

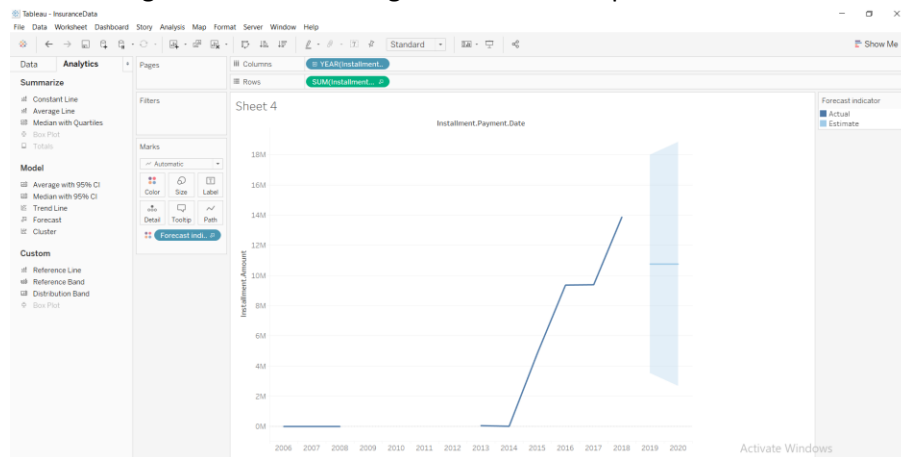
Forecasts and Clustering in Tableau

Marium Jamal – 14881

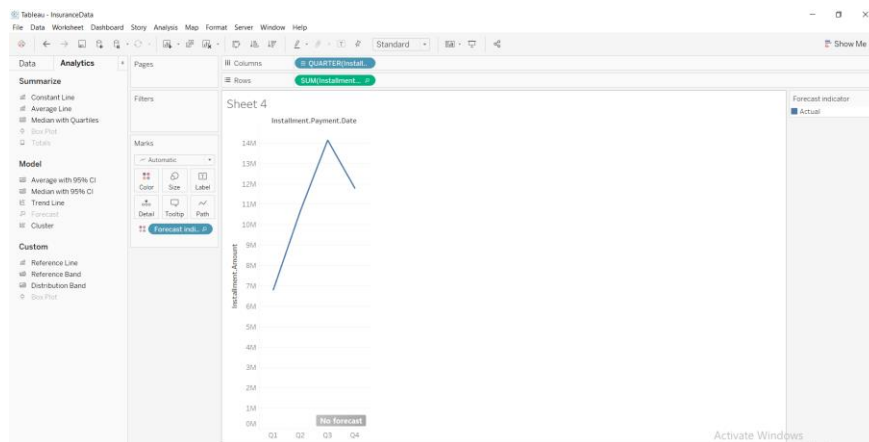
Task 1:

In this task, we are trying to forecast the installment amount w.r.t installment payment date. The steps involved to achieve the forecast are as follows:

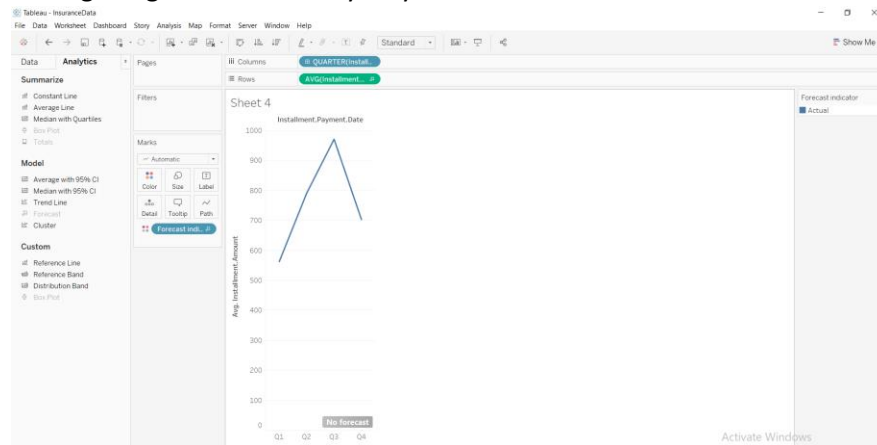
1. Imported excel file of InsuranceData into tableau.
2. Dragged and dropped the relevant columns i.e. Installment amount and Installment Payment Date into rows and columns, respectively. This is because time series column is always placed on the x-axis, and the KPI(the attribute that need to be forecasted) is placed on y-axis.
3. Selected the forecast model from the analytics field given on the left side of the page. This simply gives us a straight line of forecasting which is not acceptable.



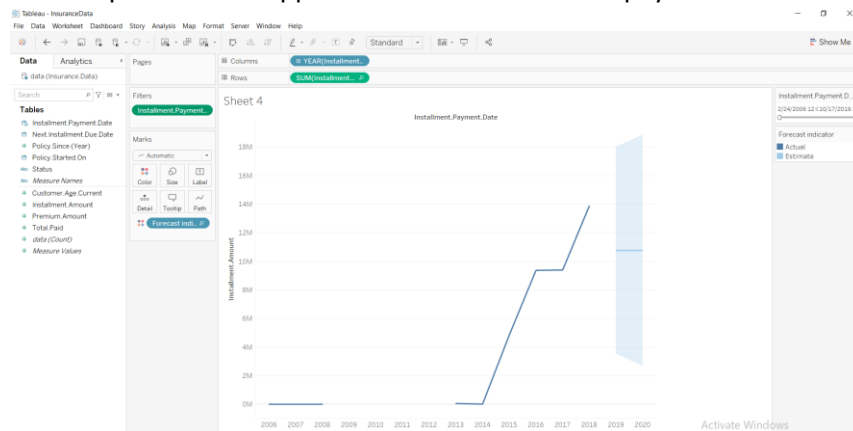
4. Drilled down the installment payment date into quarters, months and even weeks but all of them resulted in no forecast.



5. Changed the measure of installment amount from sum to average for both quarterly and yearly data, but did not get a good forecast anyway.



6. Returned back to step 3 and then applied a filter on installment payment date.



7. Made possible ammendments in the forecast options such that: confidence interval = 90% and Forecast length = exactly 3 years

Forecast Options

Forecast Length

☐ Automatic Next 8 quarters

☒ Exactly 3 Years

☐ Until 1 Years

Source Data

Aggregate by: Automatic (Quarters)

Ignore last: 1 Quarters

☐ Fill in missing values with zeroes

Forecast Model

Automatic

Automatically selects an exponential smoothing model for data that may have a trend and may have a seasonal pattern.

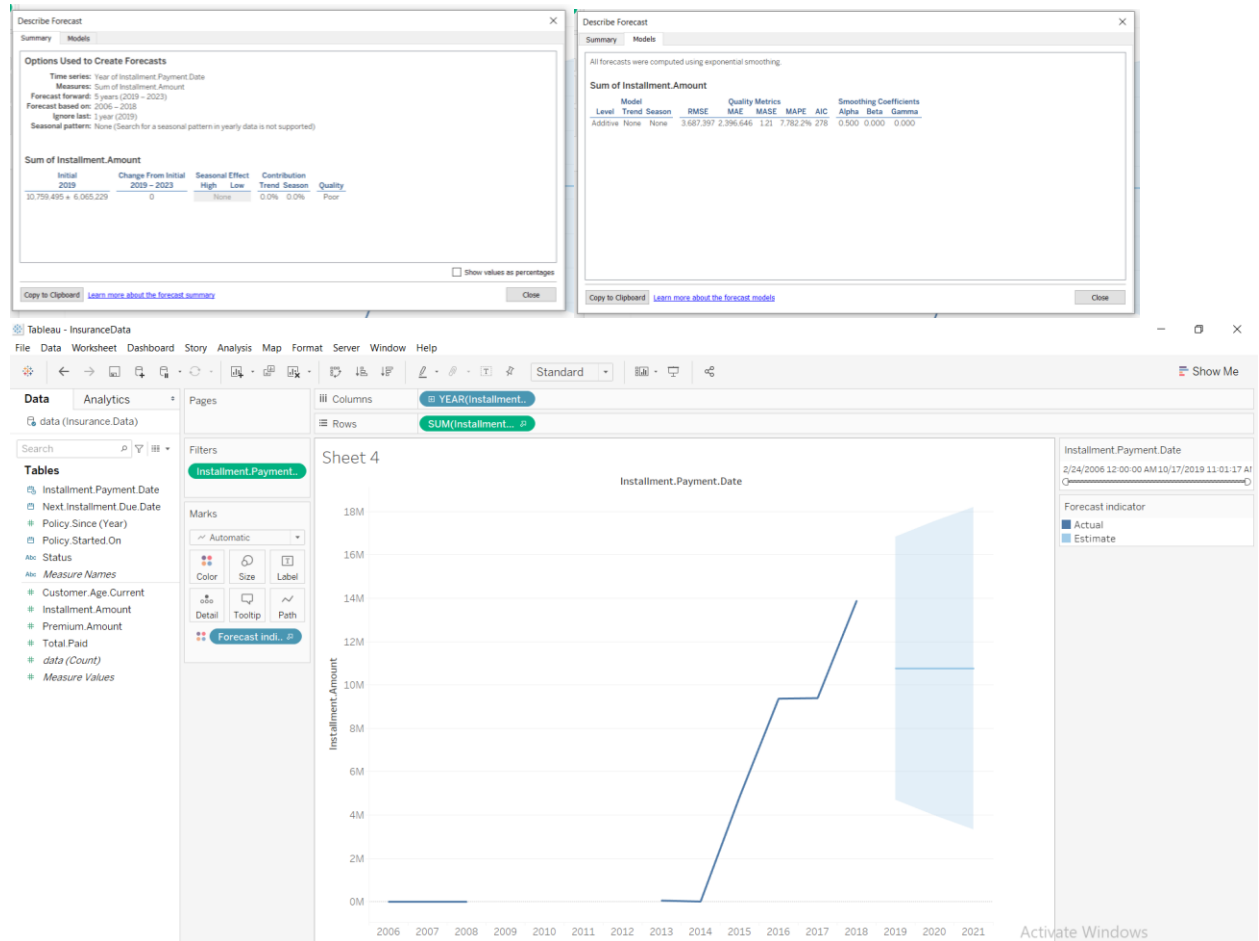
☒ Show prediction intervals 90%

Currently using source data from 2013 Q4 to 2019 Q3 to create a forecast through 2022 Q3. Looking for potential seasonal patterns every 4 Quarters.

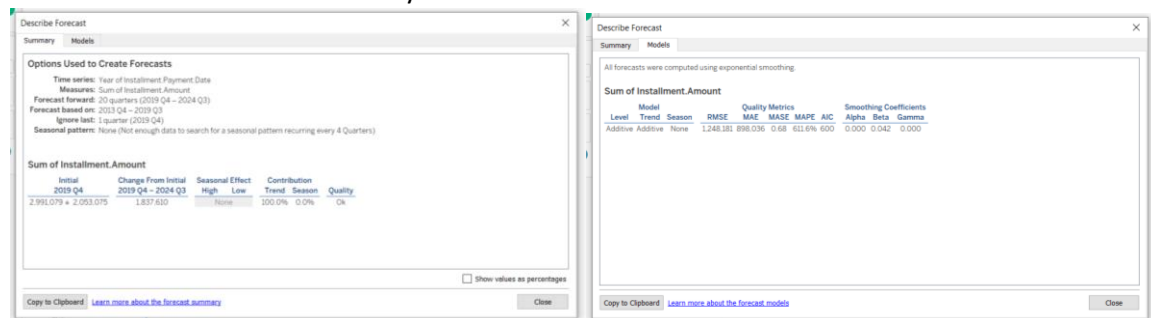
[Learn more about forecast options](#)

OK

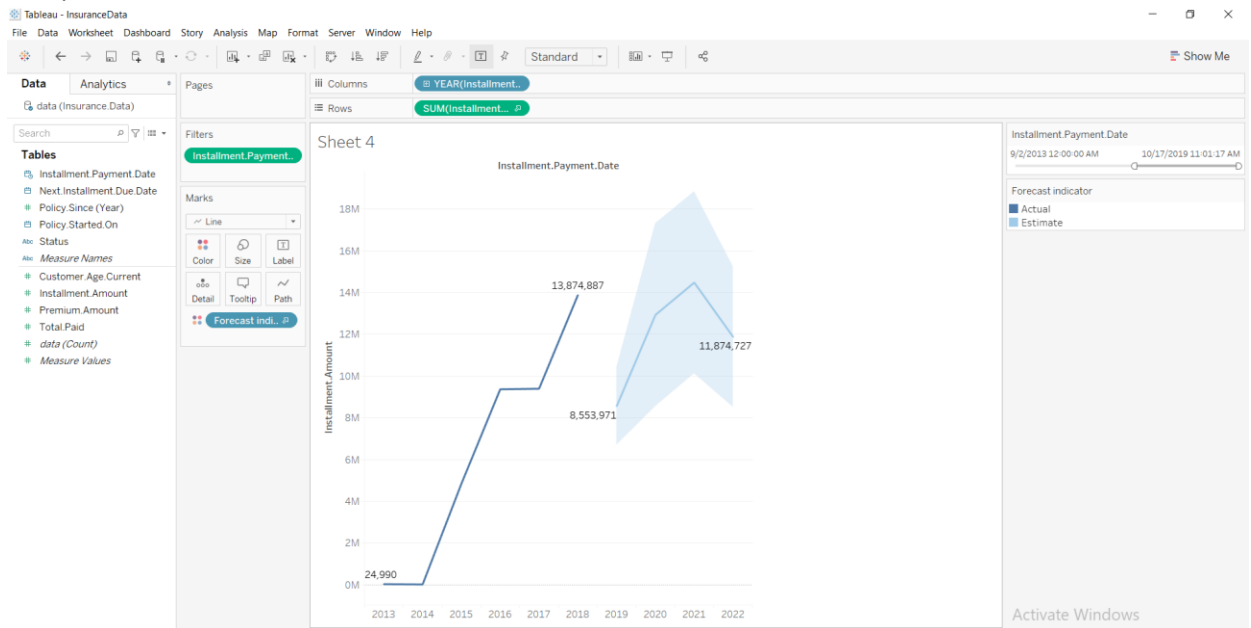
8. Modified the confidence interval and forecast length but it still returned a poor quality forecast with a very high RMSE i.e. 3,687,397.



9. Varied the installment payment date from the range filter on the right, and achieved an acceptable forecasting at 2013-2019.
10. Viewed tableau forecast summary in order to have an idea about our created forecast.

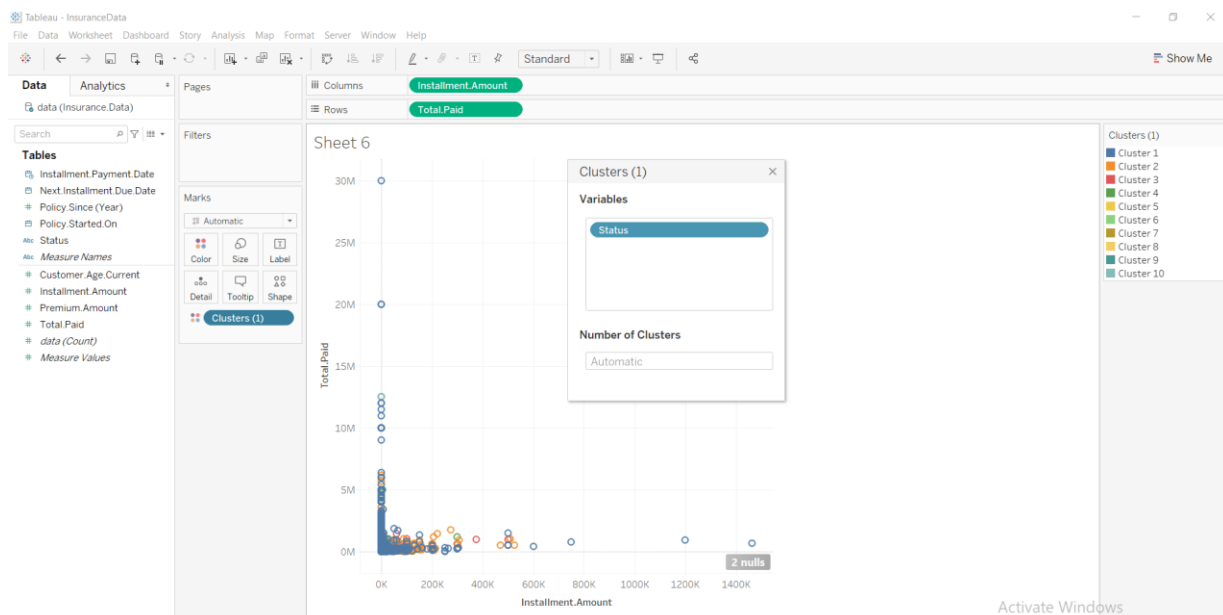


11. According to the tableau summary, forecast quality is OK, RMSE is 1,248,181 and AIC is 600. Both RMSE and AIC are low in this case as compared to other cases and Additive trend is also observed in this case; hence, we are considering this forecast as the most accurate and acceptable.



Task 2:

In order to identify KPIs for the given data, I plotted graphs for multiple dimensions such as installment amount vs. total paid where status was chosen as cluster dimension, but it did not return any proper trend for clusters as shown below. Hence, “No. of active customers” being one of the important KPIs generally cannot be considered as important KPI in this case.



The only acceptable trend of clusters was observed in the graph of avg. Installment amount vs avg. premium amount with both of these dimensions as the cluster variables as well. The graph below shows that for installment amount of or near to 0K has premium amount in different ranges such as 0M-1.2M, 1.3M-6.4M, 9M-12.5M and 20M-30M forming total of 4 clusters(cluster3, cluster4, cluster5, cluster6). Whereas, for installment amount 143K-750K, premium amount of less than 1M was observed forming another cluster, cluster2. Lastly, for installment amount in the range 1200K-1473K, premium amount was approx. 0M forming another cluster labelled as cluster1.

Hence, it can be said that the two KPIs for insurance data are Installment Amount and Premium Amount because the two of them return acceptable clusters of the data.

