# STask Details

Using the supplied predictive variables (GRE score, TOEFL score, University Rating, etc) to predict the likelihood of admission of a new candidate.

# Data Set

Obtained through Kaggle.

Link:- <https://www.kaggle.com/mohansacharya/graduate-admissions>

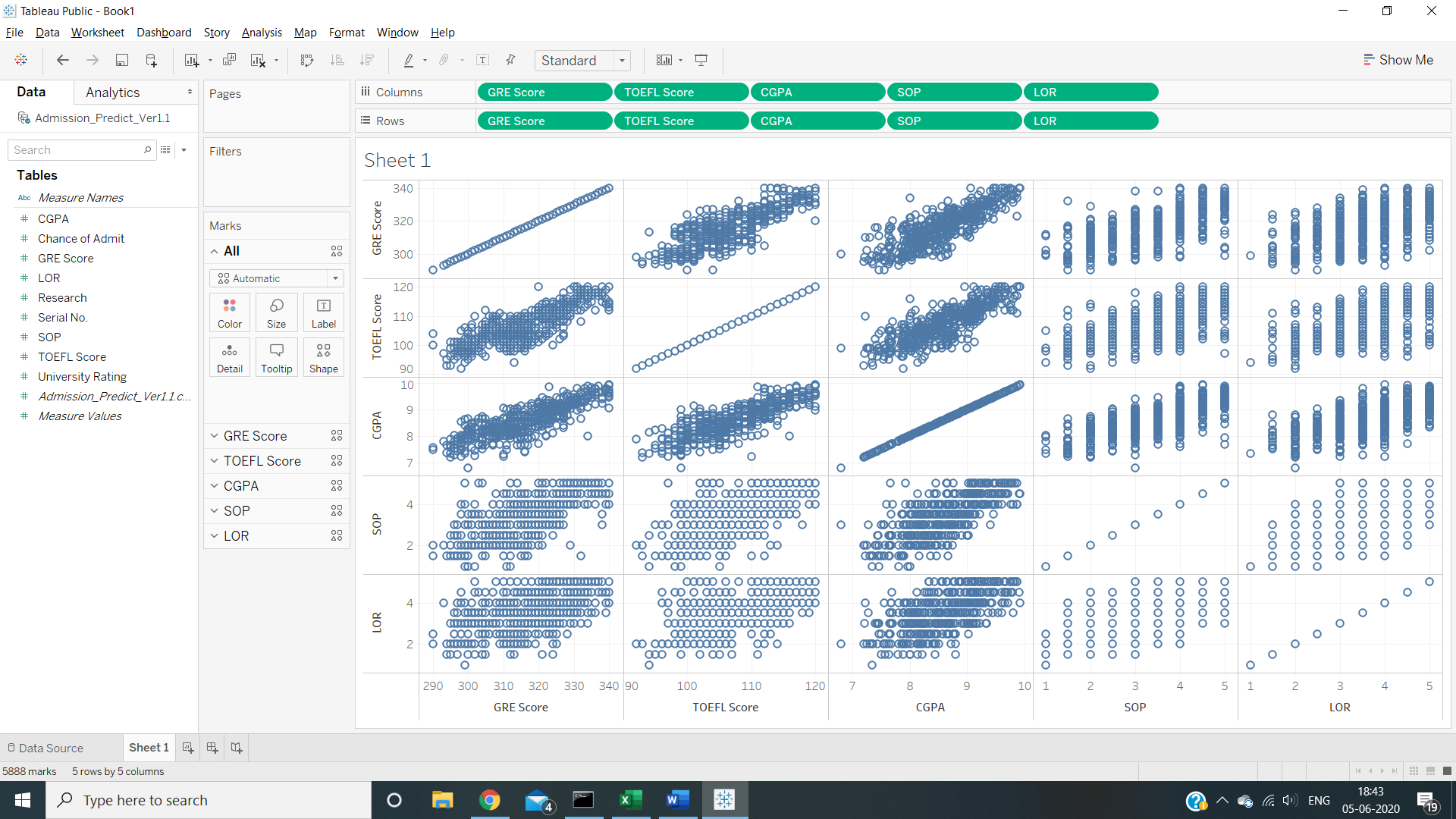
# Data Cleaning

Most of the checks on the data were done directly through excel. It wasn't a huge challenge as the data set wasn't that big. I just had to change the column name of Serial No. , LOR and Chance of Admit. Also, I made sure that Seriol No had been numbered correctly from 1 to 500.

I had the following columns in mind for making the model :-

GRE Score, TOEFL score, SOP, LOR, CGPA

I mapped each of these parameters with the other on Tableau.



I found out that GRE, TOEFL and CGPA were very strongly related to each other. Therefore, I decided to take just the CGPA.

Therefore, the parameters that will be used are CGPA, SOP and LOR.

I found the RSME for Linear Regression and Random Forests. As it turns out, Linear regression had the least RMSE of 0.05