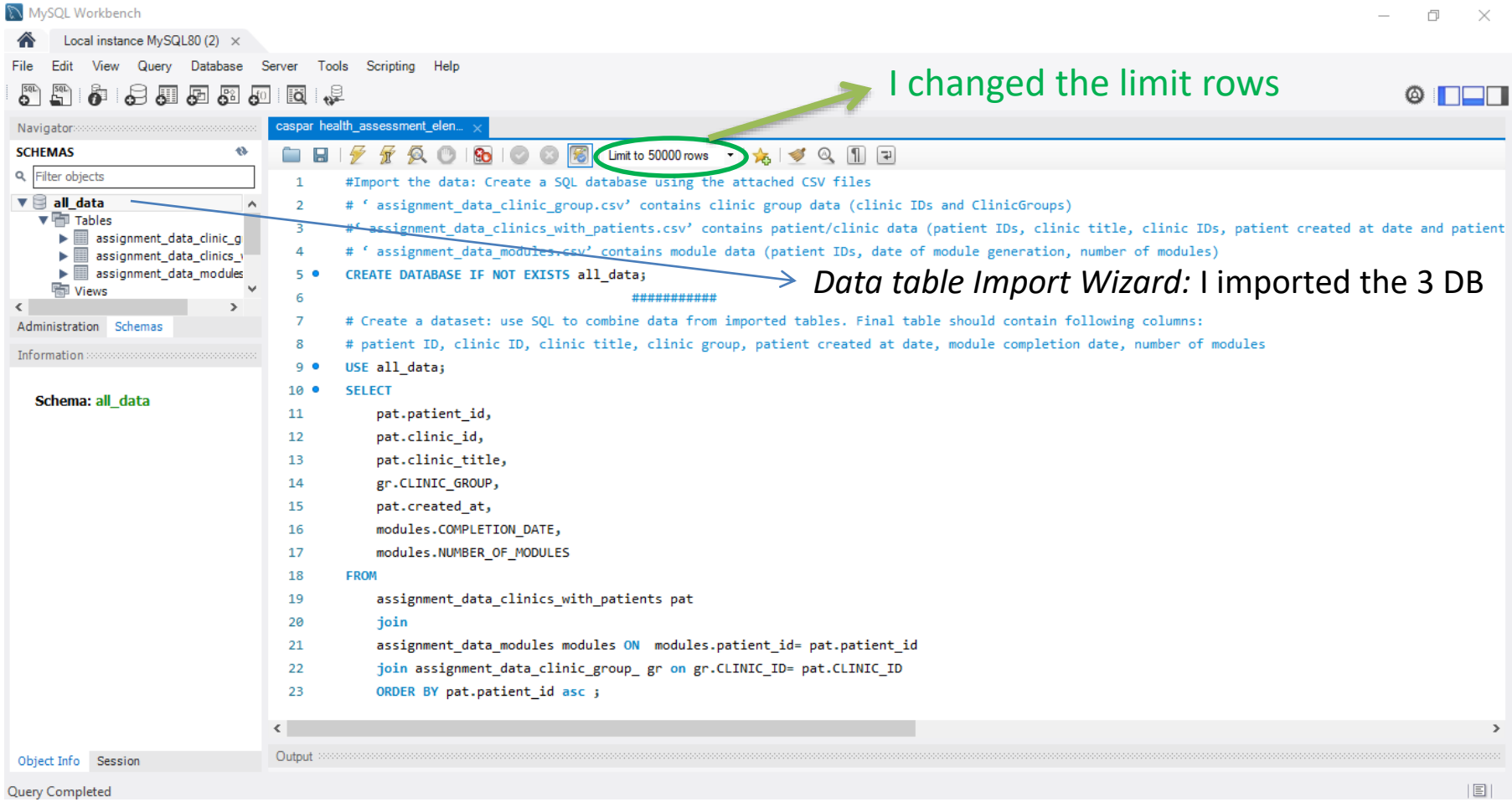
Elena Peña

Caspar Health

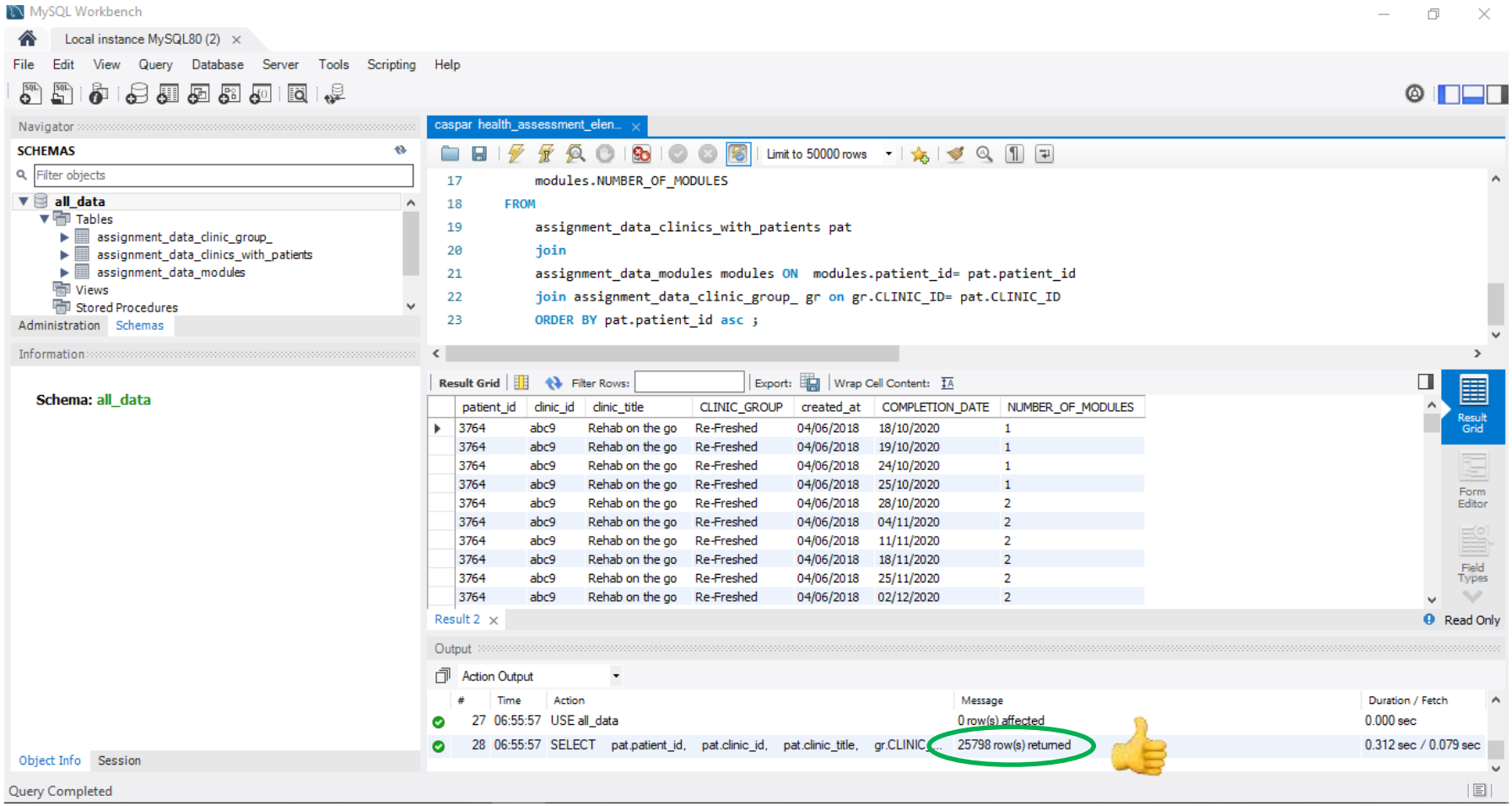
SQL and Tableau Assignment

SQL(1): Code

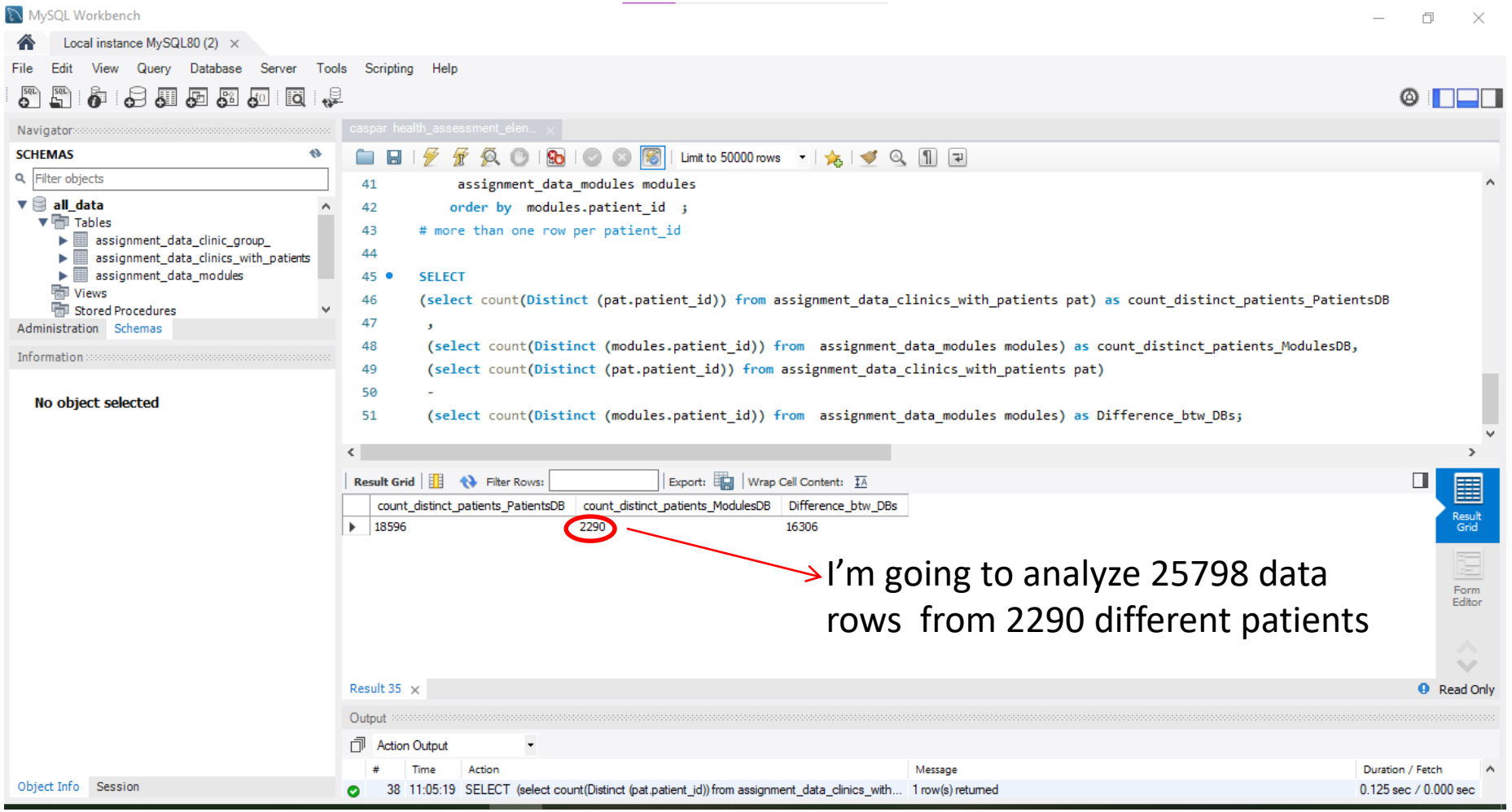
Difficulties: *Import wizard* couldn't read well the csv-files. There were problems with the encoding, so I went back to Excel reader and inside of "Save as": "Excel options" I changed the file type and "save files in web service". It did work!!



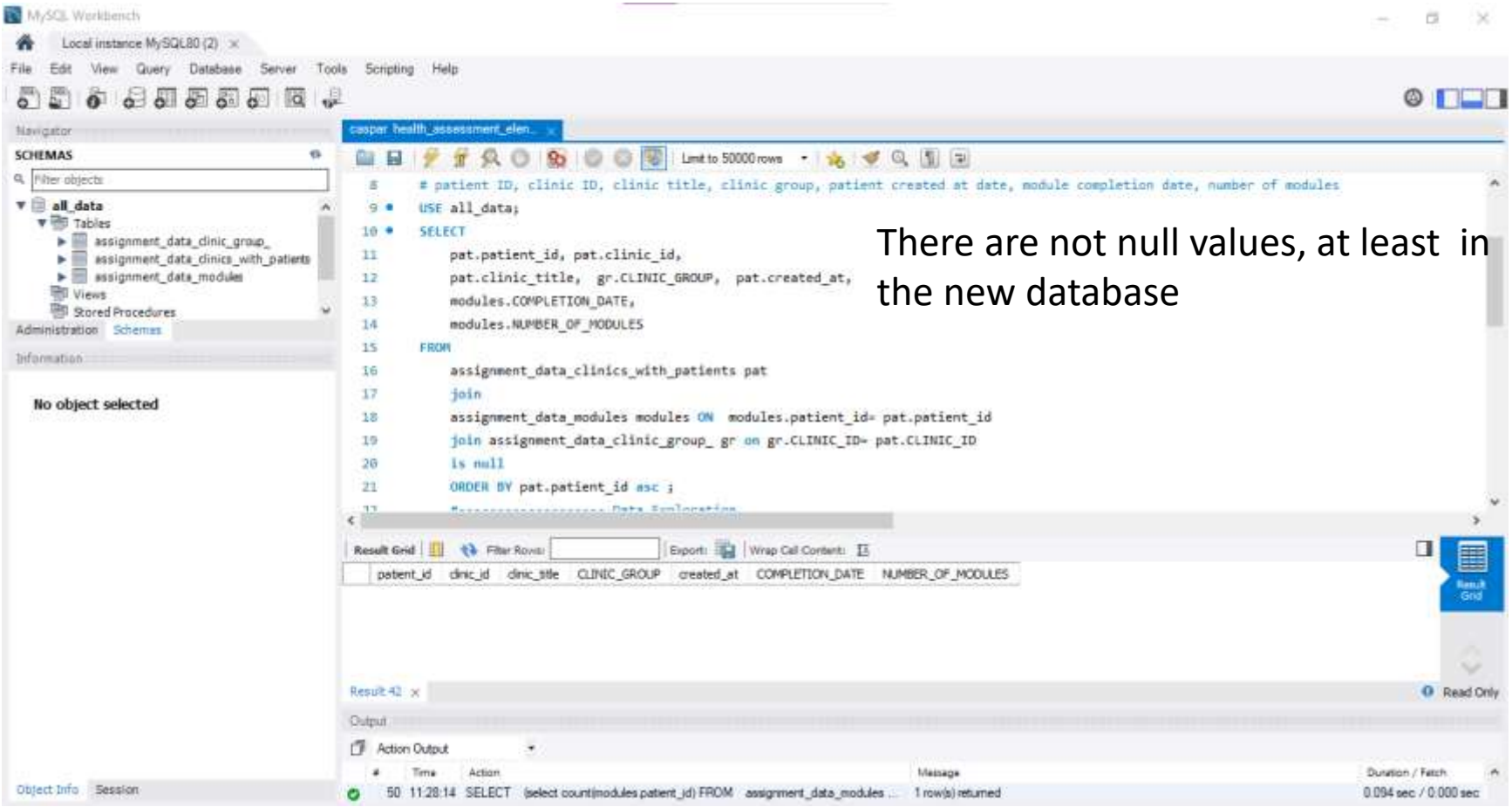
SQL(2): results



SQL(3): Data Exploration



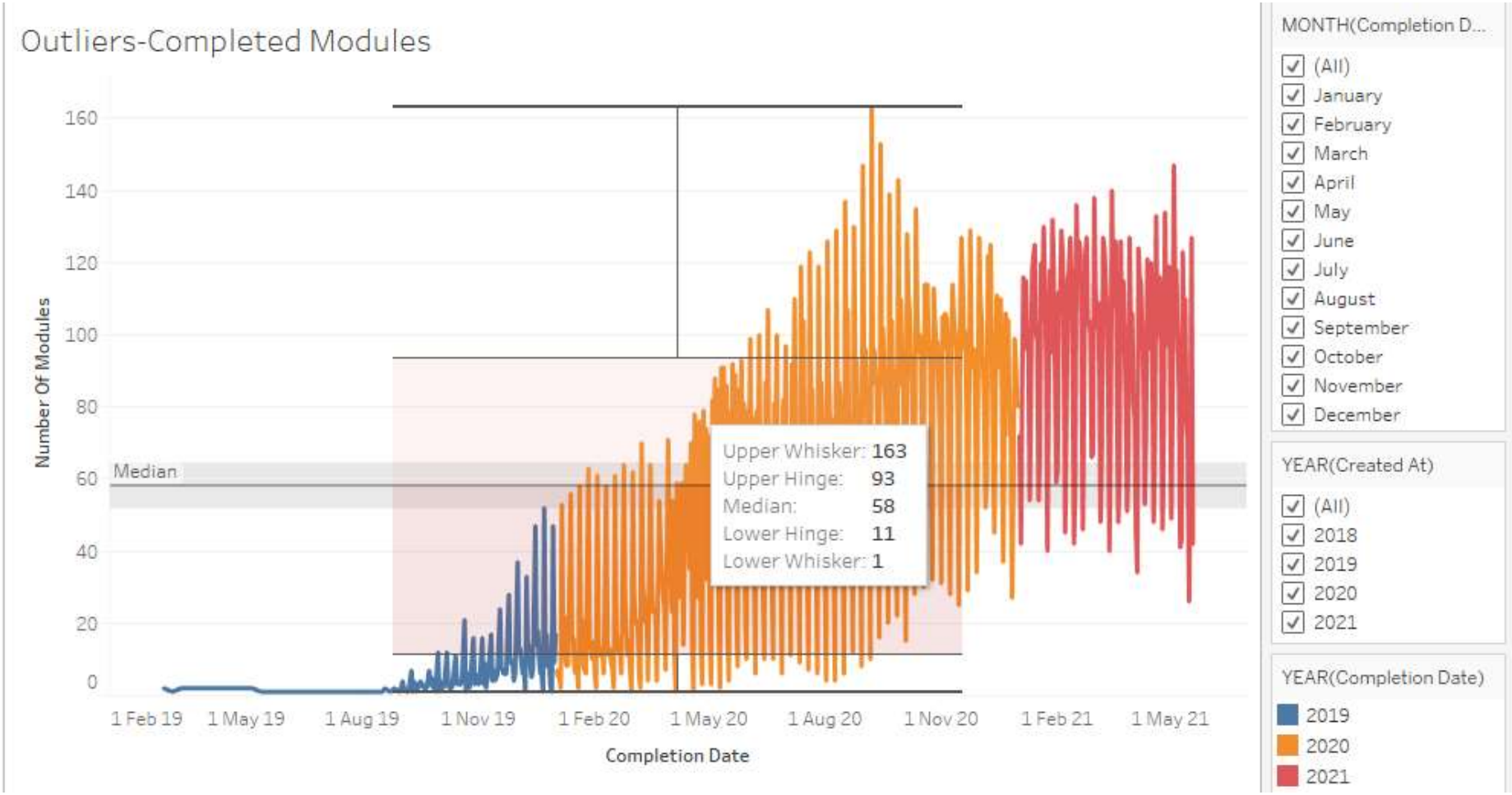
SQL(5): Data Exploration
Missing values



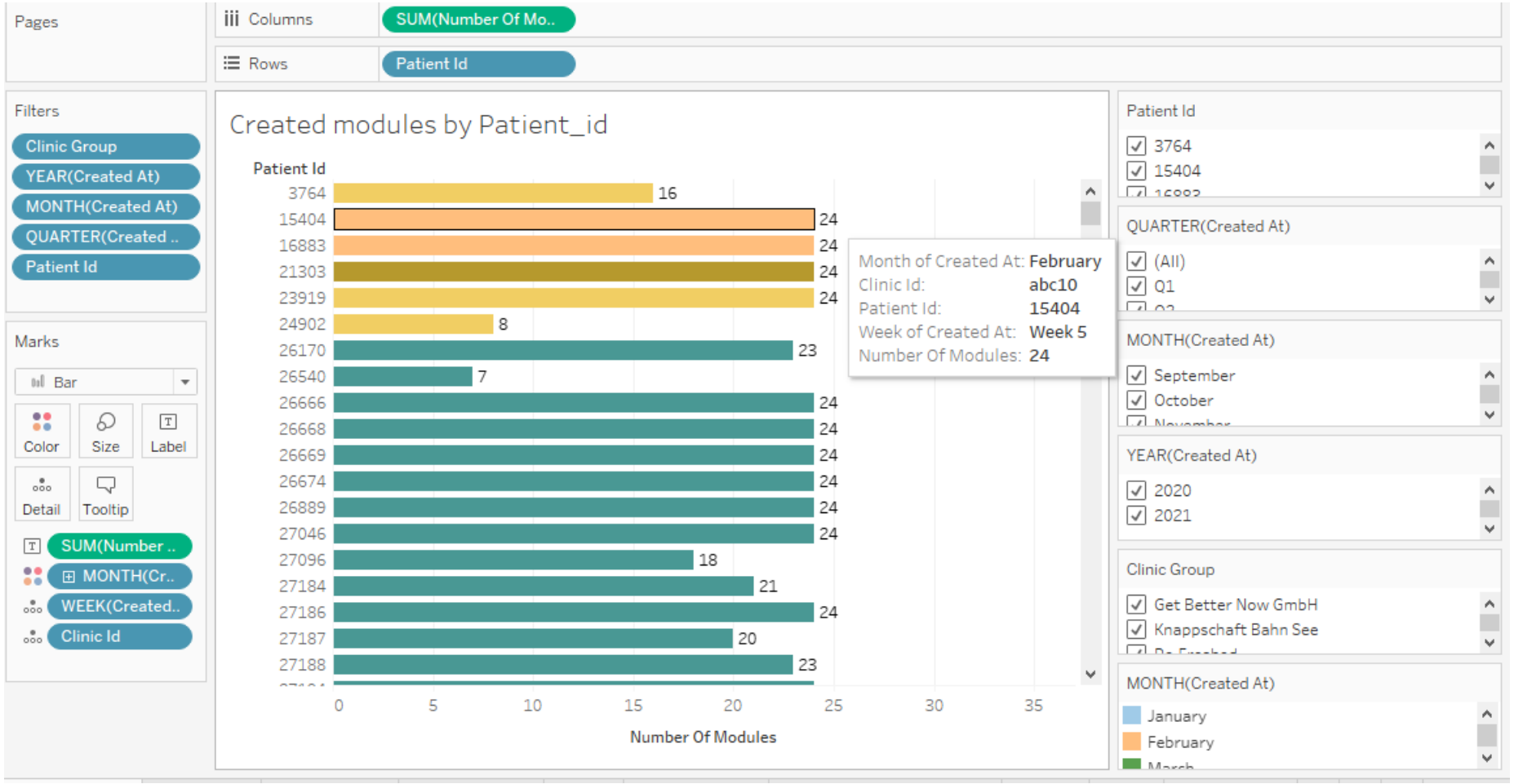
There are not null values, at least in the new database

Tableau

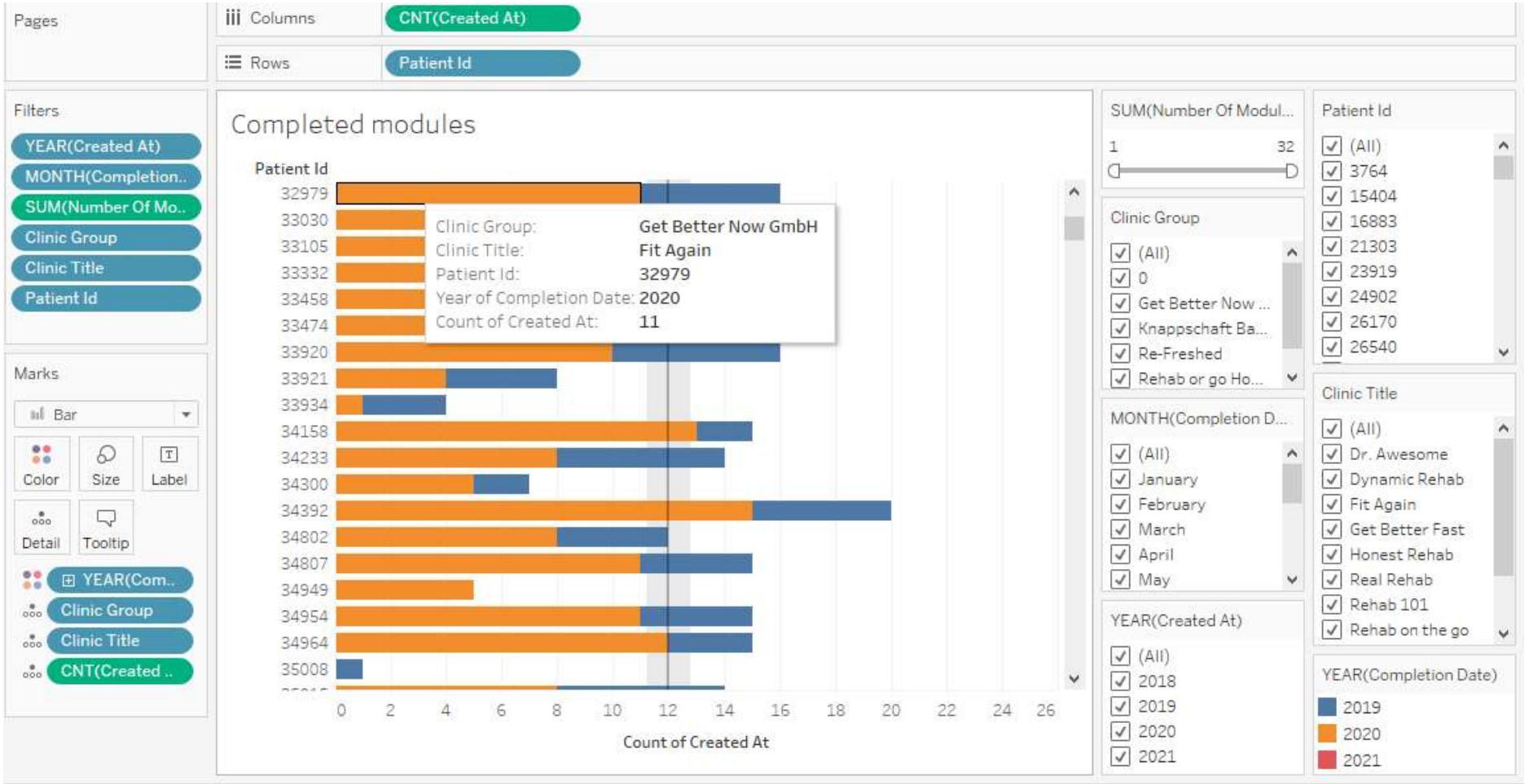
Outliers: number of completed modules



Visualizations:
Created Modules

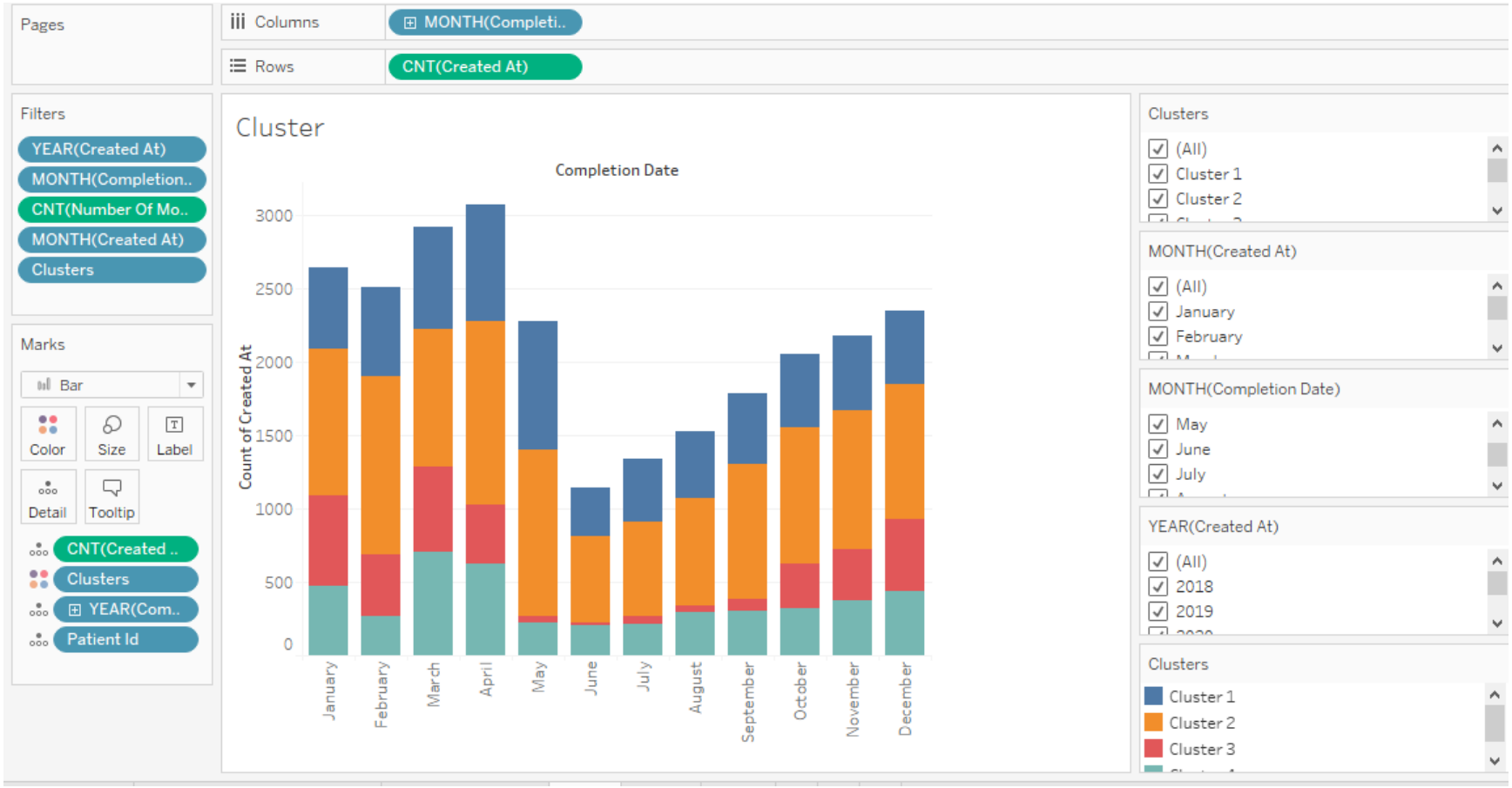


Visualizations:
Completed Modules vs created modules

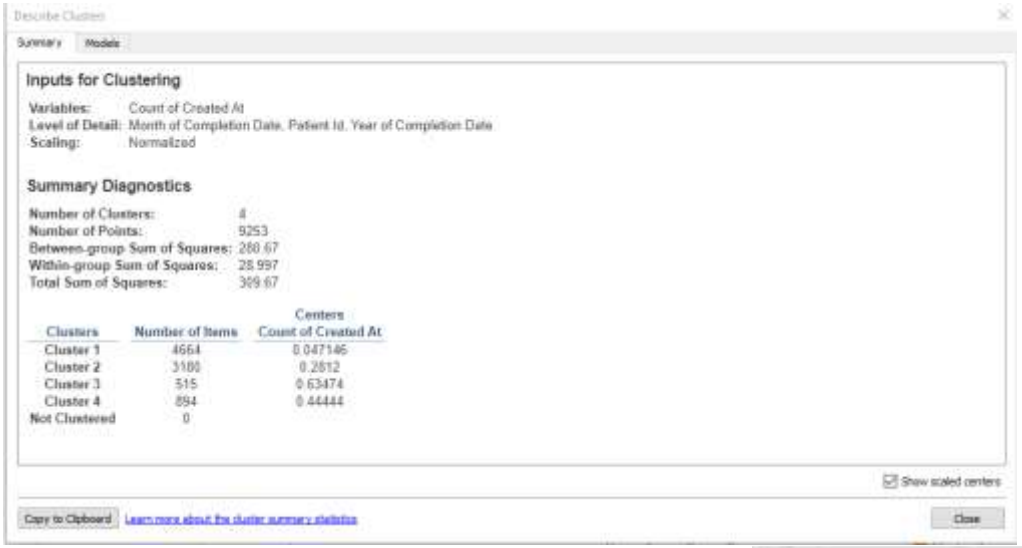


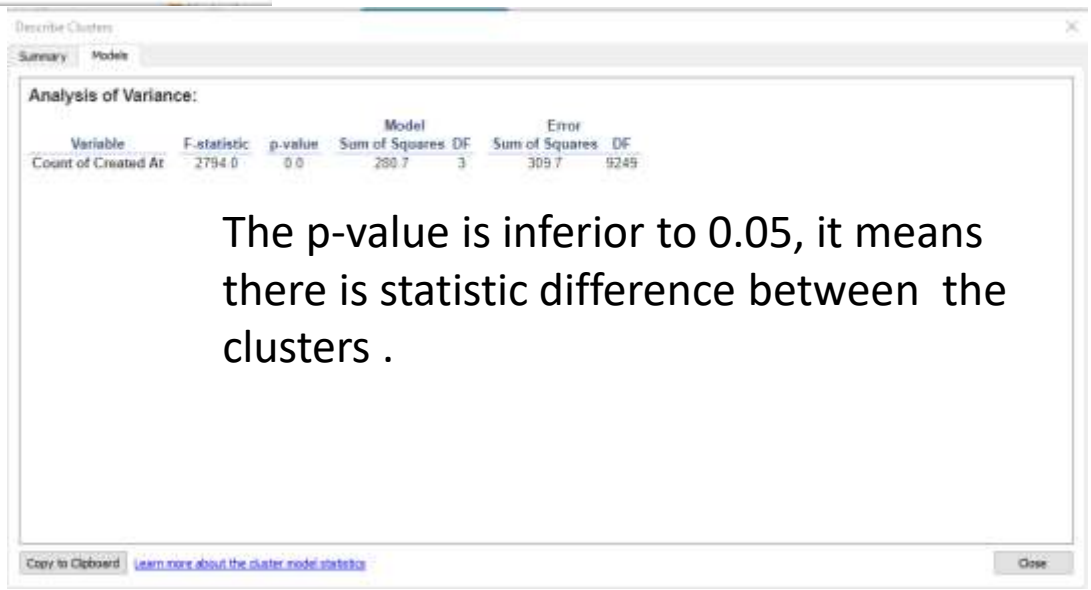
Median value with 95% CI=12

Clustering(1)



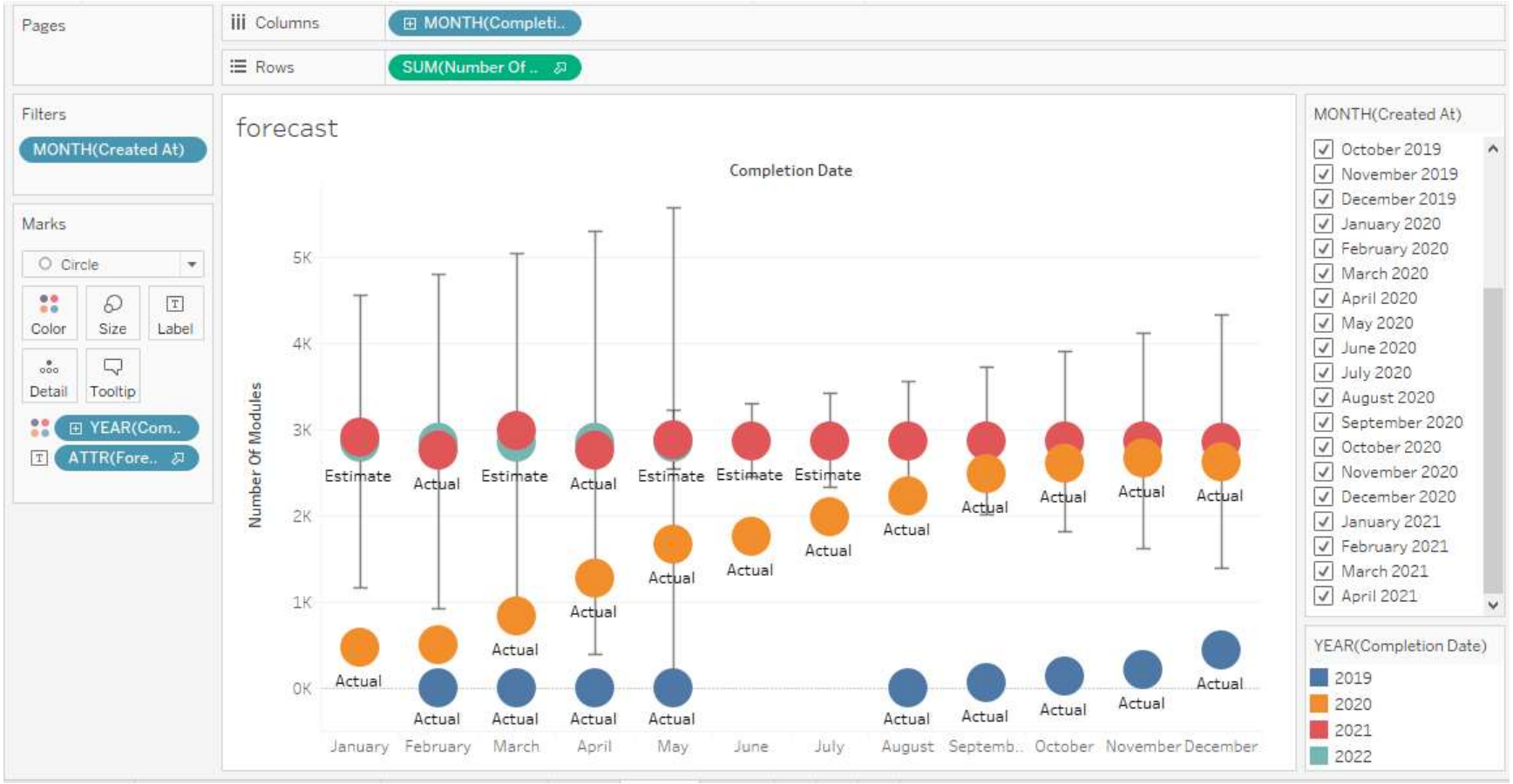
Clustering(2)



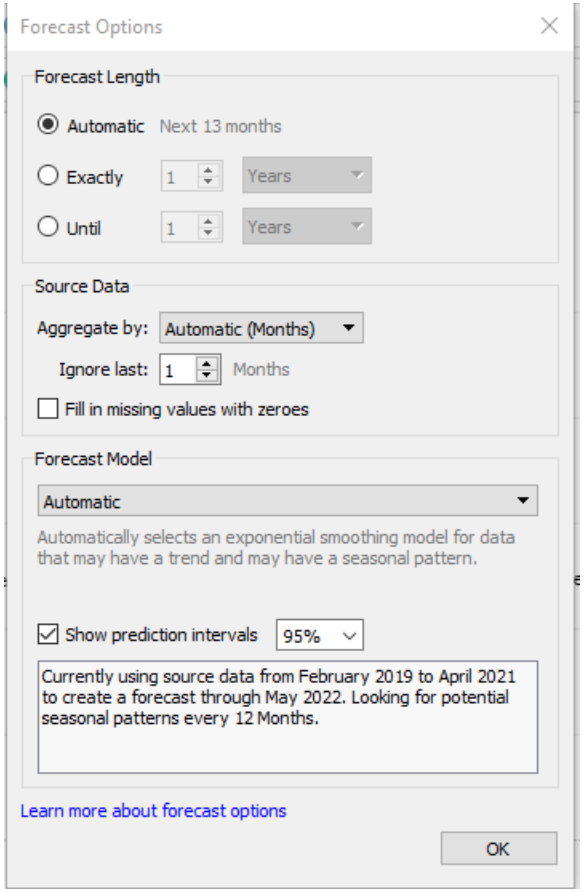


The p-value is inferior to 0.05, it means there is statistic difference between the clusters .

Forecasting (1)

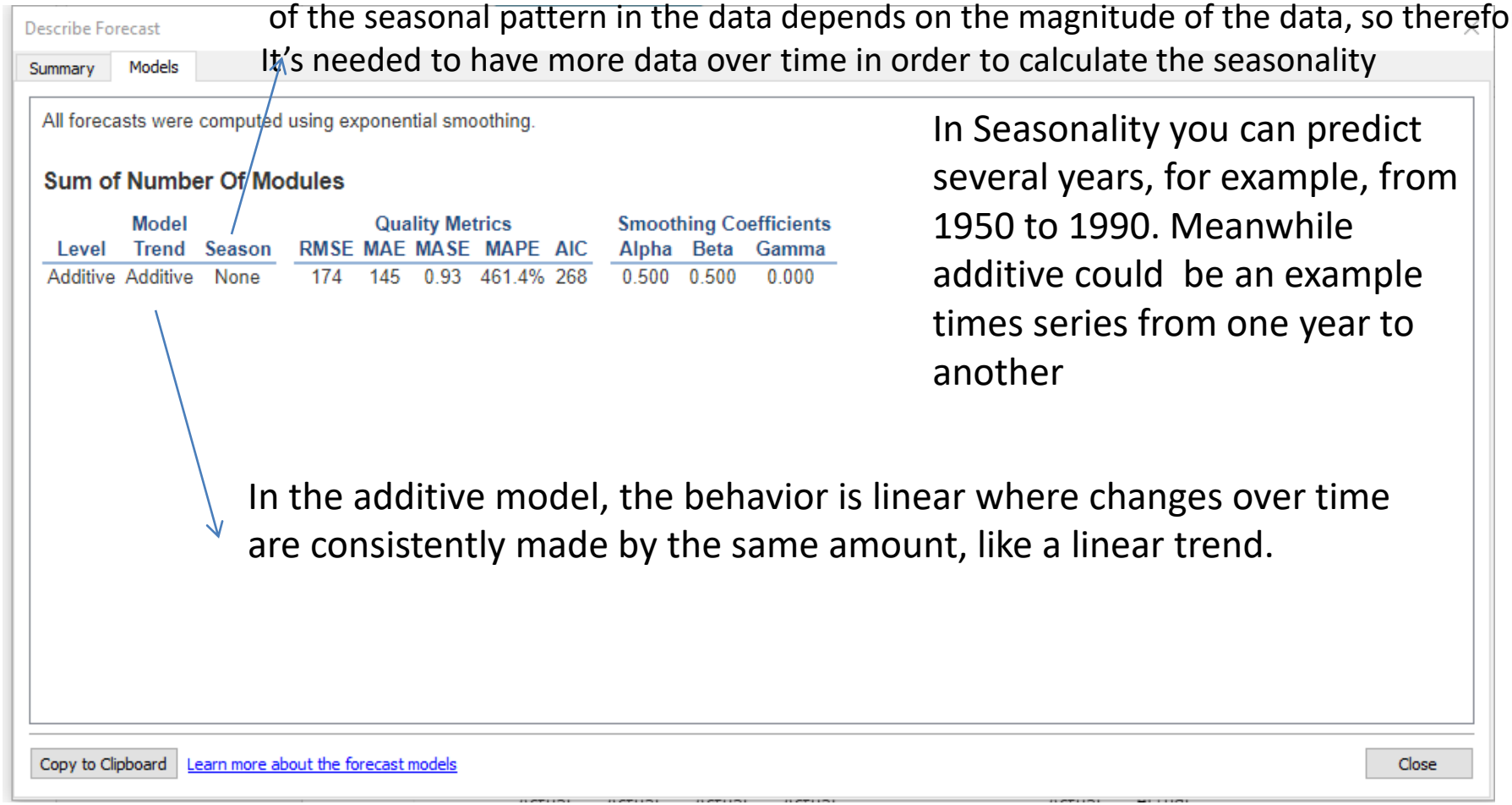


Forecasting (2)



Trend and Season depend very strong on the number of months (or other time of unit) in order to predict time series (how many modules will be done in 12 months)

There isn't a constant seasonality across the decomposition. The magnitude of the seasonal pattern in the data depends on the magnitude of the data, so therefore It's needed to have more data over time in order to calculate the seasonality



In Seasonality you can predict several years, for example, from 1950 to 1990. Meanwhile additive could be an example times series from one year to another

In the additive model, the behavior is linear where changes over time are consistently made by the same amount, like a linear trend.