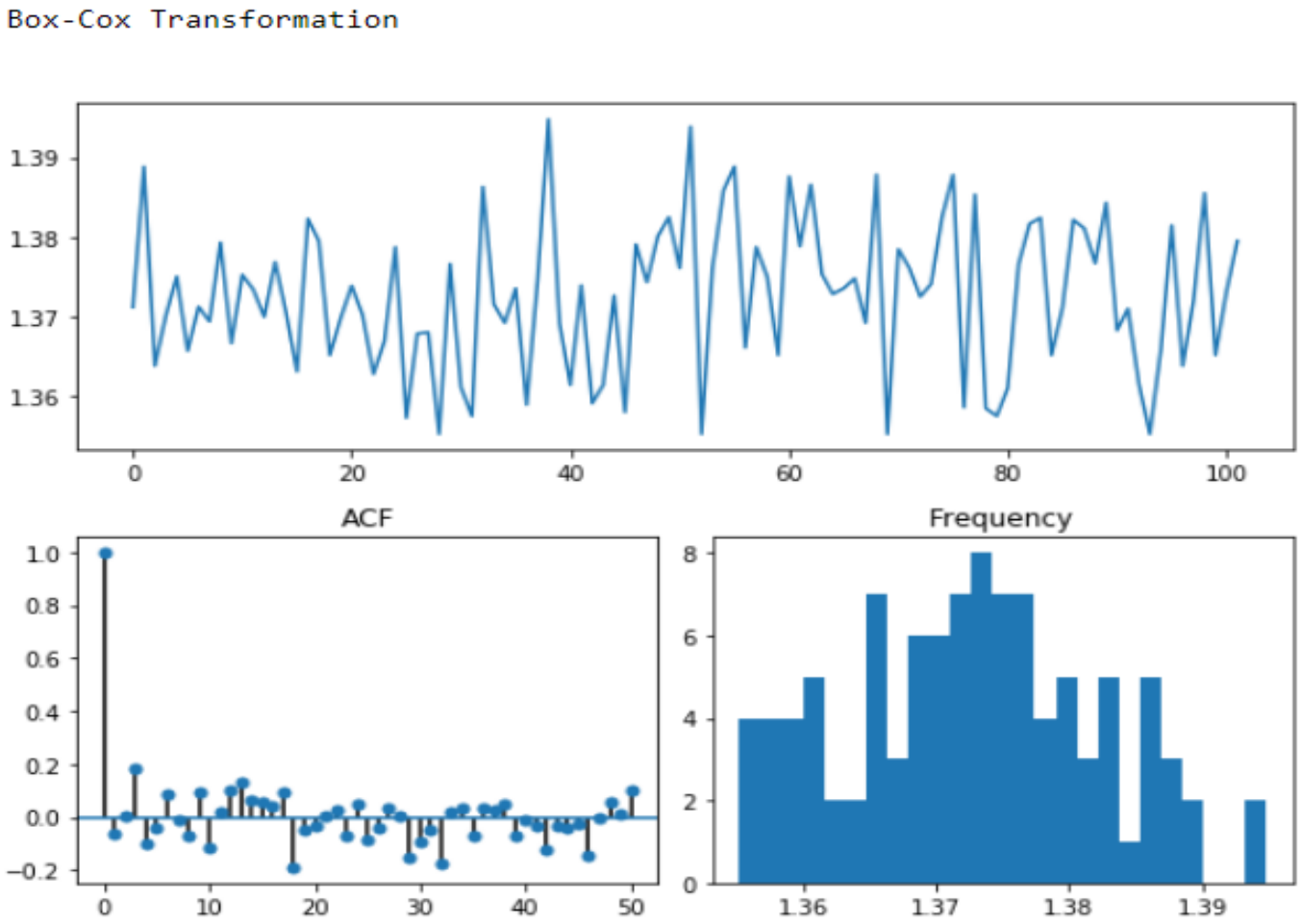
Residual Plot of Raw Data



Log Transformation



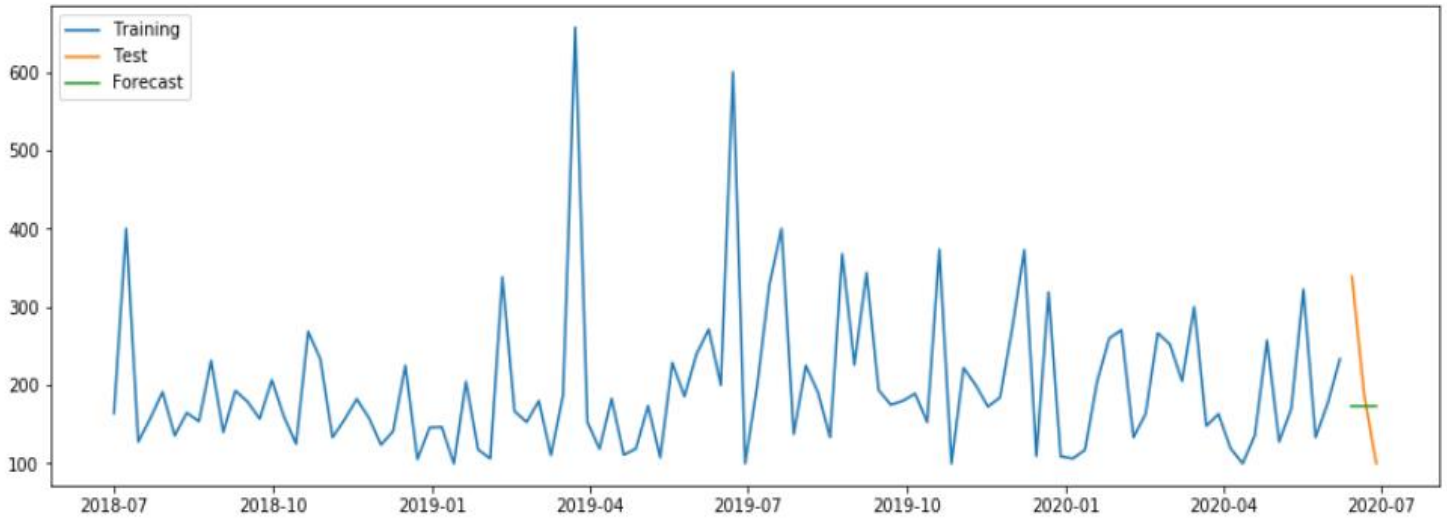
Box-Cox Transformation



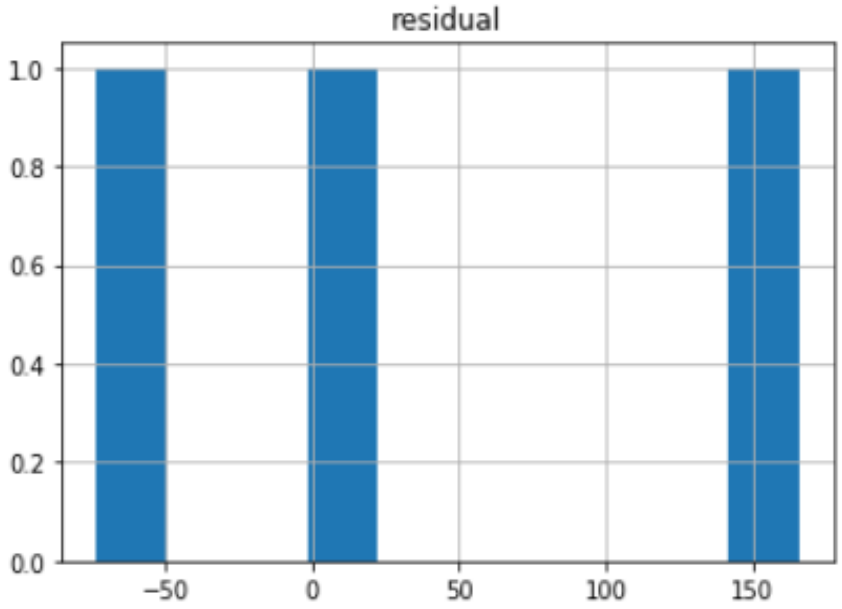
Box-Cox Predicted Quantity

Week_Start	Predicted_Quantity
2020-06-14	173.237237
2020-06-21	173.237237
2020-06-28	173.237237

Plot of Box-Cox Training Set, Test Set and Forecast



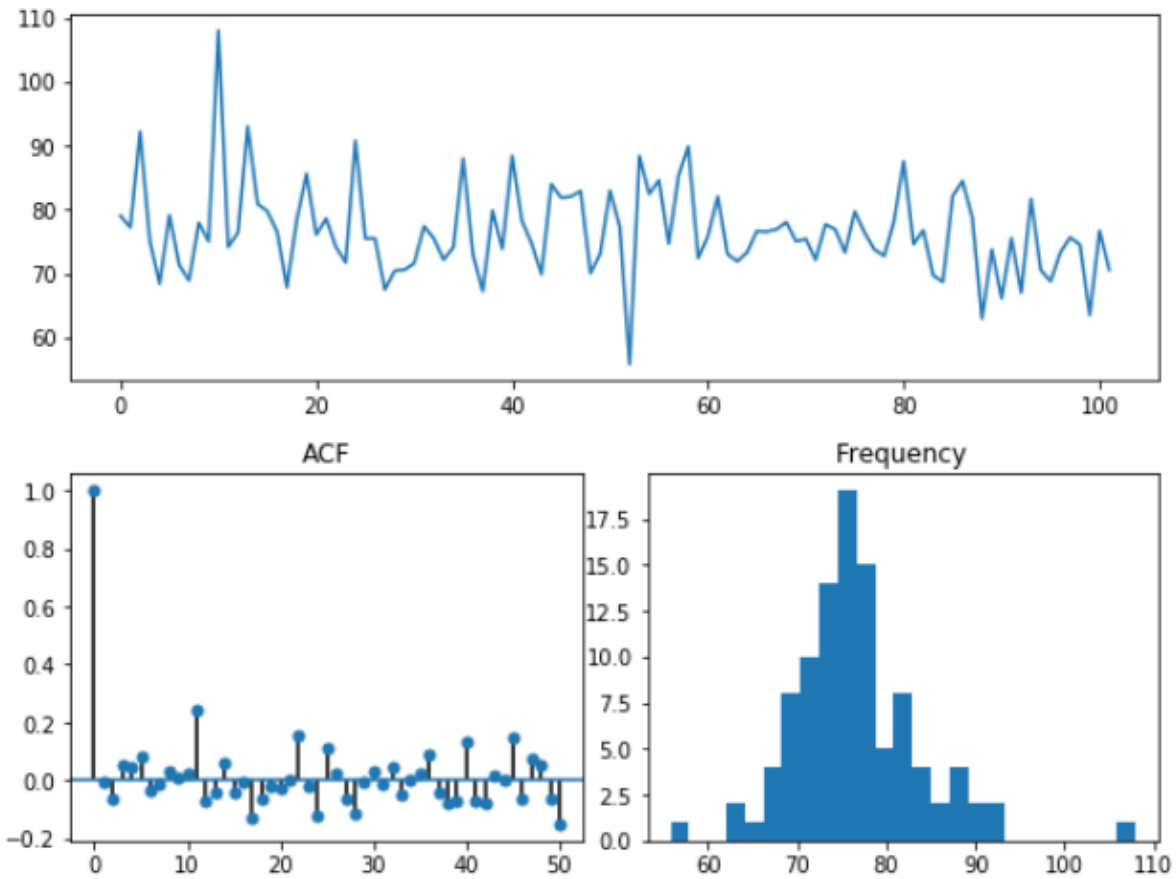
Box-Cox Residual Plot



Auto-ARIMA

Item #130173

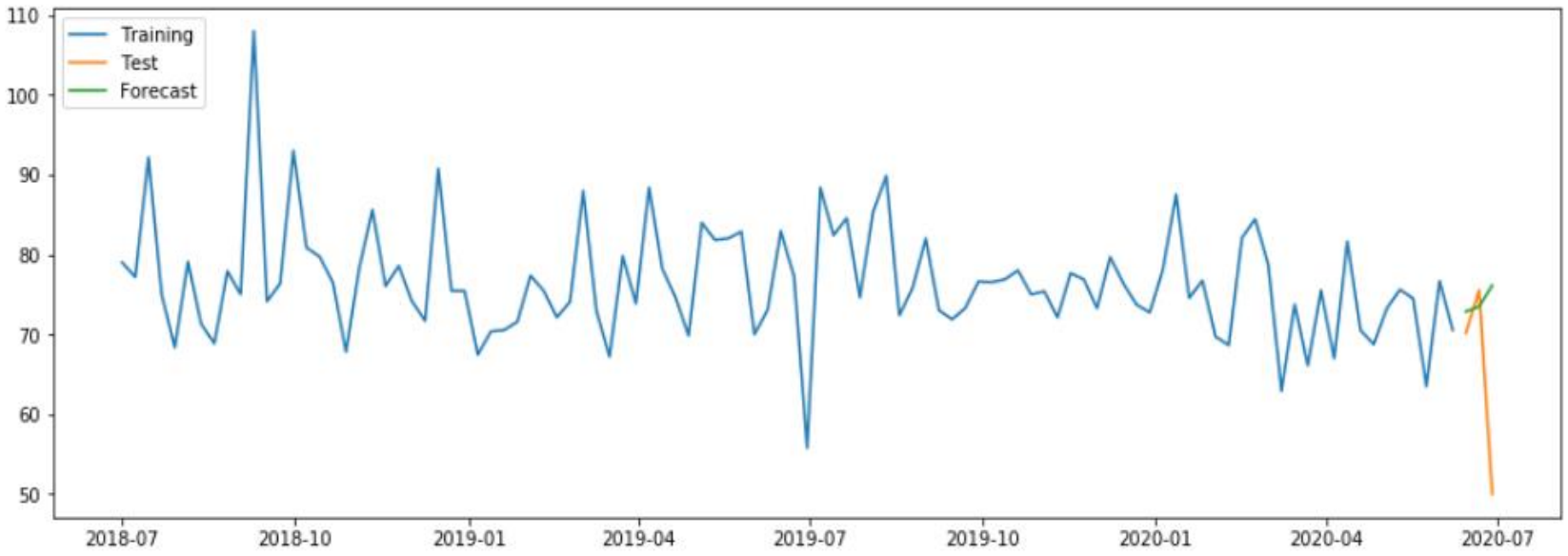
Time Plot, ACF Plot, Frequency Plot



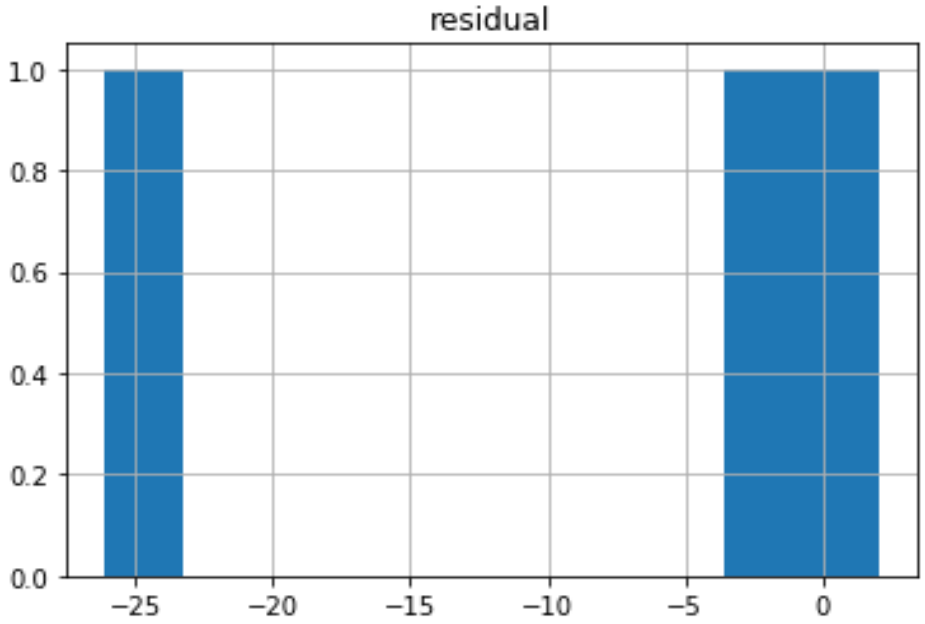
Forecast

Week_Start	
2020-06-14	72.844240
2020-06-21	73.479548
2020-06-28	76.113295

Plot of Training Set, Test Set and Forecast



Residual Plot



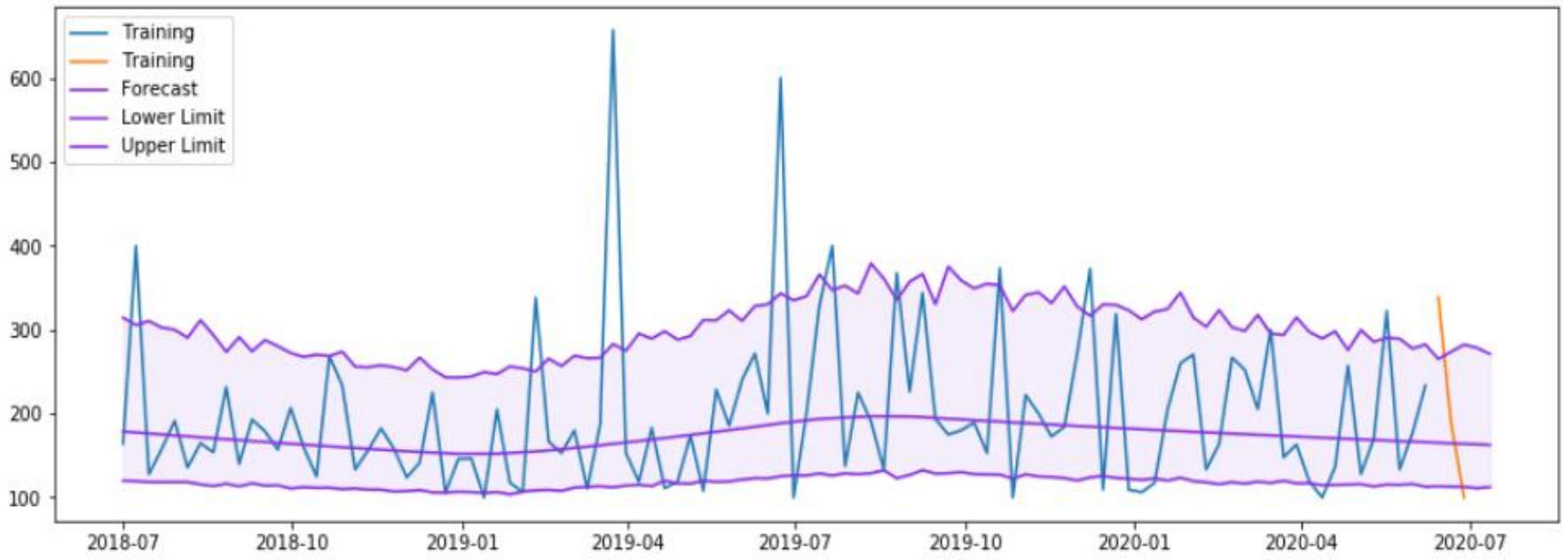
Prophet

Item #17

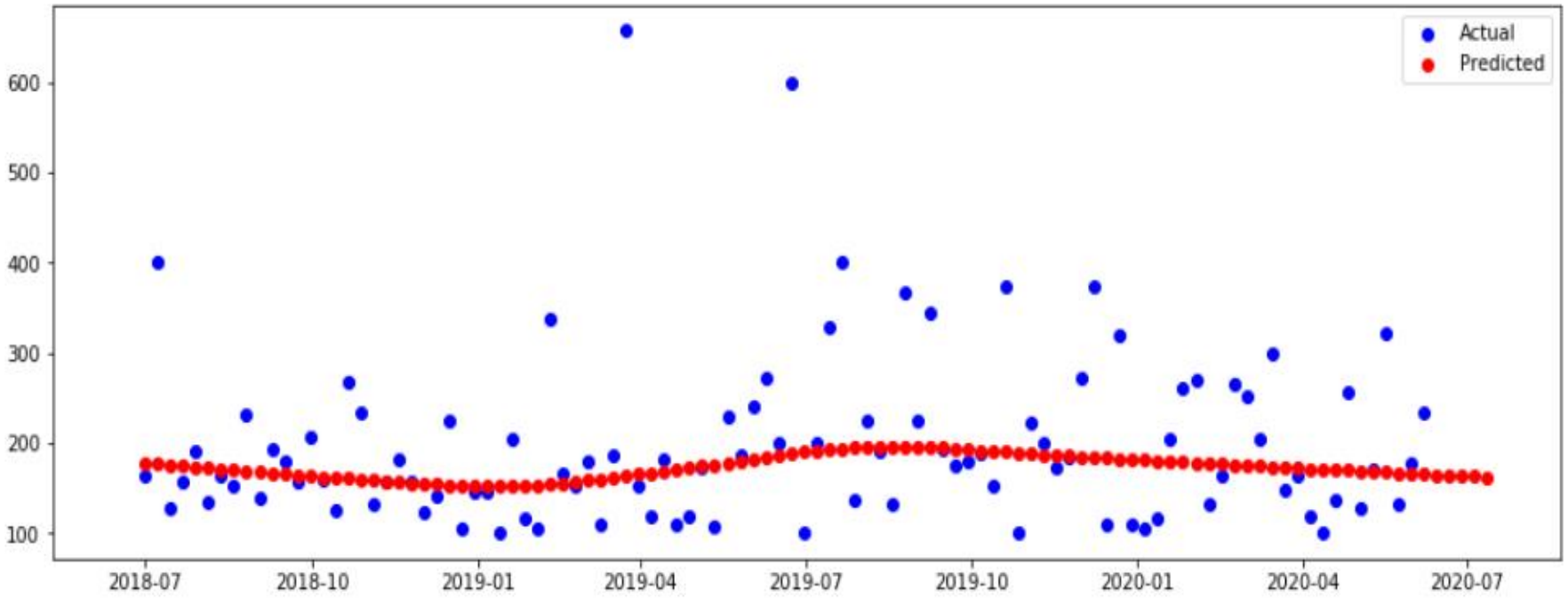
Forecast with Upper and Lower Limits

	ds	yhat	yhat_lower	yhat_upper
102	2020-06-14	165.206126	112.693411	289.951335
103	2020-06-21	164.564483	113.443048	283.172374
104	2020-06-28	163.927090	111.060929	281.858653
105	2020-07-05	163.293906	110.240731	266.603461
106	2020-07-12	162.664892	111.878927	267.716158

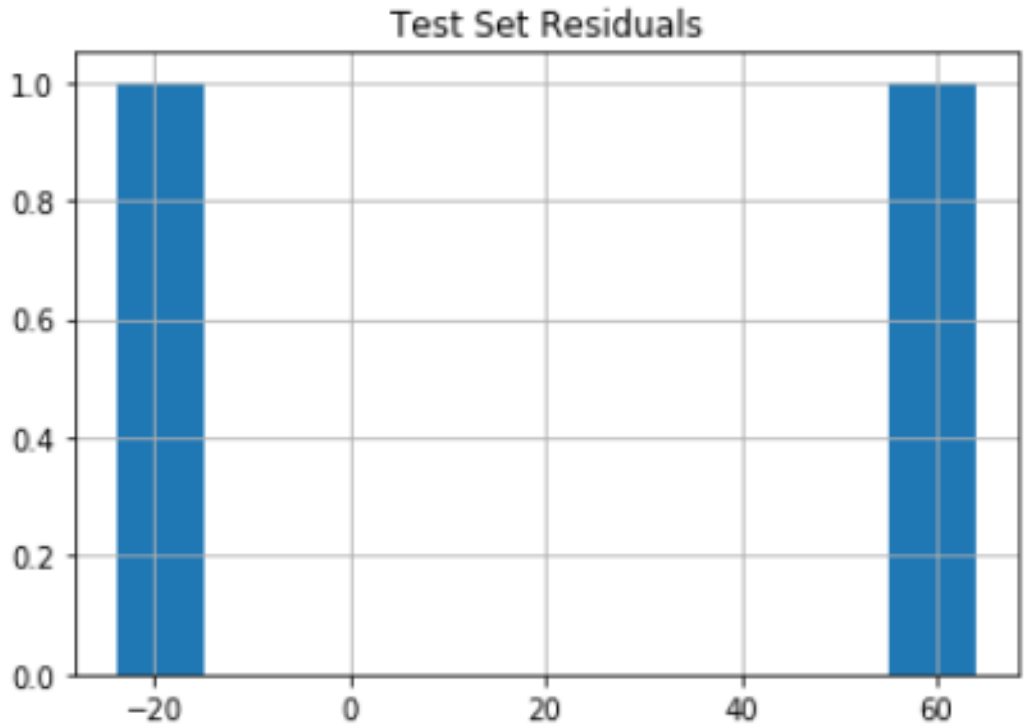
Plot of Test Set, Training Set, and Forecast with Upper and Lower Limits



Scatter Plot of Data with Forecast Line



Residual Plot



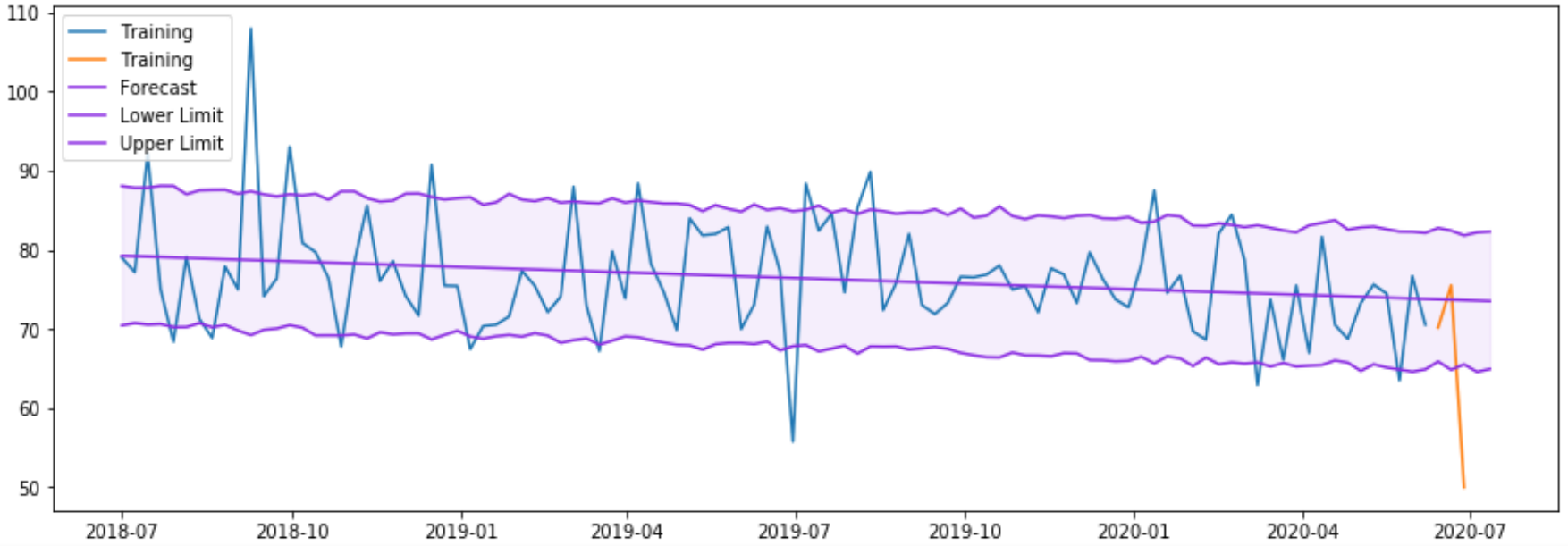
Prophet

Item #130173

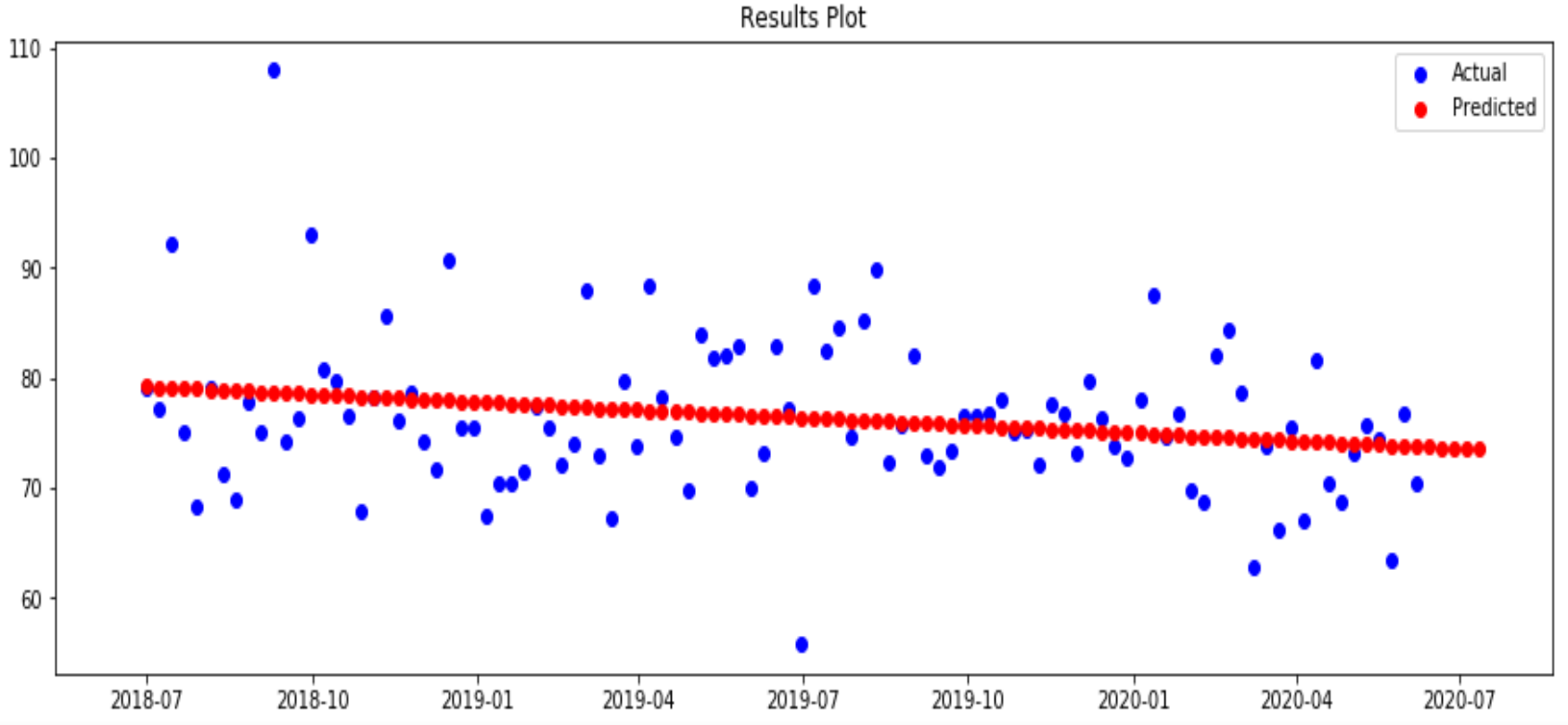
Forecast with Upper and Lower Limits

	ds	yhat	yhat_lower	yhat_upper
102	2020-06-14	73.736156	65.418828	82.937484
103	2020-06-21	73.682172	64918795	82032048
104	2020-06-28	73.628188	65.181435	82.341897
105	2020-07-05	73.574203	64.984683	82.537769
103	2020-07-12	73.520219	64.941713	82.540307

Plot of Test Set, Training Set, and Forecast with Upper and Lower Limits



Scatter Plot of Data with Forecast Line



Residual Plot

Summary

Auto-ARIMA Summary

Date	Item #17 Forecast	Item #17 Residual	Item #130173 Forecast	Item #130173 Residual
2020-06-14	173	165.65	73	-2.67
2020-06-21	173	14.99	73	2.03
2020-06-28	173	-73.24	76	-26.11

Item	R^2	MAE	RSME	MAPE
17	-0.1318	84.63	104.92	43.36
130173	-0.9143	10.27	15.20	18.76

Prophet Summary

Date	Item #17 Forecast	Item #17 Residual	Item #130173 Forecast	Item #130173 Residual
2020-06-14	165	-10.42	74	-2.82
2020-06-21	165	-67.48	74	3.25
2020-06-28	164	-173.68	74	3.56
2020-07-05	163	-23.68	74	-1.82
2020-07-12	163	63.93	74	23.63

Item	R^2	MAE	RMSE	MAPE
17	-1.8826	135.36	0.0025	0.0025
130173	-0.5861	9.67	2.75	2.75

Questions?