**Question 2**

For this exercise, I narrow down the demand prediction to a specific product ID with the most extensive data i.e. 24 months of data.

I first used the classical ARIMA model to the time-series forecasting. For the deep learning, I experimented the use of DeepAR forecasting algorithm. Unfortunately, given the time constraint, I am unable to run the training model successfully due to the issue with Apple Mac M1 chip that may be incompatible with pytorch-lightning libraries.

Nevertheless, qualitatively, the pros of deep learning models is that it is able to make use of other features to make prediction while the classical forecasting model based its prediction on trends alone. The cons of deep learning model is that it require large dataset due to large amount of network parameters and deep learning model requires extensive parameter tuning to achieve the best results.

To better prepare the input dataset, I would need to study what kind of features that may have greater impact in predicting sku and remove those features that are not and may become a noise during training.