RMSE\_table\_v2 creates a table with which you my use to analyze model performance on quantity.

1. The process begins with importing the baseproduct and basesalesinvoice tables from the combined gold database. Once the tables have been imported, the preprocessing steps begin. The table is converted to a pandas dataframe and restricted to the columns of interest. The datatypes are then changed to allow the ml practitioner to use them for modelling.
2. The dates are restricted to post covid data (after 01-01-2022)
3. A ‘coeffecient of variation’ along with a quantity is added to each product or sku. These divided into three classes and further combined to give a score which indicates the combined variation/volume (AX,AY,AY,AZ,BX,BY,BZ,CX,CY,CZ) to each product or sku.
4. Each column of interest is then formatted into a pivot table where the index is the sku the columns are the dates and the values of the table are the values of the column of interest.
5. Next, some basic formatting is done on columns for datatype, duplicates are dropped and the tables are prepared (via the transpose operation) for the four models.
6. Once the datasets are preprocessed, they can then serve as inputs for the models: ARIMA, SARIMA, Exponential Smoothing and Prophet.
7. Note these are sequential time series forecasting models and do not follow the traditional train/test.
   1. The inputs if the models are sequences (or lists) on values and using the trends, seasonality and residues of these sequences a value is forecasted.
   2. Each model has a window (or sequence) size and a forecast horizon showing how far ahead you want to predict
   3. Additionally, each model has the capability (when data is available) to begin its forecast from the first available data point. Note that if this functionality is implemented you should change the training data
   4. Once the model finishes a table is produced showing the forecast and actuals along with columns for revenue which are used to calculate an impact score
   5. The impact score is then calculated and added to the model
8. Next an rmse table is calculated for each model. This is the metric we can use to test the models accuracy. Additionally, a ‘goodness of fit’ (R^2-score) is added to each model to illustrate which model tends to fit to the given data better.
9. Once all models and rmse tables are completed, a column for the champion model, the fitscore, and the previous combined classes are added in.
10. Finally the table is output into the combinedgold database.