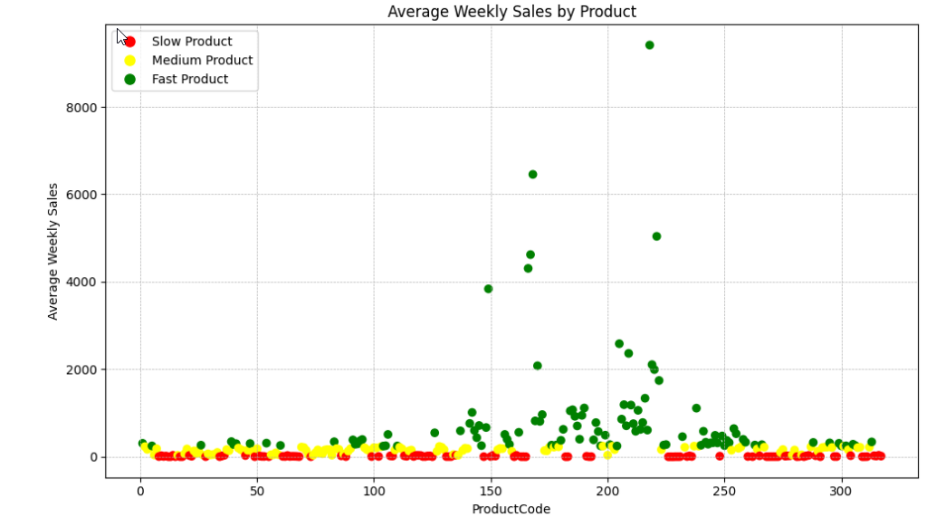
Promotion Bump Assignment:

Part A

1. What are your criteria for separating Fast, Medium and Slow items? Why?

metin, yazılım, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturulduFor this question, a K-means clustering model was implemented. Average weekly sale of an item was taken into account, meaning that this was only feature to be fitted in our model. As can be seen in the below graph (figure 1), yellow dots represents the `Fast Items`, blue ones represent ‘Slow Items’ and lastly, purple dots represent ‘Medium Items’. To compare our findings, the data was divided into thirds (quantile-based discretization) like 33.3% and 66.6%. As can be seen from the graphs (figure 1, figure 2). due to the noise in the data, the medium items was not clustered similarly.

Figure 1

Figure 2

In addition, Fast Items are likely to be more essential and in-demand items (women hygiene products, diapers, toothpaste, etc.).

1. What are your criteria for separating Fast, Medium and Slow Stores? Why?

metin, ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, diyagram içeren bir resim

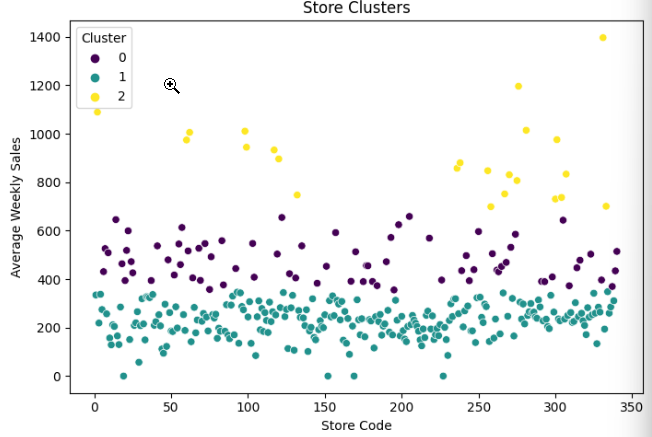
Açıklama otomatik olarak oluşturulduLike the previous question, a K-means clustering model was implemented. Average weekly sale of a store was taken into account, meaning that this was only feature to be fitted in our model. As can be seen in the below graph (figure 3), yellow dots represents the `Fast Stores`, blue ones represent ‘Slow Stores and lastly, purple dots represent ‘Medium Stores. To compare our findings, the data was divided into thirds (quantile-based discretization) like 33.3% and 66.6%. As can be seen from the graphs (figure 3, figure 4). due to the noise in the data, the medium items was not clustered similarly.

Figure 4

Figure 3

1. Which items experienced the biggest sale increase during promotions?

metin, ekran görüntüsü, sayı, numara, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturulduTo answer this question, for each product, average sales is calculated during promotion and non-promotion period and the difference between the sales were compared. The biggest difference in sales during promo and non-promo was for the product number 218.

1. Are there stores that have higher promotion reaction?

metin, ekran görüntüsü, sayı, numara, ekran, görüntüleme içeren bir resim

Açıklama otomatik olarak oluşturulduTo answer this question, for each store, average sales is calculated during promotion and non-promotion period and the difference between the stores were compared. The biggest reaction in for a store during promo and non-promo was for the store number 92.

1. What is the biggest effect explaining sales change during promotions?

We can observe that sales quantities tend to be higher during promotional periods compared to non-promotional periods. However, to decide which is the biggest effect explaining sales change during promotions, we might need to more information, such as the specific store or product.

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. Is there any significant difference between promotion impacts of the Fast versus Slow items?metin, yazı tipi, ekran görüntüsü içeren bir resim

   Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

For the Fast Items difference in sales between promotion dates, and non-promotion dates is significantly higher compared to Slow Items’ difference.

Part B

1. What measure would you use for goodness of fit?
   1. When evaluating the goodness of fit, we need take Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) into account. Coefficient estimates and their p-values also might help. However, for time series models, we should consider the AIC and BIC Key metrics would help a lot about determining goodness of fit.
2. How good is your model developed in step 1?
   1. metin, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, çizgi, ekran görüntüsü içeren bir resim

      Açıklama otomatik olarak oluşturulduWhen we consider Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The MAE indicates that, on average, our forecasts deviate from the actual sales by around 2217 units. he Mean Absolute Percentage Error (MAPE) for our predictions is approximately 69.15%, suggesting that the forecasts deviate from the actual values by this percentage. When we compare with Linear Regression model, SARIMAX gives way more accurate predicts.
   2. Additionally, we can consider p-value for our model’s success. The p-value for the `Is\_Promotion` coefficient is statistically significant (less than 0.05), which suggests that the promotional effect on sales is statistically significant which can be seen below figure.
3. metin, çizgi, yazı tipi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

   Açıklama otomatik olarak oluşturulduWhat are the main problem points causing bad fits?
   1. Our SARIMAX model’s parameters were chosen based on common practices, however they might not be optimal. Adjusting these parameters could improve the fit.
   2. Additional Exogenous Variables: Consider including more exogenous variables that might be fitted into the SARIMAX model to get better test results. For example, categories that influence sales, such as holidays, weekends, or other special events.
4. What would you change in step 1?
   1. I would solve this assignment with other time series model to compare more effectively the success rate of SARIMAX model.