1. A python notebook named “preprocess.ipynb” is created to do some initial processing before using Tableau
2. Raw data “data\_v1.xlsx”
3. Some fields are renamed as "Country Name": "Country", "Time": "Year", "Ease of Doing Business Score": "Business Score", "Rail Lines”: "Rail Lines Length (km)", "Railways, Passengers Carried": "Paassenger Distance Traveled”, “International Tourism, Number of Arrivals”: "Incoming Tourists" for better understandings
4. Unnecessary information is removed from the tail part
5. The data type of “Year” column is changed to integer from float
6. 25% of the data have no Business score. We cannot ignore them as they contain some other useful information like GDP, Incoming Tourists.
7. Formatted data is saved as “data\_v2.xlsx”
8. Only 2019 (latest in the dataset) data are analysed.