DtonesOD1 / **Capstone**   Public

<> **Code**   ⊙ Issues   ⅊ Pull requests   ▷ Actions   ⊞ Projects   📖 Wiki   ⊘ Security   📈 Insights   ⚙ Settings

ᛉ main ▾   ···

Capstone / README.md

🐢 **DtonesOD1** Update README.md   🕐 History

🐾 **1** contributor

☰   54 lines (34 sloc) │ 4.79 KB   ···

Capstone

Business Overview & Understanding

Business Overview

Fire Maul Tools is a firefighter-owned company serving firefighters, tactical teams, and military personnel worldwide by providing innovative tools that were designed for what they do. The company started by producing handmade firefighting tools such as axes and halligan bars, but since then has expanded their product breadth to include items such as tool lube and grip kits. The grip kits are just what they sound like, it is a fiber tape that comes with a lacing ring system that greatly improves grip on the tools. The item has proved to be very popular and the business owner wants to make sure he is getting the most out of the product while keeping his cost down. That is why Fire Maul Tools has tasked me with forecasting the future demand of the grip kits. This way the company can plan its purchasing and inventory levels more effeciently which will lead to lower costs associated with excessive inventory or stockouts. The business owner has stated that since he is a start-up small business his shop is relatively small and does not have a lot of room to inventory the grip kit materials which are shipped to him by the box on pallets. In the past he has been short on supplies when orders come in, but can not afford to take up excess space by over stocking. By following the Data Science methodology and utilizing time-series machine learning techniques, I plan to accomplish solve this problem and help Fire Maul Tools grow as a company so that they can continue to serve First Responders.
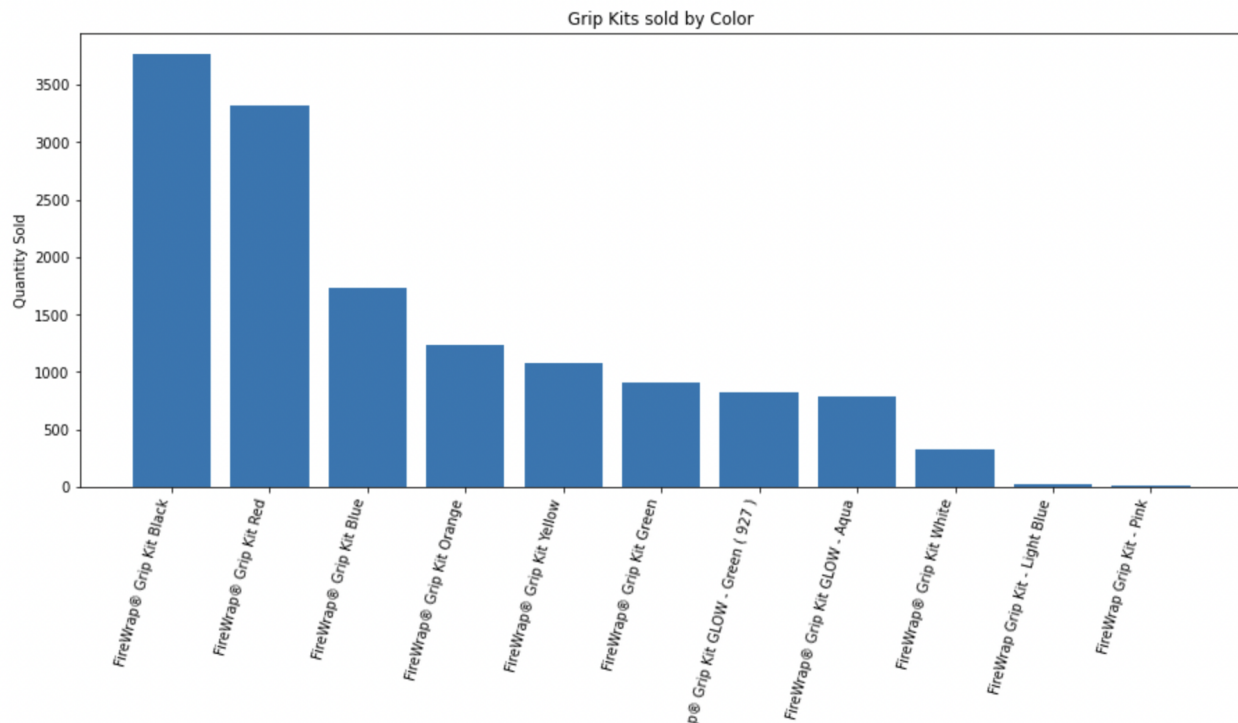
Data Understanding

To complete this task, I have been given a dataset with the needed information that will be used with our models. Again, the standard Data Science methodology will be followed:

- Obtain the Data
- Clean the Data
- Exploration
- Model
- Interpret

A look at the distribution of grip kits sold('Qty)by color:

Data Prep

After the data was cleaned and we had a grasp of what the data entailed we began to prepare for the modeling process. I was attempting to be able to predict the average number of units sold per week using the quantity column. After an initial run at it, it was decided to remove the outliers and to split the data into a 80/20 train test set. With the dataset being 5 years long it was easy enough to make the most recent year as the test set. After that the dataset had be entered into datimetimeindex, formatted to W-sat, which sets the dates to weeks beginning on Saturday

```
df_black_outlier = df_black[df_black['Qty']<=300]
df_black_outlier['Qty'].plot()
```
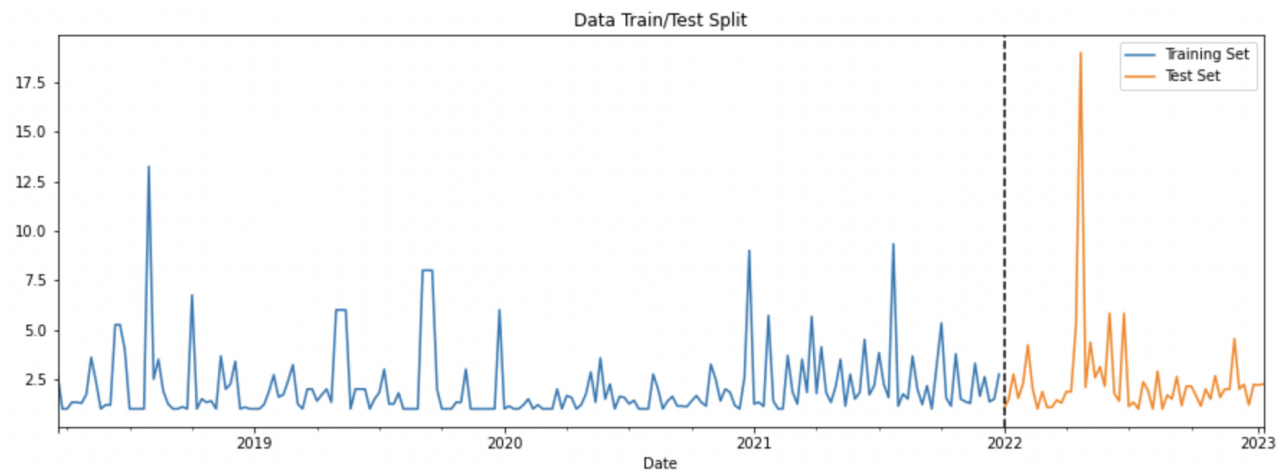
```
<AxesSubplot:xlabel='Date'>
```
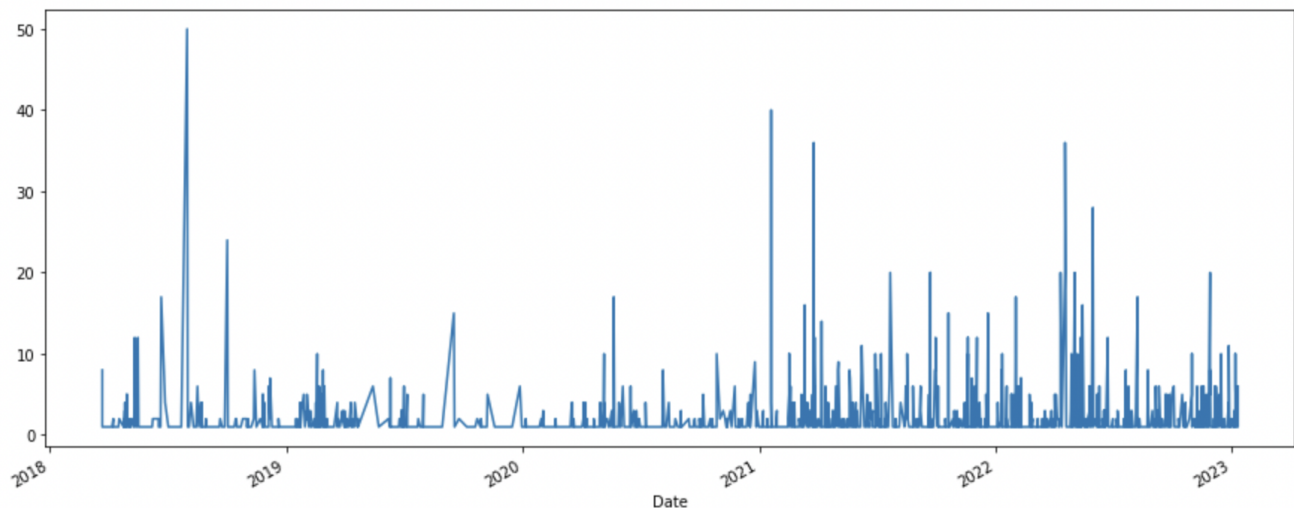


```
check_stationarity(df_black_outlier['Qty'])
```

```
The series is stationary
```

Models

The method used for the models was based off the SARIMA appendix from Phas3 4. A range of parameters were set for (p,d,q) and (P,D,Q,s) and a SARIMAX grid search was performed to find the optimal paramters. These parameters were fit to the model and then checked that they did not violate assumptions.

```
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ma.L1         -1.0404      0.039    -26.636      0.000      -1.117      -0.964
ma.S.L51      -0.8068      0.295     -2.734      0.006      -1.385      -0.228
sigma2         3.8509      1.113      3.460      0.001       1.669       6.033
==============================================================================
```

```python
# Call plot_diagnostics() on the results calculated above
output_1.plot_diagnostics(figsize=(15, 18))
plt.show()
```
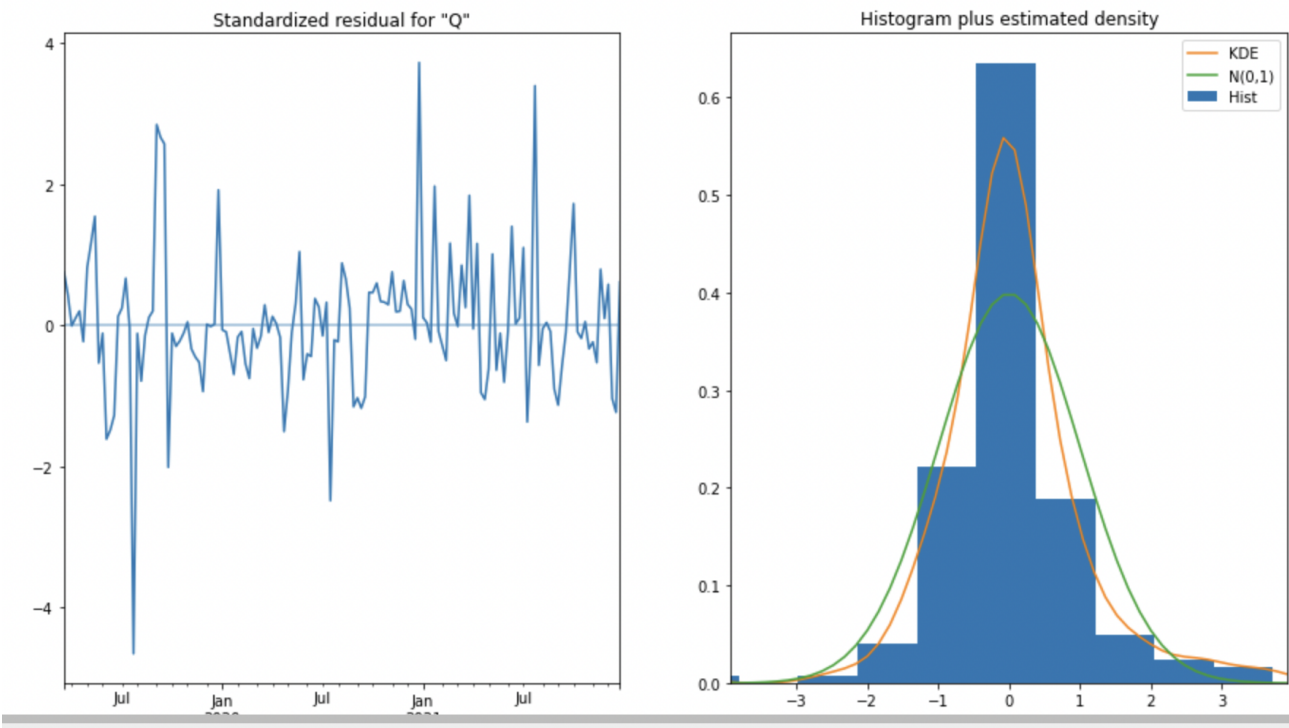

After the data was fit we used prediction methods such as 'One-step ahead', 'Dynamic' and 'get_forecast' to visualize and print out predicted values and confidence intervals for 'Qty' which is the number of units sold.
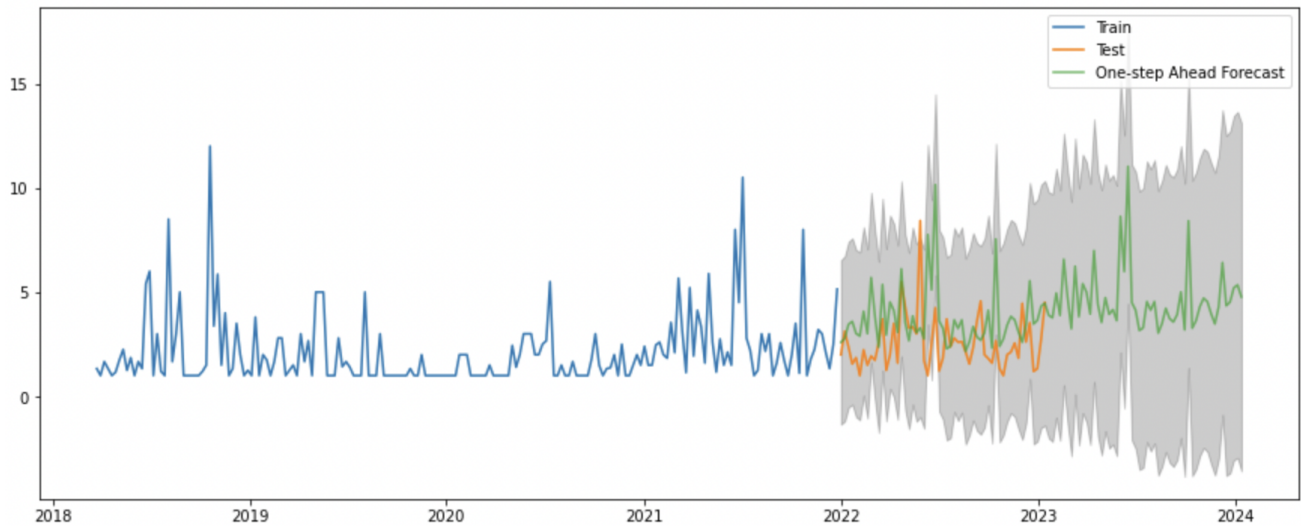
This method was performed 3 times for the top selling grip kit colors, which were Black, Red, and Blue, these 3 made up over 60% of sales. Therefore there are 3 SARIMA time series models in the notebook.

Evaluation

After the models were run and the forecasting methods compared, it was decided that the 'get_forecast' method produced the best results for all 3 models. Here are visuals that show the predicted average number of units that will be sold each week for the next year, broken down by color.

```
The average quantity prediction per week for the next year for Black Grip Kits is:  2.377730540886904
The average quantity prediction per week for the next year for Blue Grip Kits is:  2.3650596983399303
The average quantity prediction per week for the next year for Red Grip Kits is:  3.707200237295367
```

A lot of effort went into this and I hope it can be useful for FireMaul Tools!

Contact Info: Brian O'Donnell briodonn1021@gmail.com

Give feedback