# Project: Forecasting Sales

## Step 1: Plan Your Analysis

1. Does the dataset meet the criteria of a time series dataset? Make sure to explore all four key characteristics of a time series data.

   Ans. The dataset has each time unit with one data point or more, it is dependent on time, the time interval is continuous, sequential and of each interval.

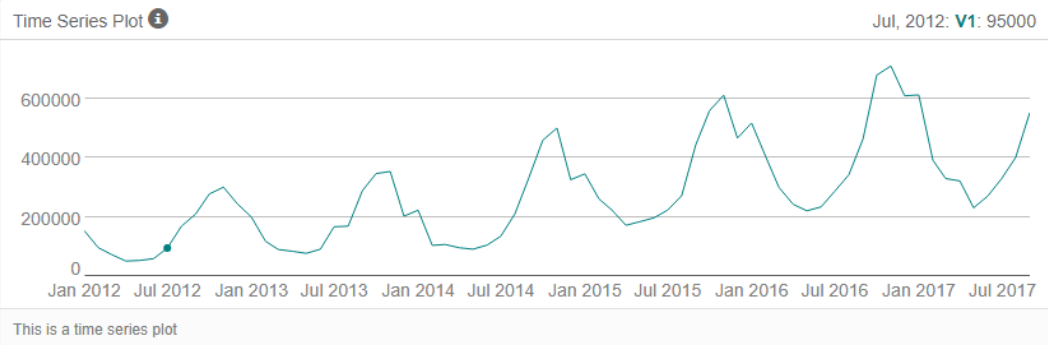
2. Which records should be used as the holdout sample?

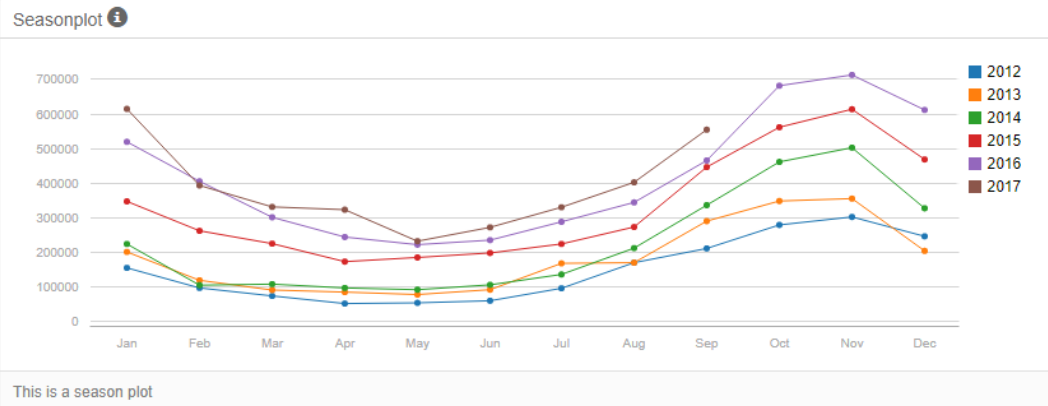
   Ans. The required sales of 4 months are to be predicted so holdout sample of 4 month should be used from June 13 to September 13.

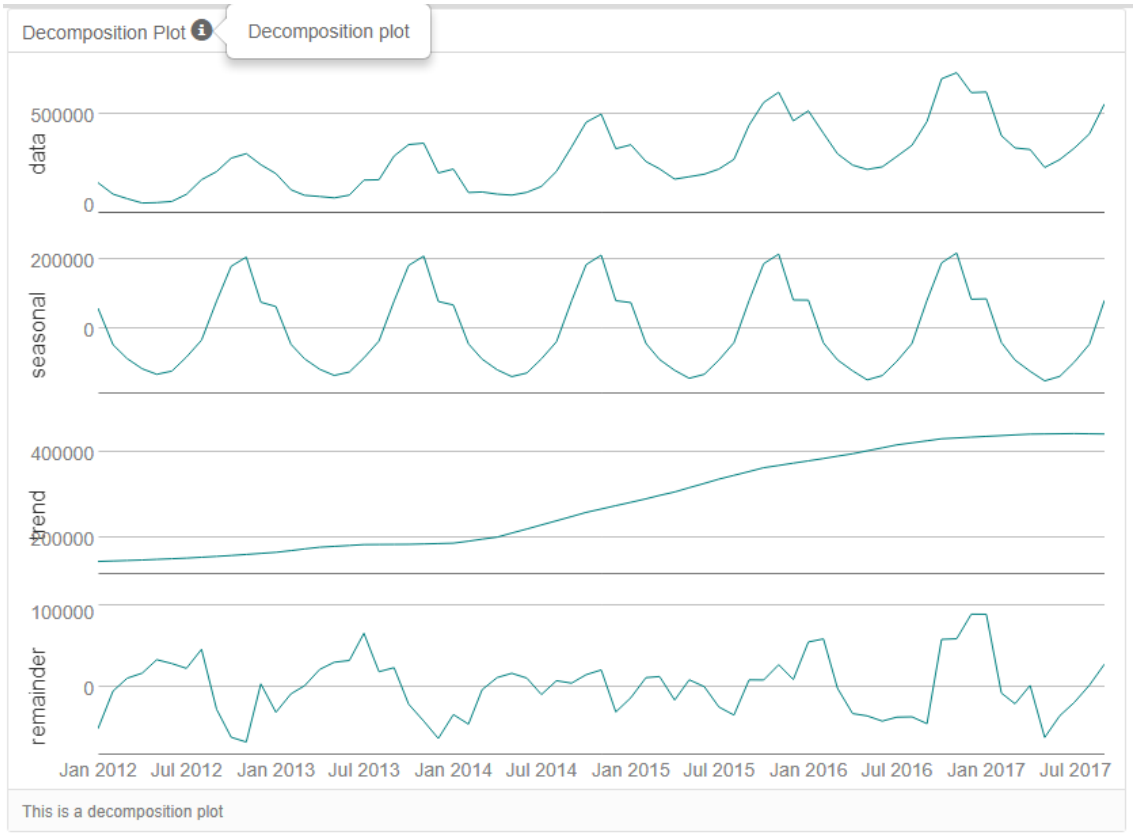
## Step 2: Determine Trend, Seasonal, and Error components

*1.* What are the trend, seasonality, and error of the time series? Show how you were able to determine the components using time series plots. Include the graphs.

Ans. Trend and seasonality show an increasing while only fluctuation in error (remainder). Should apply multiplication and addition, respectively, on trend and seasonality and multiplication on error.





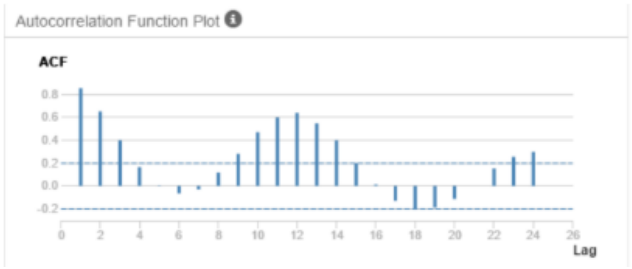

## Step 3: Build your Models

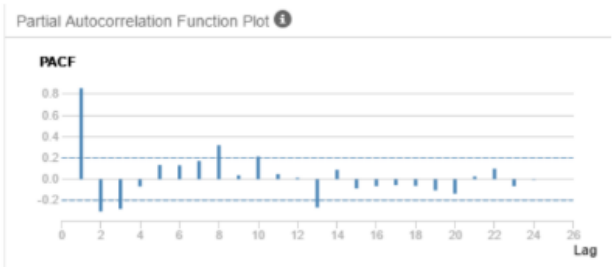
1. What are the model terms for ETS? Explain why you chose those terms.

   Ans. The models terms for ETS(M,A,M) were selected from the time decomposition plots. Both dampened and non-damped models were compared. The dampened model had an AIC value of 1639.47, RMSE of 33153.53 and MASE of 0.3675. The non-damped model AIC value of 1639.74, RMSE of 32992.73 and MASE of 0.3727. The comparison results in favor of dampened model as it has higher accuracy, due to lower RMSE and MASE.

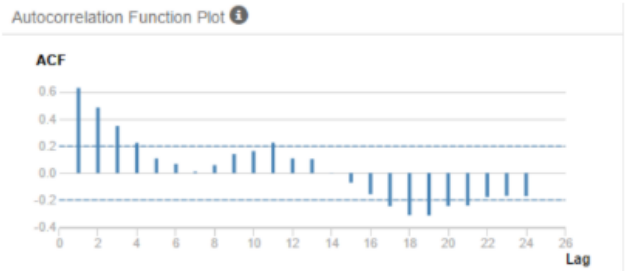
2. What are the model terms for ARIMA? Explain why you chose those terms. Graph the Auto-Correlation Function (ACF) and Partial Autocorrelation Function Plots (PACF) for the time series and seasonal component and use these graphs to justify choosing your model terms.

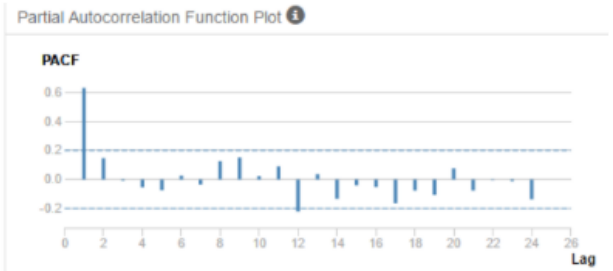
   Ans.
   Without differencing.



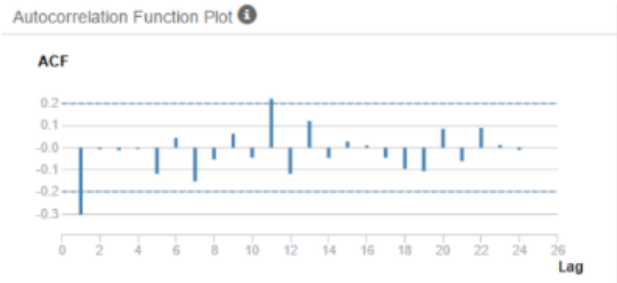



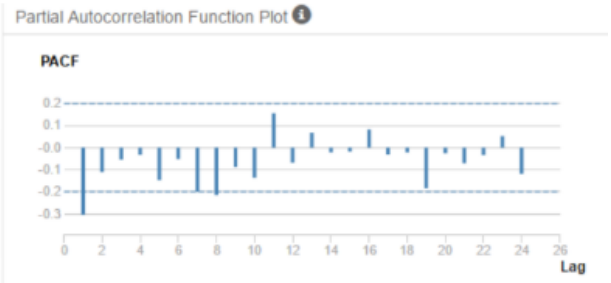

With seasonal differencing




With seasonal first differencing.




The value of AIC is 1256.60, RMSE is 36761.53 and MASE is 0.3646. Both plots do not show correlation.

## Step 4: Forecast

1. Which model did you choose? Justify your answer by showing: in-sample error measurements and forecast error measurements against the holdout sample.

   Ans. The ARIMA model computed lower values of MAPE and ME than ETS model. This is why ARIMA model is better at forecasting sales for the holdout sample.

ETS model

Actual and Forecast Values:

| Actual | ETS_damped |
|---|---|
| 271000 | 255966.17855 |
| 329000 | 350001.90227 |
| 401000 | 456886.11249 |
| 553000 | 656414.09775 |

Accuracy Measures:

| Model | ME | RMSE | MAE | MPE | MAPE | MASE | NA |
|---|---|---|---|---|---|---|---|
| ETS_damped | -41317.07 | 60176.47 | 48833.98 | -8.3683 | 11.1421 | 0.8116 | NA |

ARIMA model

Actual and Forecast Values:

| Actual | ARIMA |
|---|---|
| 271000 | 263228.48013 |
| 329000 | 316228.48013 |
| 401000 | 372228.48013 |
| 553000 | 493228.48013 |

Accuracy Measures:

| Model | ME | RMSE | MAE | MPE | MAPE | MASE | NA |
|---|---|---|---|---|---|---|---|
| ARIMA | 27271.52 | 33999.79 | 27271.52 | 6.1833 | 6.1833 | 0.4532 | NA |

2. What is the forecast for the next four periods? Graph the results using 95% and 80% confidence intervals.

   Ans. The forecasting from 13 Oct to 13 Jan are 754,854, 785,854, 684,654 and 687,854.