Springboard—DSC Program

Capstone Project 2 Proposal

Demand Forecasting for a Healthcare System Supply Chain

By Laura Eshee

June 2020

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]

The healthcare supply chain starts with the manufacturers of medical products, goes through the healthcare system's stocking and distribution process and ultimately to the provider's office, the hospital, or the administrative support areas of the system for usage.[1]

"Healthcare supply chain management is unique because each stakeholder has their own interests to protect. Different stages in the supply chain flow may be focused on their own goal. Providers may want to use a specific product because they were trained with it, whereas hospital executives aim to purchase the most affordable quality items.

Since supply chain goals are not always aligned within an organization, the healthcare supply chain management process can be inefficient and fragmented. Healthcare organizations must take into account numerous requests and viewpoints to settle on specific product budgets."[1]

According to an article on the Recycle Intelligence website titled '5 Ways to Improve Healthcare Supply Chain Management,' one way to improve health care supply chain management is to develop effective inventory management. Inventory levels must be managed effectively. Too much stock will result in high costs for storing and maintain it. Too little stock will result in not having enough to meet demand and having to pay a premium to get what is needed quickly. As a result, good forecasting for inventory levels is needed.[2]

Problem Statement

The client is the largest healthcare system in its state. It has eleven hospitals, five health parks, more than three hundred medical offices, nine cancer centers, fifty-five rehabilitation centers, three hospice facilities, twenty-one imaging centers and fifteen urgent care locations. It also has more than twenty-four thousand employees.

In 2018, there were 114,750 hospital admissions across the system. Additionally, it had $4.1 million in assets and received $3.2 million in revenue.

The Supply Chain department of this client consists of one distribution center, as well as, multiple storerooms at the facilities. There are approximately 3,600 items that are stocked at the distribution center. Currently, forecasting for how much of each item should be stocked is done by hand by the manager.

Dataset

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

The approach that will be used to model this problem will be a 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, and all models built will be compared with respect to performance metrics that align with the business problem. In general, the classical phases associate with many data science problems will be implemented, namely: data acquisition and wrangling, storytelling and applications of inferential statistics, model building and ranking.

Deliverables

As required, all Jupyter notebooks, a written final report, and a presentation slide deck will be submitted at the end of this project.

Data Wrangling

The 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. 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. Next, the client requested that the unit of measure (UOM) used for forecasting be the tracked UOM. The tracked UOM is denoted in the data by an X in one of the BUYFL columns. Therefore, the transaction UOM was converted to the tracked UOM using the BUYFL columns. The data wrangling for this data set was fairly simple. First, the date column data type was converted from object to datetime. Then, the transaction UOM was converted to the tracked UOM and the tracked quantity was computed. The data set was then saved to a csv file. It is now ready for the next step.

Data Story

With the data wrangling completed, it is time to see what story the data tells. The following metrics were calculated on the data set:

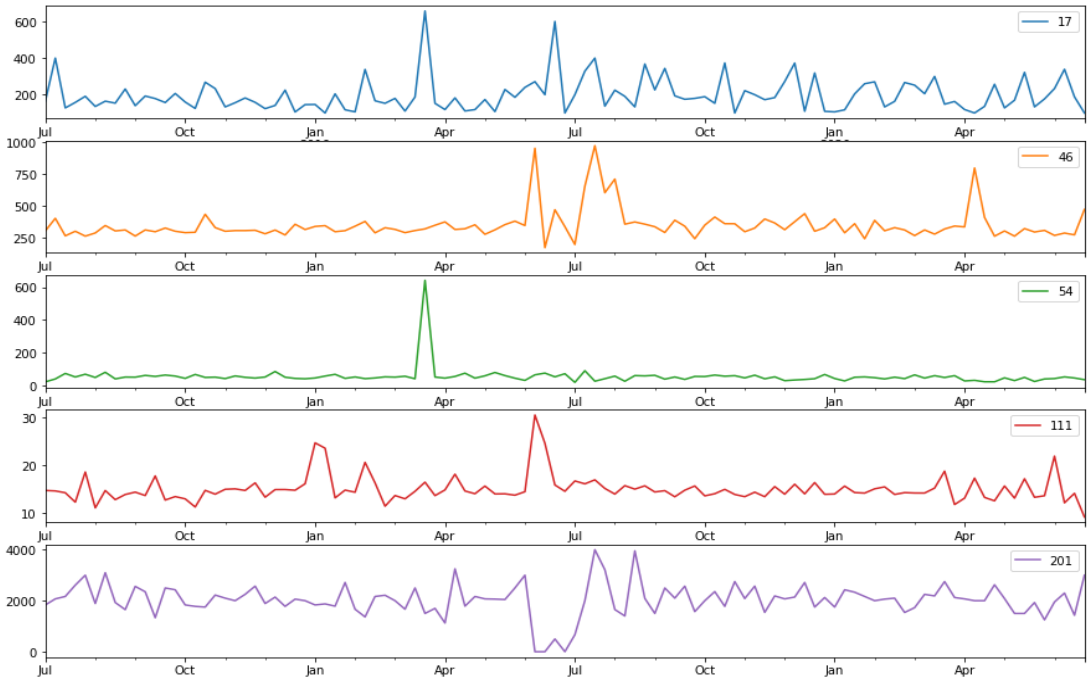
- Weekly forecast: the client asked for the forecasting to be done weekly. Therefore, the week starting date for each Trans Date was calculated. Then, the data set is checked to determine if there are any missing values.

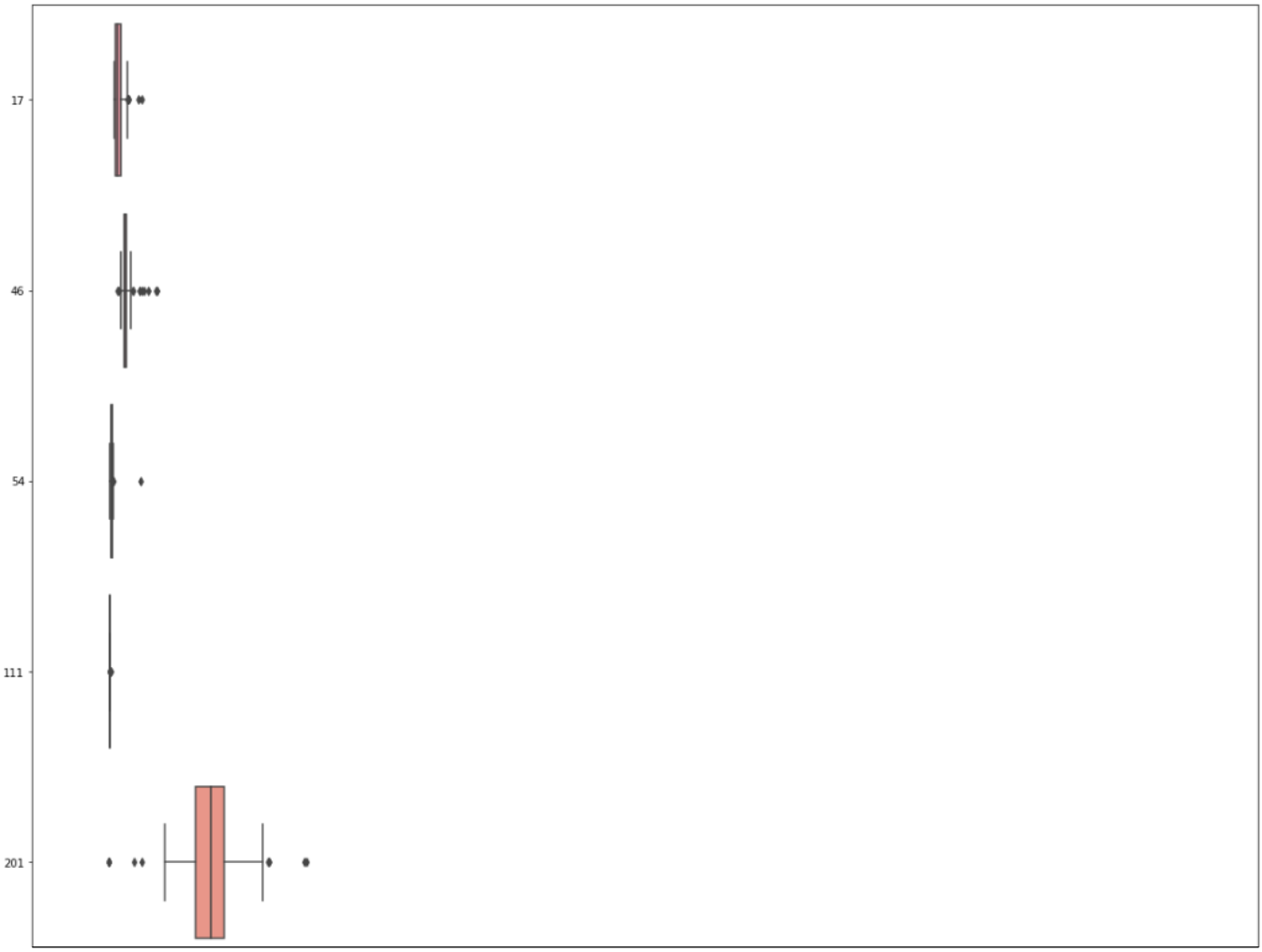
- Missing values: it was found that several of the conversion columns were missing values, since not every item has the same number of UOMs. This finding was expected.

- 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.

- Plotting: the data frame was now reshaped so that time plots and box plots could be created for each item. A sample of the time plots and box plots are included here. To see the plots for all of the items, go to https://github.com/lookingglass01/springboard/blob/master/CapstoneProjects/CP2/code/Demand%20Forecasting%20Data%20Story.ipynb.





In summary, this data set has 2 years' worth of orders for 92 items in the warehouse. The data was sorted by item and week starting date. Then, it was inspected for missing dates. The data was reshaped so that it could be plotted with time plots and box plots, and the values for the missing dates were filled in with zeros. The final data set was saved so that it will be ready for the next step in the project, which in inferential statistics.

Hypothesis Testing

Next, hypothesis testing was performed on the data set. It was decided to test the mean orders of items in the same class to see if the means were equal. The results are as follows:

- 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 means are not equal.
- 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 means are not equal.
- 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 means are not equal.

Conclusion

In summary, this milestone report shows the process for getting the data set ready for analysis, the plots done that show the story the data is telling, and the inference testing that was done. The project is now ready to move into the next phase, which focuses on machine learning and forecasting.

Footnotes:

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)

2. '5 Ways to Improve Healthcare Supply Chain Management', *Recycle Intelligence*, May 19, 2016, https://revcycleintelligence.com/news/5-ways-to-improve-healthcare-supply-chain-management, (accessed June 16, 2020)