

Claim Forecasting

Instructions:

- You have to use Python for the complete analysis. Use Jupyter notebook for documentation and python code.
 - In case of any issue with the code, please mail to rahul@awesomestats.com
-

Forecasting the demand of services or products leads to better management of short term or long term planning. In this case, we are looking at the warranty related issues reported, on a particular brand of two-wheeler. The data is a monthly roll-up of approximately half a million issues reported by the customers over a four year period.

We will be using Claim forecasting data in this exercise. Refer the **Exhibit 1** to understand the feature list. Use the data and answer the below questions.

1. Load the time series dataset in Jupyter Notebook using pandas.
2. Plot the time series data to visualize trend and seasonality in the data.
3. Decompose the claim data to report Trend, Seasonality, and irregular component. Find the Seasonality window using the claim data.
4. Test stationarity of data using augmented dickey fuller test.
5. Use data differencing as a strategy to make the data stationary.
6. Plot the ACF and PACF plot. How will you inspect the plot to arrive at the p-lags and q-lags?
7. Split the data into training set and test set. Use walk forward validation strategy for model building and evaluation.
8. Given recent claim, what is the expected claim for the next time period? Build a model with statsmodel.api to forecast the amount claimed in next time step.
9. How do you interpret the model outcome? Report the model performance on the walk forward validation set.

Exhibit 1

Sl. No.	Name of Variable	Variable Description
1	date	Date of Claim
2	rate	Amount claimed
3	item	Number of claims