Introduction

PUMA is the world's premier dynamic lifestyle brand that designs and develops a wide range of footwear, apparel and accessories. The company was founded in the sporting goods industry and eventually moved into the world of fashion. Puma shoes and apparel are extremely popular with young people around the world in the hip-hop graffiti culture. As a result, Puma has a first-line appeal and influence in the world. In the high-end fashion boutiques PUMA offers the most unique and creative fashion products. It always succeeds in combining professional sports and fashion. It has always succeeded in perfectly blending professional sports and fashion trends together, making people with diversified needs, get multi-functional satisfaction to guide the trend of sportswear and shoes worldwide.

Industries are facing supply chain disruptions as a result of the COVID-19 pandemic. There is a shortage of shipping containers, a shortage of labor, a shortage of parts because of reduced shipping. This affects the company's inventory management and is not conducive to cost control and growth. Therefore, we can use machine learning predictive analytics to solve supply chain problems. According to Gartner, demand forecasting is the most widely used machine learning application in supply chain planning. The study highlighted that 45 percent of companies are already using the technology, with 43 percent of them planning to use AI-driven demand forecasting within two years. Demand forecasting helps companies reduce supply chain costs and make significant improvements in financial planning, capacity planning, profitability and risk assessment decisions. At the same time, AI predictions can reduce errors in supply chain networks by 30 to 50 percent. On the one hand, machine learning not only improves the accuracy of demand forecasting, but also automates a large number of planned tasks and processes huge data sets. On the other hand, demand management can be better managed by analyzing the growth of the company. Companies use it to anticipate and plan how to meet demand for services and products. Demand management improves the link between operations and marketing. The result is a close alignment of strategy, capabilities, and customer needs.

With the development of the market economy, market competition is becoming more and more intense. In order to win the competition and win customers, companies must provide their products to customers in the shortest time and at the lowest cost, which requires companies to make correct and timely product sales forecasts to generate decisions in order to accomplish their goals better faster. Sales forecasting has become a key element of success for all businesses, especially in the apparel industry. It is not easy for enterprises to make high-quality forecasts, so it is necessary to understand the factors that affect sales forecasts. The sales forecast is mainly affected by two major factors: one is external factors, including demand trends, economic changes, industry competition trends, government and consumer groups trends; The second is internal factors, including marketing strategy, sales policy, sales personnel, production status.

This report demonstrates the importance of forecasting sales for the retail industry, while analyzing the factors that affect sales to help companies increase sales and promote the company's growth. With the development of the market economy, different companies are using different methods to compete in the market and increase their market share. In order to win the competition and attract more customers, companies must deliver their products to customers in the shortest time and at the lowest cost, which requires companies to make correct and timely product sales forecasts to generate decisions to accomplish their goals better and faster. We solve two main problems. Firstly, help PUMA predict future sales. Secondly, how to do inventory management in advance to avoid both inventory storage and overstocking. Therefore, this project will analyze the activity dataset as this dataset contains variables such as orders, cancellations, and shipments by product and date. The ARIMA model and LSTM Network to analyze which factors affect sales and to provide some actionable recommendations.

Data Cleaning

The Activity data set contains PUMA sales data from January 2020 to April 2022. And the Product data set contains information about all the products currently being sold by PUMA. Since there is a lot of raw data, we decided to merge the two datasets and do some processing on the data. First, we removed some useless columns, such as ‘IPC Flag’, ‘Cust ID’ and ‘ASIN’.

After normalizing the name of each column, we merged the "Activity" and "Product" datasets according to the UPC (product's 12-digit identifier). Some rows missing the value of Reporting Business Unit were removed. And the missing values in the variable Main Color Group were replaced with "unknown'. We also converted the time variable into the standard datetime variable. After that, we converted the missing values to 0 in all variables involving order quantity and order value, which is beneficial to our subsequent calculation.

We combined all normally completed or pending sales and canceled order data to get the total number of orders placed by all customers in each period of time, which is the variable we are going to predict (Total Ordered Qty). The Cleaned dataset has 2,064,371 rows and 10 effective variables. After getting the cleaned data set, we will make a preliminary interpretation of the data set.

Exploratory Data Analysis

In the EDA part, we will conduct a preliminary analysis of the current dataset. When we just look at the surface of the data, we cannot know some of the actual operation conditions of PUMA. So our goal is to discover some facts hidden behind the surface of data. After cleaning the dataset, we selected some variables that can be discussed and made some crosstabs. Through these charts, we have found some operational strategies and rules of PUMA.

Figure 1: Total Ordered Qty by Product Division

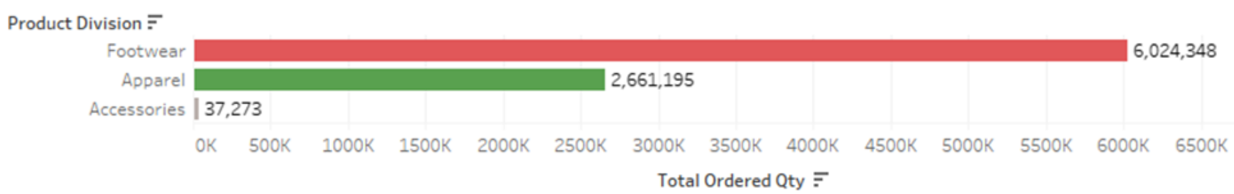
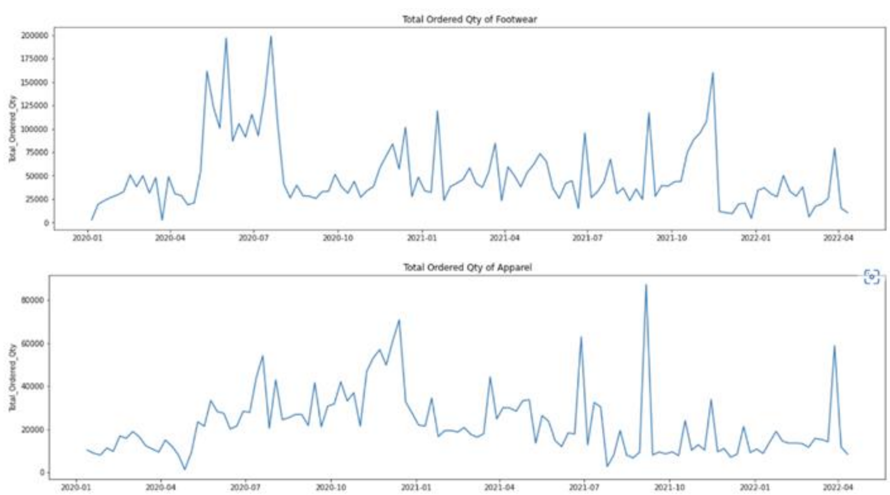
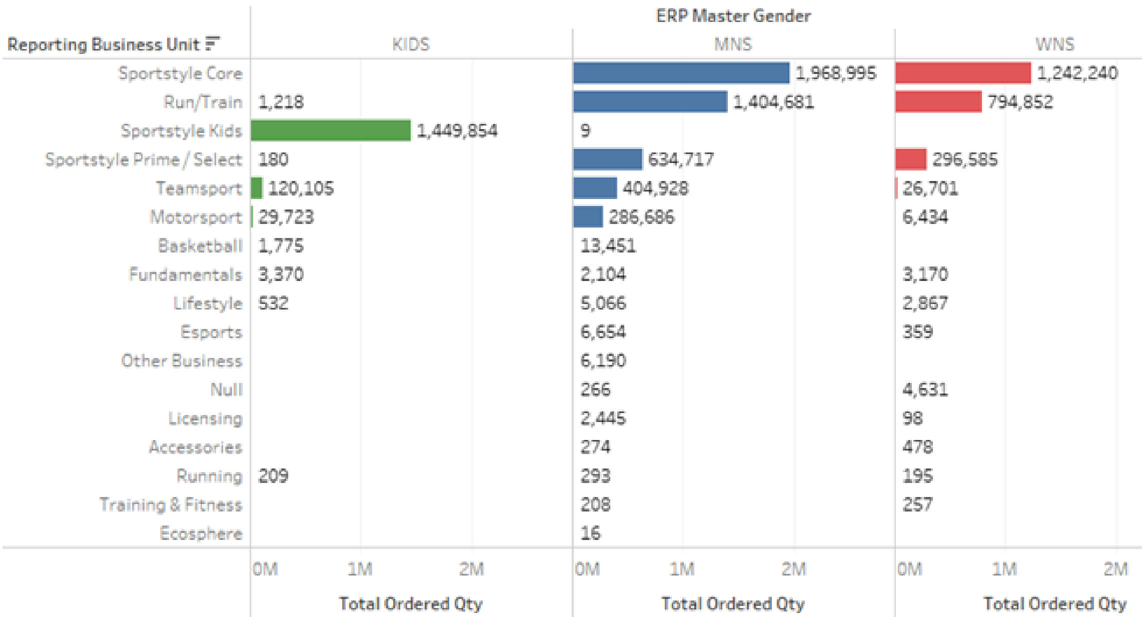


Figure 2: Time Series of Footwear and Apparel



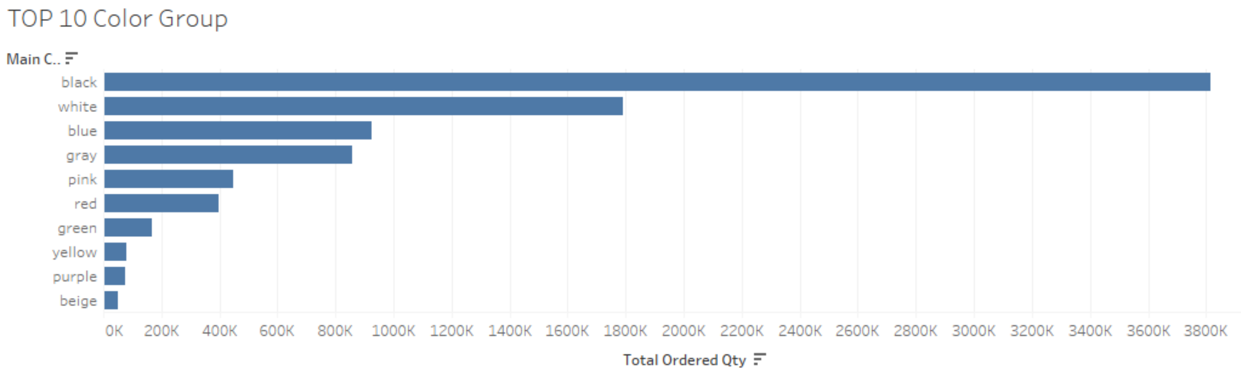
When we divide product division in more detail, we can detect that the main commodity composition of PUMA is footwear. It accounts for more than 50% of total orders and even almost twice as much as the other two parts. The second chart is the time series of footwear and apparel. Obviously the trends of these two charts are not very similar with each other. So if we combine them together to forecast, the prediction model will not perform well. So we will pick footwear as a sample in the prediction part.

Figure 3: Total Ordered Qty by ERP Master Gender



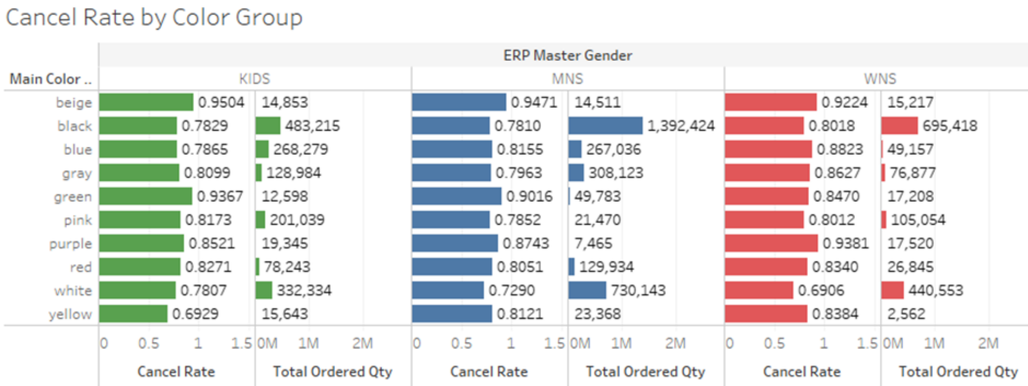
This is the crosstab of all Reporting Business Unit and Master Gender. There are six main business units which make up the most orders of PUMA. Others only accounts for a small part. The classification of these reporting business unit is more like what we called “series”. Some series are designed for specific groups. Like Sportstyle Core is more like a general series for adults because men and women’s order are evenly distributed. Kids have their own series which is called Sportstyle Kids. But like Motorsport rarely has orders from women and kids. It is a unit mostly for man customers.

Figure 4: Top 10 Total Ordered Qty of Color Group



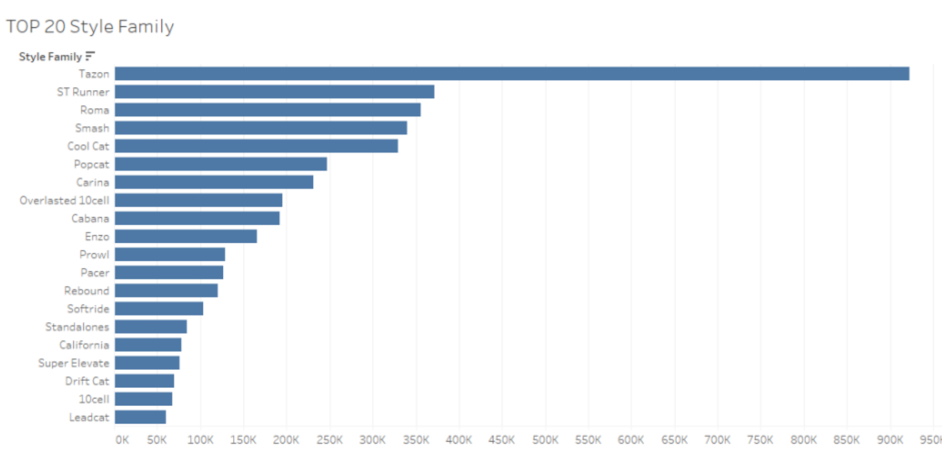
This is about the sales comparison of color groups. What we can observe is that black and white still dominate total sales. Among them, the sales volume of black is also twice that of white, so it can be said that black is the most suitable color for the public customers. The following groups are blue and gray, these two color groups are also very suitable for the public customers. Then there are pink and red, which we think have better sales in certain groups, but they are not suitable for all consumer groups, which is the reason why their sales are not as good as other popular colors. But after dividing sales by gender group (Figure 4), we found some interesting conclusions. As can be seen from Figure4, there is no doubt that black and white are very popular colors in all groups. Sales of blue and gray are high, but relatively speaking, sales of gray among children are somewhat lower than expected. This may be because children prefer bright colors. This is also the main reason why gray is actually lower than blue in total sales. As for red, we were surprised to find that the number of red orders for men was unexpected, while women didn't seem to be interested in red. Other colors belong to minority colors, and there is not a big gap between them.

Figure 5: Cancel Rate by Color Group



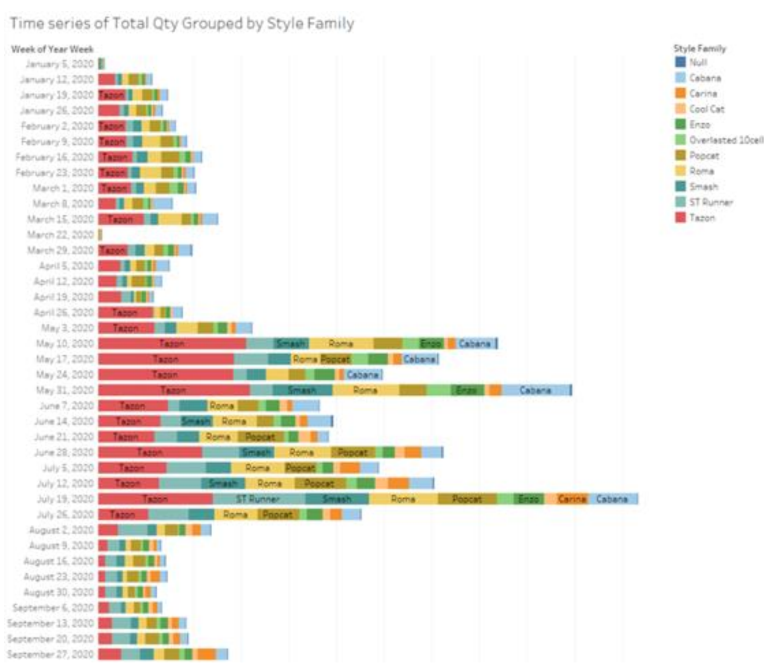
This chart is about the comparison between the cancellation rate and the color sales distribution. For all colors, the cancellation rate is in a relatively high position, which we think is due to the overall lack of supply preparation. There are two colors with a cancellation rate of more than 90% which are green and beige. But PUMA does not need to worry too much, because the total sales of these two colors are relatively low. We need to pay more attention to the color categories that sell more. Among these categories, the overall cancellation rate is about 80%. This is still a very high cancellation rate, so we need to constantly improve our supply. We should ensure the supply of colors with high sales first so it can bring PUMA a higher profit.

Figure 6: TOP 20 Style Family



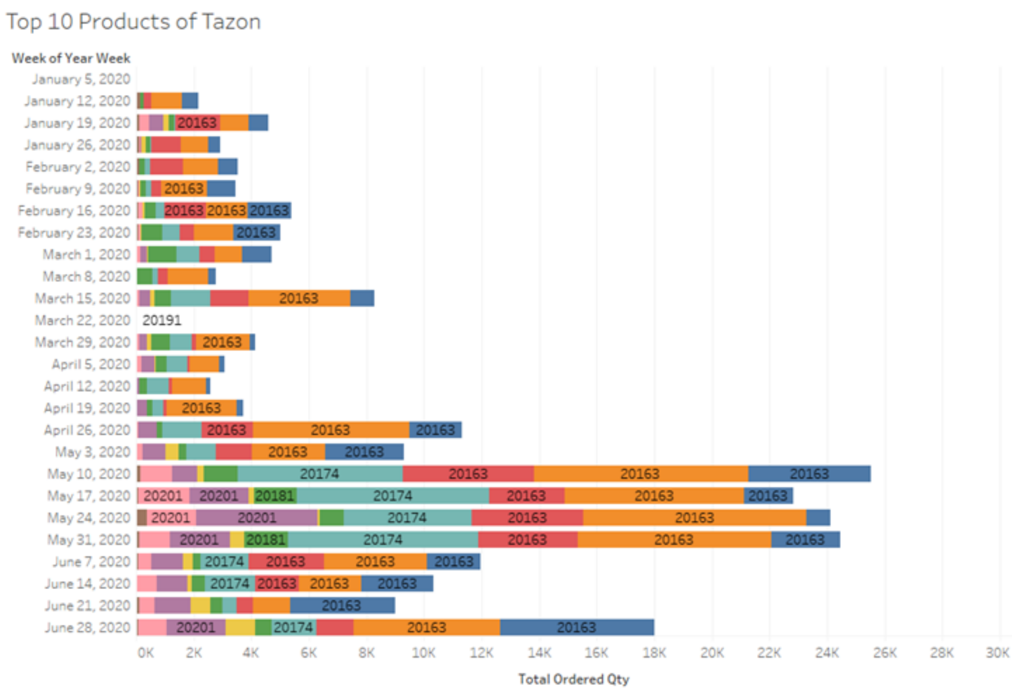
Now let's take a look at the data about style families. Style family is a very valuable attribute for research, because the appropriate style for everyone is different. And customers usually choose shoes that have been proved to be suitable for them to continue to buy. So when a customer thinks a style family is right for him, he will consider buying follow-up products all the time. We selected the shoe type with the highest sales of top20 to study, trying to find some rules of PUMA in the shoe type. Among the many style families, Tazon has the highest sales and is more than double the second style ST Runner. Considering that ST Runner is a style specially designed for children, the advantage of Tazon in shoe style can be regarded to be very obvious. This is a style that is very suitable for most of the customers. ST Runner represents the most popular style among children. So we think that Tazon can represent the consumption trend among consumers to some extent, which is one of the main reasons why we choose Tazon as a sample in the following study. For other styles, the difference between them is not very large. This also shows that different people have different suitable styles, and customers will gradually find their own.

Figure 7: Time Series of Total Qty Grouped by Style Family



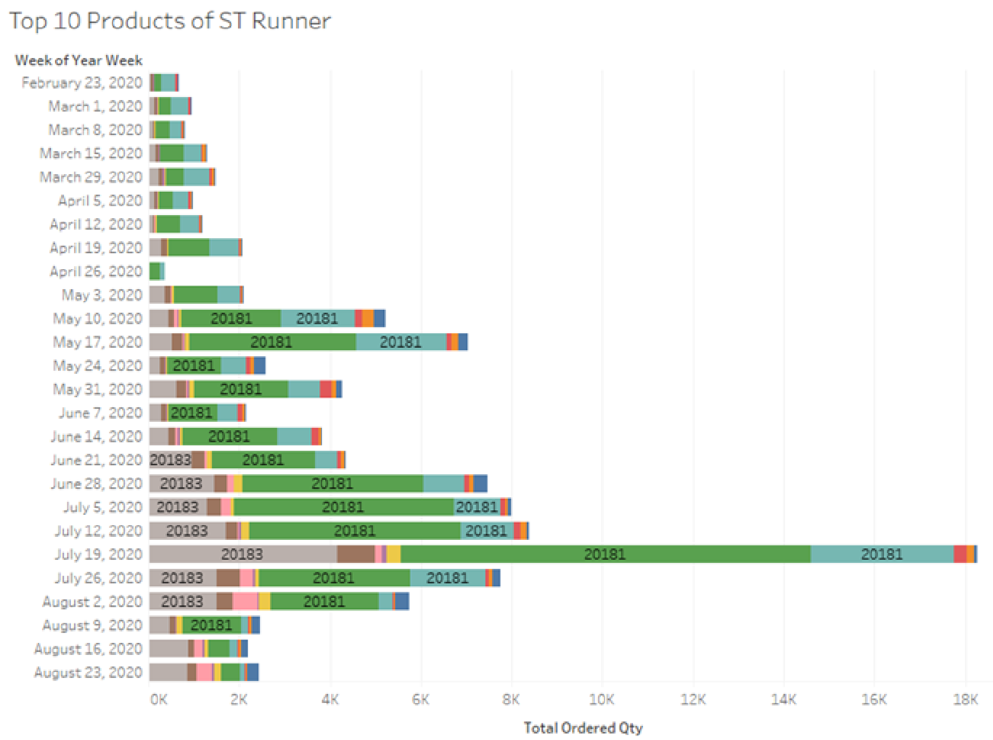
In the times series chart, We noticed that there's a popular period in time series data. So we want to know if it was because of a few popular products that drove sales. So we picked the Top 10 Style Families' data. We can see that as the whole products sales get promoted, Tazon also get in line with this trend. And Tazon accounts for almost 25% of total sales. So we want to know if the appear of this rising period is due to some certain popular products that drive the whole sale.

Figure 8: Top 10 Products of Tazon



We picked the data of Top 10 products especially in Tazon. The number marked on the chart is the first season that this product was introduced to the market. Here on May 10, the most popular items are from 2017 4th quarter, 2016 3rd quarter, and 2016 3rd quarter. And also there's one product from 2020 first quarter. So it is a new product. But most of the others are old products. And we can see the color here, the same color represents the same kind of style family. Those popular items in such period are also popular in the previous period. So there's no evidence that sales were driven by certain new products. Because all the products are sold more in the popular period. They have already proved relatively popular among all the products. They increased evenly. And we cannot know which one contributes the most.

Figure 9: Top 10 Products of ST Runner



Also we picked another style family to test our idea. Here we picked the same date which is May 10. The first product is from the 2018 first quarter. It is also an old item. So the reason for the increasing sales is not some of the certain products, there's maybe other reasons that drive the whole sale and we need to check if PUMA has some market strategy in such period.

Modeling

ARIMA Model

Since footwear sales account for a large portion of total sales, and weekly sales of footwear and apparel and accessories vary widely, we used footwear as an example to forecast sales for the next 40 weeks. First, we aggregate the footwear sales data by date. All categorical variables are invalidated because these categorical variables such as color and gender cannot be aggregated. Therefore, we used the ARIMA model to fit weekly footwear sales data and forecast sales for the next 40 weeks.

Figure 10: Result of Stationarity Test

```
# Stationarity Test
result = adfuller(train)
print('ADF Statistic: %f' % result[0])
print('p-value: %f' % result[1])

# p-value is less than 0.05, the time series in training set is stationary.
```

```
ADF Statistic: -3.855739
p-value: 0.002386
```

We used 80% of the data as training set and the remaining 20% as the test set. We note that the time series data does not have any seasonality. So, we do not need to consider seasonal decomposition. Before modelling with the ARIMA model, the stationarity of the training set time series needs to be tested. The results show that the p-value (0.002386) is less than 0.05, so the data is stationary. Because the time series is already stationary, no difference is needed , that is, the order of differencing (or the parameter d) = 0.

Figure 11: Autocorrelation

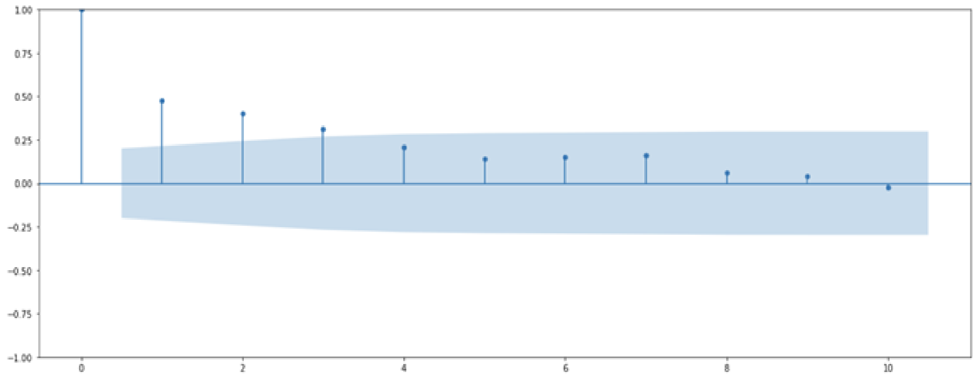
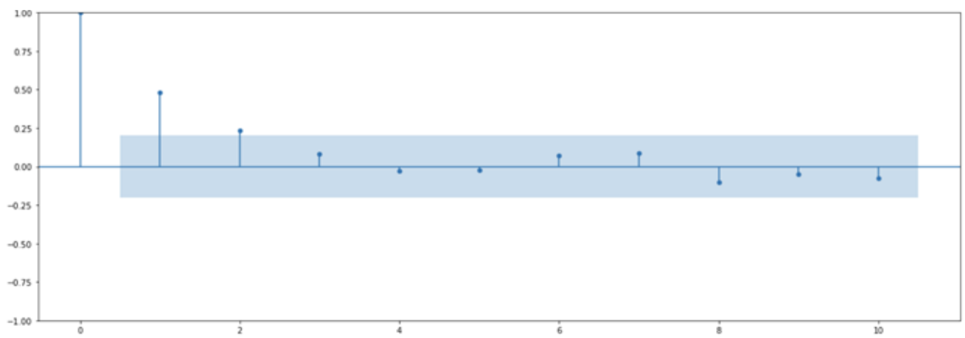


Figure 12: Partial Autocorrelation

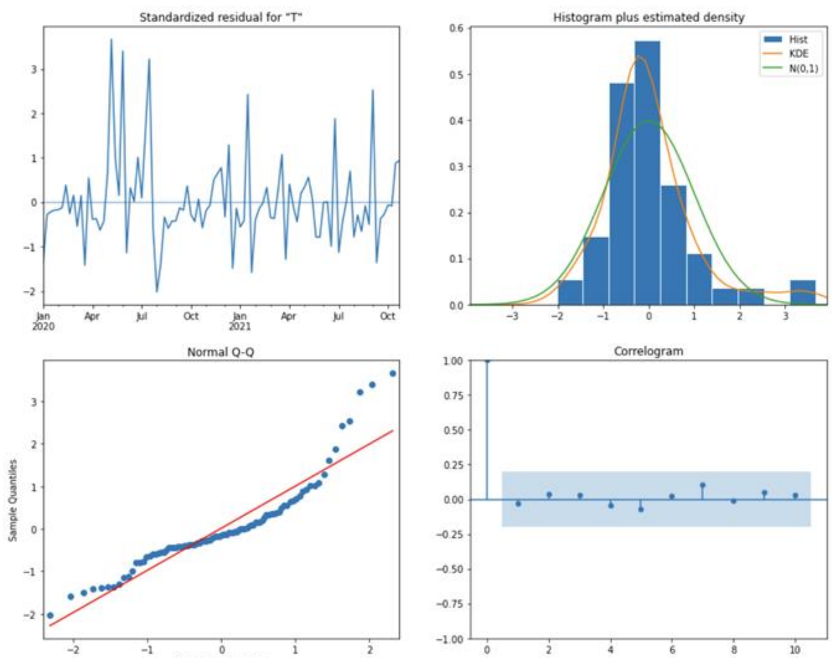


Then, Autocorrelation and partial autocorrelation plots of the smooth time series were plotted and examined to determine the parameters q and p of ARIMA. The results showed trailing autocorrelation (ACF) and insignificant partial autocorrelation (PACF). And the parameter q should be within 3 and parameter p should not exceed 2.

Figure 13: Summary Results of the Optimized ARIMA Model

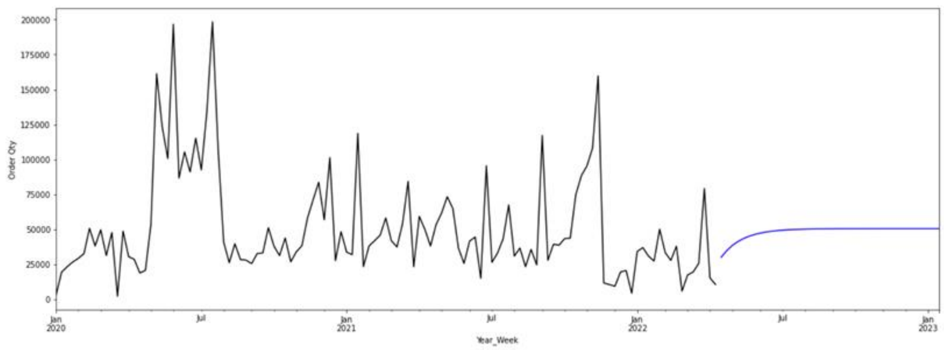
```
==============================================================================
Dep. Variable:       Total_Ordered_Qty   No. Observations:                   95
Model:                  ARIMA(1, 0, 1)   Log Likelihood                -1118.517
Date:                 Sun, 03 Jul 2022   AIC                            2245.034
Time:                         07:56:15   BIC                            2255.249
Sample:                     01-06-2020   HQIC                           2249.161
                          - 10-25-2021
Covariance Type:                   opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       5.396e+04   9833.330      5.488      0.000    3.47e+04    7.32e+04
ar.L1          0.7971      0.085      9.325      0.000       0.630       0.965
ma.L1         -0.4188      0.151     -2.782      0.005      -0.714      -0.124
sigma2      9.997e+08      0.569   1.76e+09      0.000       1e+09       1e+09
===================================================================================
Ljung-Box (L1) (Q):                   0.07   Jarque-Bera (JB):                73.25
Prob(Q):                              0.79   Prob(JB):                         0.00
Heteroskedasticity (H):               0.42   Skew:                             1.43
Prob(H) (two-sided):                  0.02   Kurtosis:                         6.22
===================================================================================
```

Figure 14: Model Diagnostic Results



Next, we used grid search to find the parameter combination with the lowest AIC. where p was set to range from 0 to 2, q was set to range from 0 to 3, and d was equal to 0. The results show that the optimal combination of p, d, q is 1, 0, 1 and we built the ARIMA model. The results of the model show that all terms are significant and therefore the model is valid. Also, the residual term was diagnosed and found to be close to normal distribution, so it was white noise.

Figure 15: Footwear Sales Forecast for the Next 40 Weeks



We used this model to predict sales in the test set and calculated the mean square error (MSE) and the root mean square error (RMSE). The MSE of the prediction results for the test set is 2,182,090,394, and the RMSE is 46,713. Since the predicted values have a large difference with the actual values in the test set, so the ARIMA model does not perform well. Then we predicted the weekly total sales for the next 40 weeks. However, the sales forecast for the next 40 weeks is close to a horizontal line, so the ARIMA still performed poorly. We speculate that this may be because total sales are mixed with a lot of information that might affect sales, so some patterns are hard to recognize by the model.

Note that we did not use other variables, because, as we mentioned, all the categorical variables cannot be aggregated by weeks, so we tried to split the dataset into subsets by Gender, Color, and US Size, and then predicted the sales separately and finally, added them up to get one total model to see if we can improve the prediction performance. After grouping, patterns that were once mixed together may be able to be distinguished and successfully captured and fitted by the model, and ultimately, we can improve the predictive performance of the model.

The variable Gender contains 3 categories ('WNS', 'MNS', and 'KIDS'). The variable US Size contains 11 major sizes from '7' to '12'. The variable Color Group contains the 6 main colors ('black', 'white', 'blue', 'gray', 'pink', and 'red'). First, we divide the sales of footwear into 3 groups according to gender, then split them into training set and test set respectively and use ARIMA model to fit the training set of 3 groups and use grid search to find the optimal parameter combination . Then use the optimal model to predict the weekly sales in the 3 test sets separately and then add up to get the total weekly shoe sales and draw the plot.

Figure 16: Summed Prediction Results for the Test Set Grouped by Gender

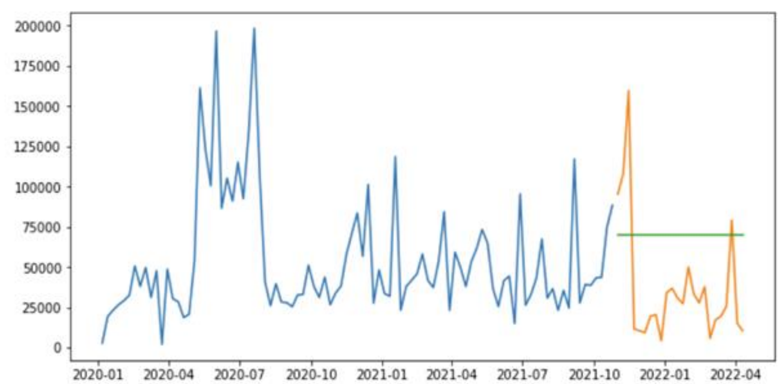


Figure 17: Summed Prediction Results for the Test Set Grouped by US Size

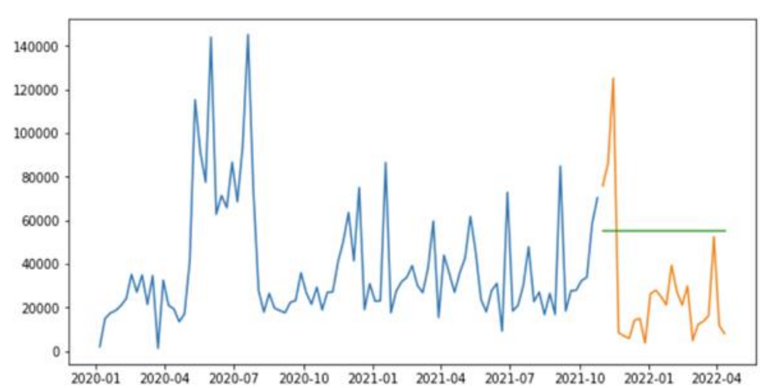
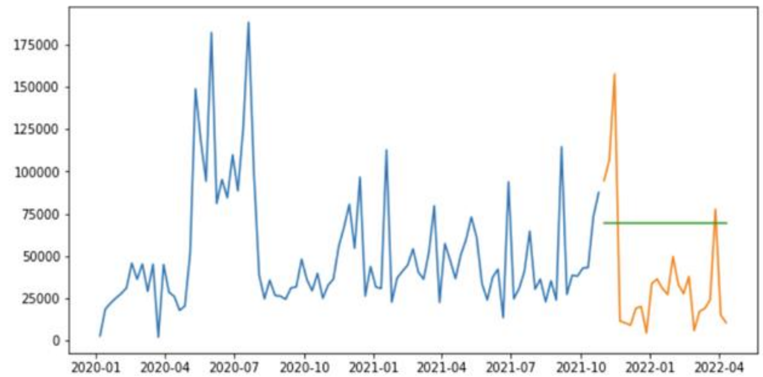


Figure 18: Summed Prediction Results for the Test Set Grouped by Color Group



We also divided the data into 11 groups and 6 groups according to US Size and Color Group, respectively, and then fitted, optimized, predicted, and summed up all the prediction results for each test set. The prediction results of these three grouping methods are all close to a horizontal line, which means that the prediction performance of using these three variables alone is not good.

Figure 19: Summed Prediction Results for the Test Set Grouped by Gender x Color Group

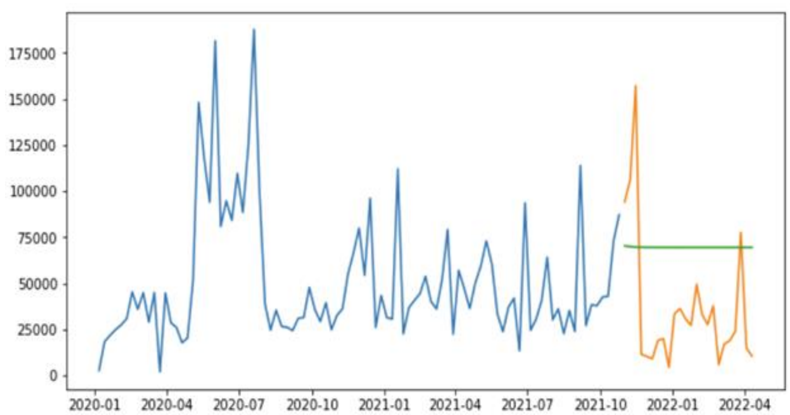


Figure 20: Summed Prediction Results for the Test Set Grouped by Gender x US Size

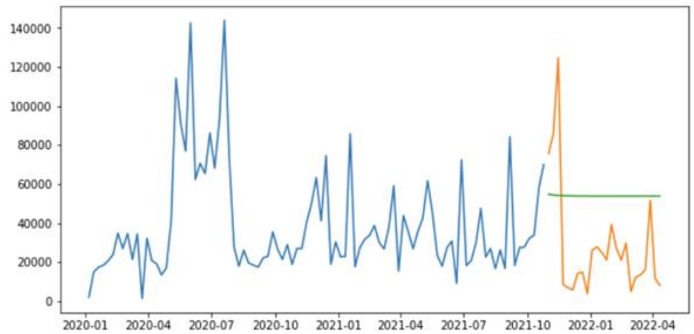
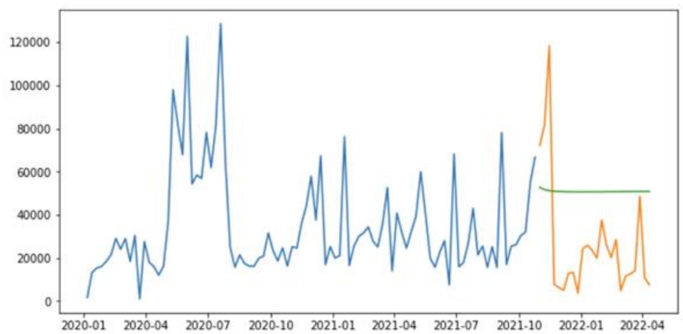


Figure 21: Summed Prediction Results for the Test Set Grouped by US Size x Color Group

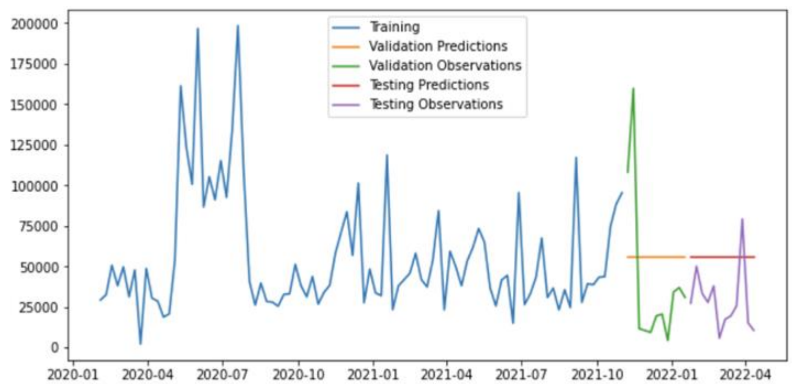


In addition, we also consider combining these three variables in pairs, and then split the data according to Gender x US Size, Gender x Color Group, and US Size x Color Group, and then fit, optimize, predict, and sum the predicted results. The results show that this bivariate cross-classification method still cannot improve the prediction performance of the model, so it may be that the ARIMA model is not suitable for predicting footwear sales data.

LSTM Network

Another model that is often used to predict time series data is the Long Short Term Memory (LSTM) network. First, in order to apply the deep learning model on the time series data, we used the time sliding window method to divide the data into time windows of length 4 weeks. That is, use the sales of the previous 4 weeks to predict the sales of the 5th week and so on. Then we selected the first 80% of the time series as the training set, the middle 10% as the validation set, and the last 10% as the test set.

Figure 22: Prediction Results for Validation Set and Test Set



For the parameters of the LSTM model, I used relu as the activation function. The learning rate of the model is set to 0.01 and the number of iterations of the model is 100. The results show that both the prediction on the validation set and the prediction on the test set are a horizontal line, so the performance of LSTM is also not good.

Conclusion

Through the above analysis, the result of the ARIMA model is not very good, we analyze that there may be many other factors affecting sales. Subsequently, we need to create more variables that may affect sales through feature engineering, and use deep learning models, such as convolutional neural network, LSTM model, etc. ARIMA principle is moving average and autoregressive, so the prediction results are close to the historical average. If the real value fluctuation is not very violent, ARIMA prediction may be more applicable. The neural network LSTM will store the past data in the 'memory nerve'. It's not just looking at an average, so the prediction may be a little radical, but the effect may be better when the original data fluctuates greatly.

On the other hand, though PUMA's products are very competitive in the field of footwear, the overall cancellation rate is very high. In our study of the composition of goods during the period of rising sales, there is no evidence that sales are driven by popular products or new products. There is no seasonality in the time series data, and the ARIMA model may not be suitable for predicting future sales. The prediction performance of the LSTM model is also poor, so there may be many other factors affecting sales.

Recommendations

Puma has a relatively stable position in the sports brand market, so consolidating the traditional market is a very important part. For Puma, footwear products increased the most on different product lines, with a year-on-year increase of 36%; Clothing and accessories products increased by 28.6% and 27.2% year-on-year respectively. First of all, we will do some research on the reasons for the different peaks. There may be the following reasons: Joint brands, some large-scale promotions and social media influence. Since puma officially returned to the basketball market in June 2018, Puma has successfully launched a series of products with both performance and fashion, such as triple and fusion nitro. More and more NBA players wear puma shoes. Star effect has always been an important strategy for Puma to improve its performance.

In addition, in terms of data, we suggest that enterprises introduce more variables that may affect sales or create variables through feature engineering, and try to use other classification variables to predict from the subdivision dimension, such as predicting and reporting the sales of business departments or style series. Although the data in 2021 have recovered and even exceeded the pre epidemic level, the continuing challenges and uncertainties brought by the COVID-19 are still the problems puma needs to solve in 2022.

References

Aditya Shastri, Elaborated Marketing Strategy of Puma + SWOT Analysis, Retrieved from https://iide.co/case-studies/marketing-strategy-of-puma/

Autoregressive Integrated Moving Average (ARIMA). (2021, October 12). Investopedia. https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp

Goyal, S. (2020, November 19). *Credit Card customers*. Kaggle. https://www.kaggle.com/sakshigoyal7/credit-card-customers

Theodore Levitt, Exploit the Product Life Cycle, Retrieved from https://hbr.org/1965/11/exploit-the-product-life-cycle

Tighe, D. (2022, March 1). Puma: Net sales share, by segment worldwide 2021. Statista. Retrieved April 18, 2022, from https://www.statista.com/statistics/254019/share-of-pumas-consolidated-sales-worldwide-by-segment/

Z., Adrians, & Pieter; Zantinge, D. A. (1996, January 1). Data Mining., 15, 2021.