Data Science Tableau Project

By

Sidak Pal Singh Woodwal

102003149

COE-7

Thapar University

Submitted to

Dr. Kashish Goyal

Thapar University

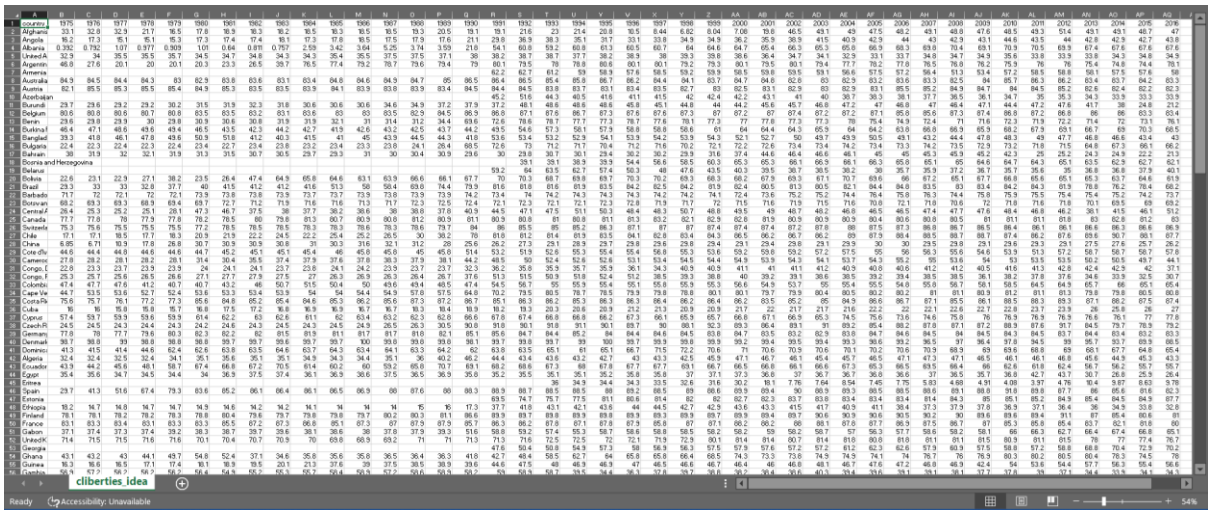
Objective: To analyse and try to establish the relationship between Civil liberty and Judicial independence across nations and highlight the importance of an independent judiciary in maintaining and protecting the liberties of its people.

Datasets used: The datasets have been sourced from gapminder.org. Gapminder is an independent educational non-profit fighting global misconceptions. 2 datasets have been used in this project, namely:

ATTRIBUTION: FREE DATA FROM WORLD BANK VIA GAPMINDER.ORG, CC-BY LICENSE

1) **Judicial independence index:** The judicial independence sub attribute denotes the extent to which the courts are not subject to undue influence from the other branches of government, especially the executive. This sub attribute includes five indicators: High Court independence, Lower Court independence, compliance with High Court, compliance with judiciary and law and order. These five indicators were aggregated into the judicial independence sub attribute using IRT.
2) **Index of the civil liberties: "This is one of five subindexes of the democracy index composed from the following indicators estimated by experts in each field:**
   1. Is there a free electronic media?
   2. Is there a free print media?
   3. Is there freedom of expression and protest (bar only generally accepted restrictions, such as
   banning advocacy of violence)?
   4. Is media coverage robust? Is there open and free discussion of public issues, with a reasonable diversity of opinions?
   5. Are there political restrictions on access to the Internet?
   6. Are citizens free to form professional organisations and trade unions?
   7. Do institutions provide citizens with the opportunity to petition government to redress grievances?
   8. The use of torture by the state.
   9. The degree to which the judiciary is independent of government influence. Consider the views of international legal and judicial watchdogs. Have the courts ever issued an
   important judgement against the government, or a senior government official?
   10. The degree of religious tolerance and freedom of religious expression. Are all religions permitted to operate freely, or are some restricted? Is the right to worship permitted both publicly and privately? Do some religious groups feel intimidated by others, even if the law requires equality and protection?
   11. The degree to which citizens are treated equally under the law. Consider whether favoured groups or individuals are spared prosecution under the law.;
   12. Do citizens enjoy basic security?

Civil Liberties



Judicial Independence

Data Cleaning Process

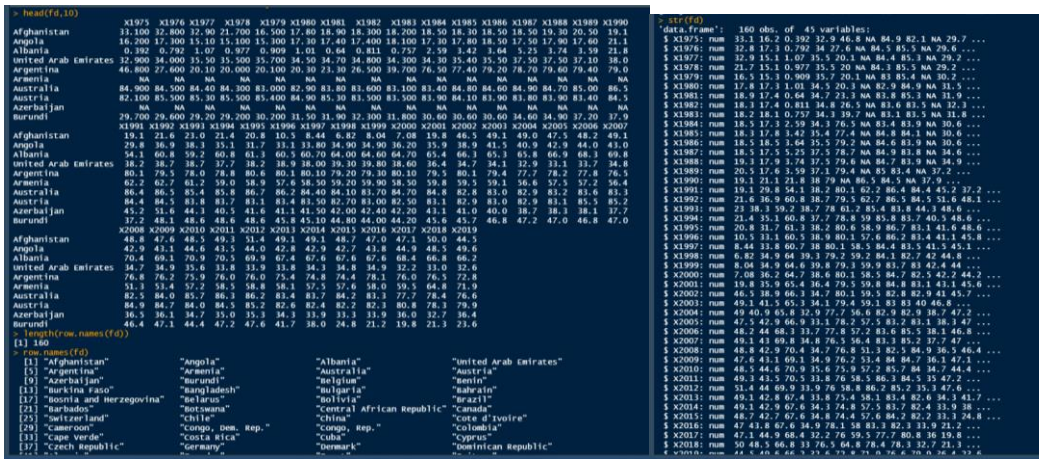
Summary:

1) Rename variables
2) Replace NA values
3) Derive new columns from existing ones
4) Merge cleaned datasets into one
5) Use merged set to create final variables used in visualization

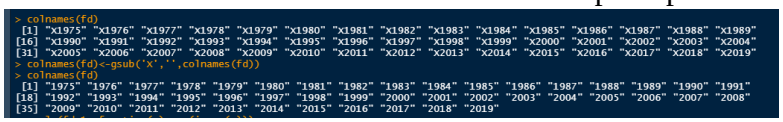
Civil Liberties Dataset:

The zoo library is used to plot the na_pareta plots

The dataset was read into RStudio and I carried out an evaluation of the summary and structure of the dataset.

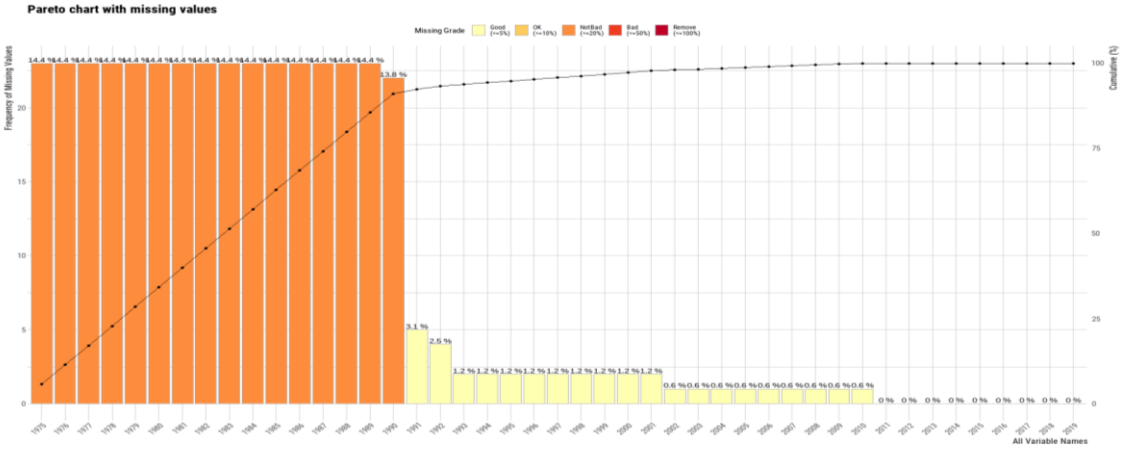


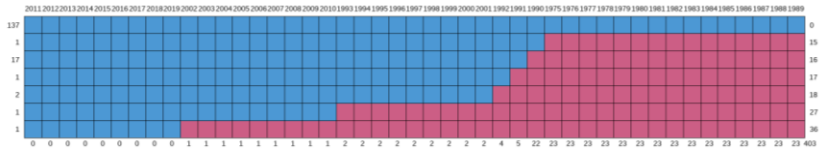
Then the variable names were modified to drop the prefix "X" from the year names.



Following this I measured count of missing values country-wise as well as sum of missing values. These values were handled using the na_locf which means last observation carried forward as in case of these indices, year by year value for a nation is much more likely to be similar to the previous year rather than an aggregate measure.

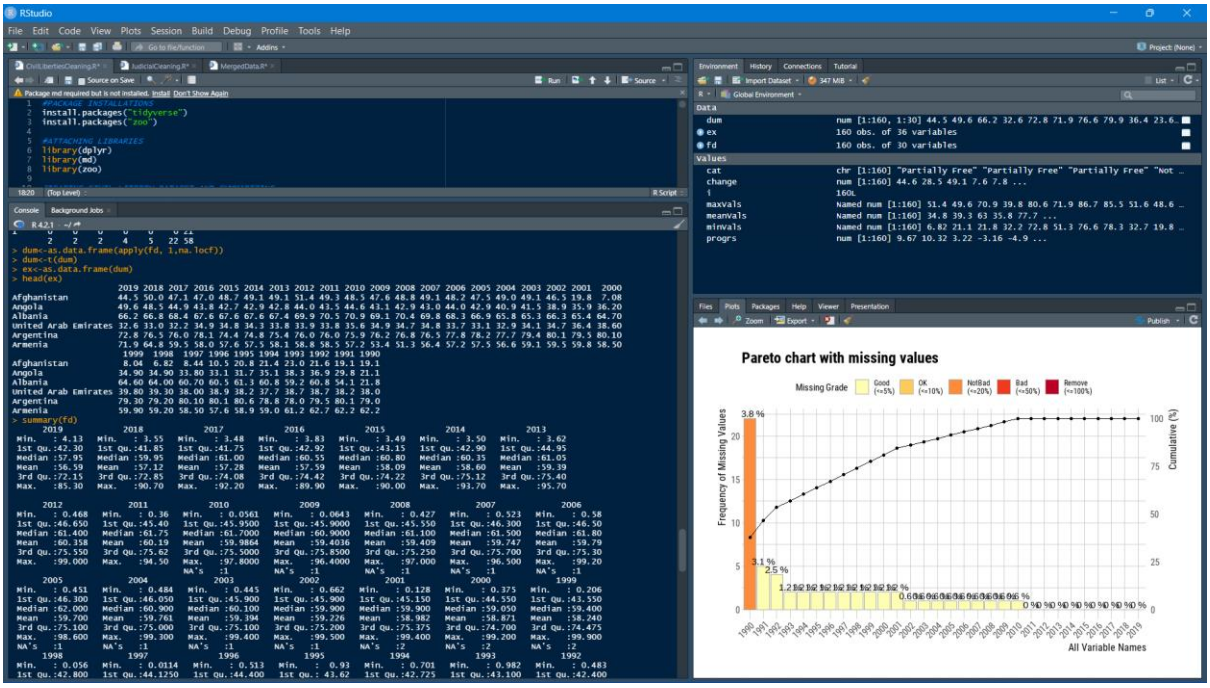
Here is a visualization of the missing values using a pareta plot





Plot of non-missing(blue) vs missing(red) values

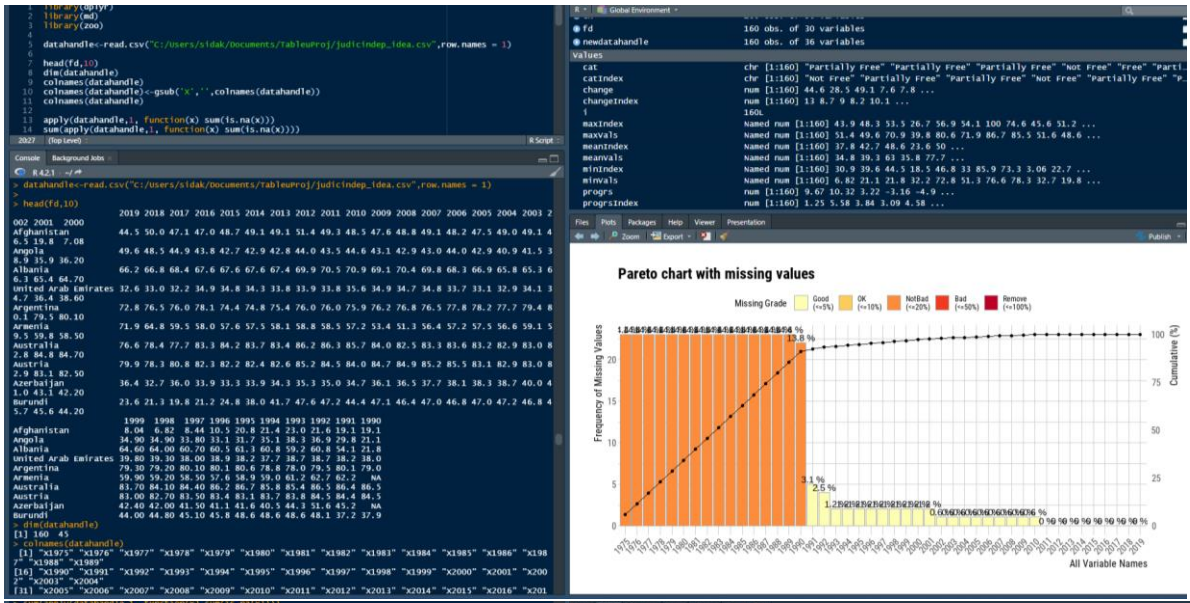
Since there are many missing values before 1990, I truncated the data to only include data from 1990 to 2019.

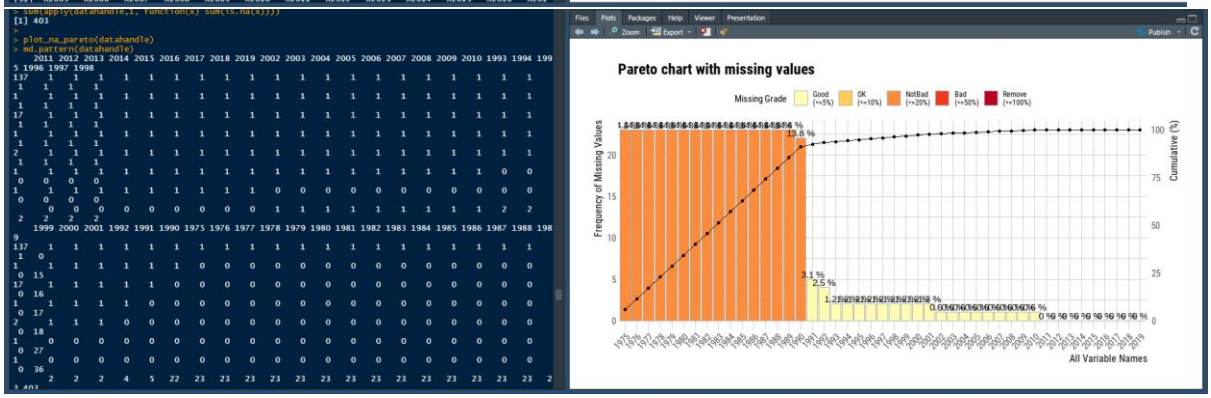


After truncating data, I removed the NAs by replacing them with the previous year's index for the country. Following this, I created new columns in the data frame including minimum, maximum and mean values of liberty indices of each country . Using these values, I created new columns which calculate the difference between the maximum and minimum value of indices to enumerate the maximum change a nation has been through. I created another column which measures the latest index of a nation (2019) against the mean of values from 1990 to 2019 to show how a nation performed lately. Finally, I created a categorical column on basis of maximum index of a nation and classified the first and last quantile as Free and No Free respectively.

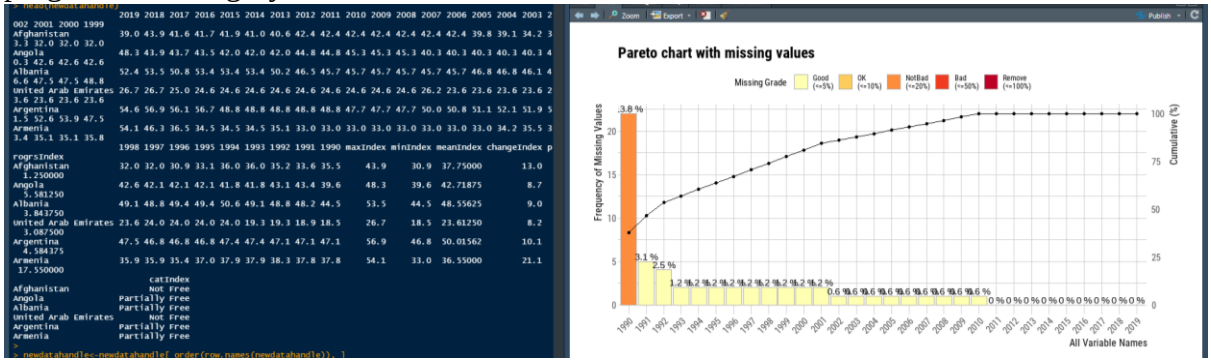
Judicial Independence Index

This dataset was cleaned similar to the previous dataset where I viewed and visualized the dataset, renamed the variables and calculated the total number of NA values in the dataset.



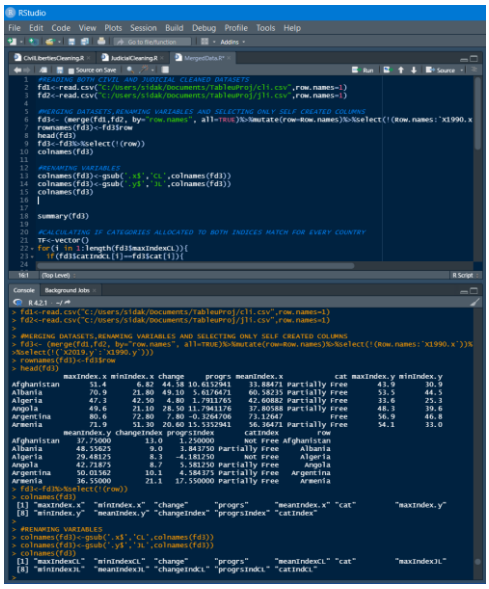


As visible above, again the years before 1990 had large amounts of missing data and was removed, the remaining was filled using locf method. The complete data was used to create similar new variables such as min, max and mean as well as change, progress and category columns were created for the dataset.

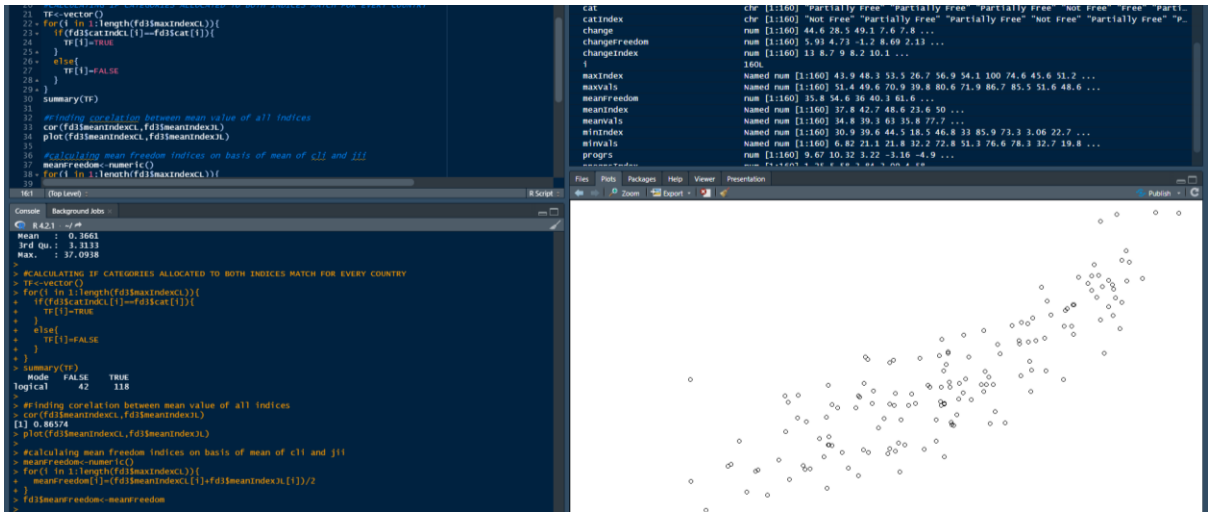


Merging the datasets together

I merged the datasets together using natural join and pipelined the output into dplyr select functions to remove the year wise data and only keep the columns that were previously defined by me. The variables were renamed.



After this I took the freedom category columns from both the datasets and evaluated whether the data agreed on the category allotted to each country. Here are the results:

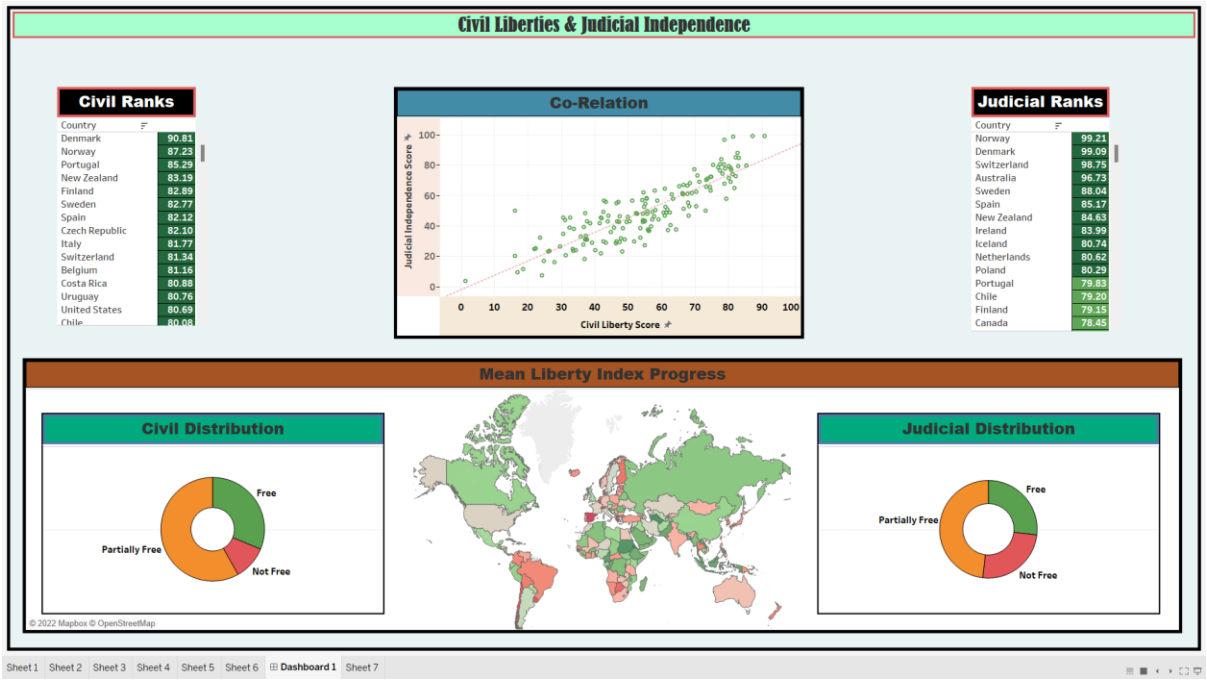


118 instances agree upon the categorisation while 42 disagree. Also the corelation between civil liberty and judicial independence is +0.86574 meaning a strong positive correlation .

Lastly, I have calculated the mean of civil and judicial liberties to find the mean change in index and the mean civil-judicial freedom value

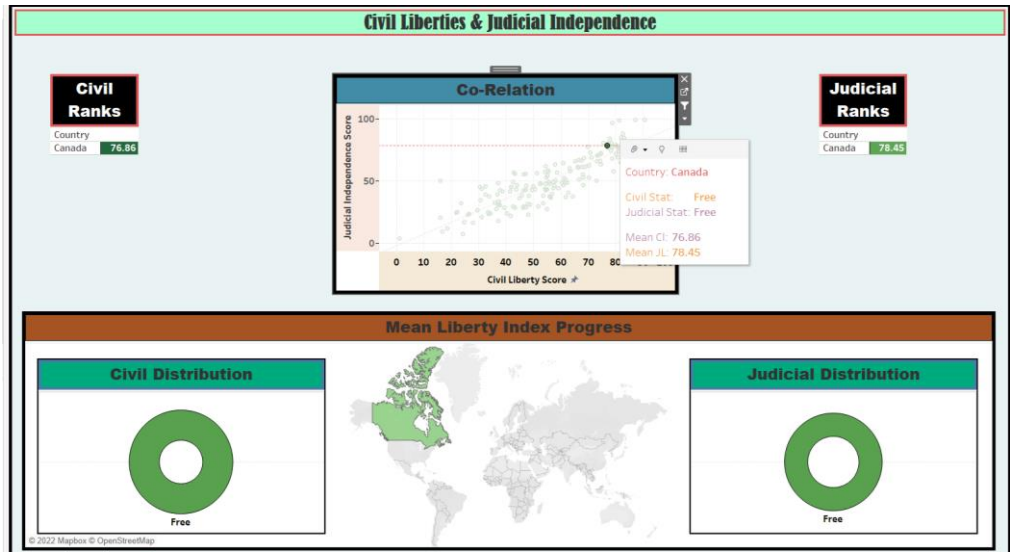
Tableau Dashboard

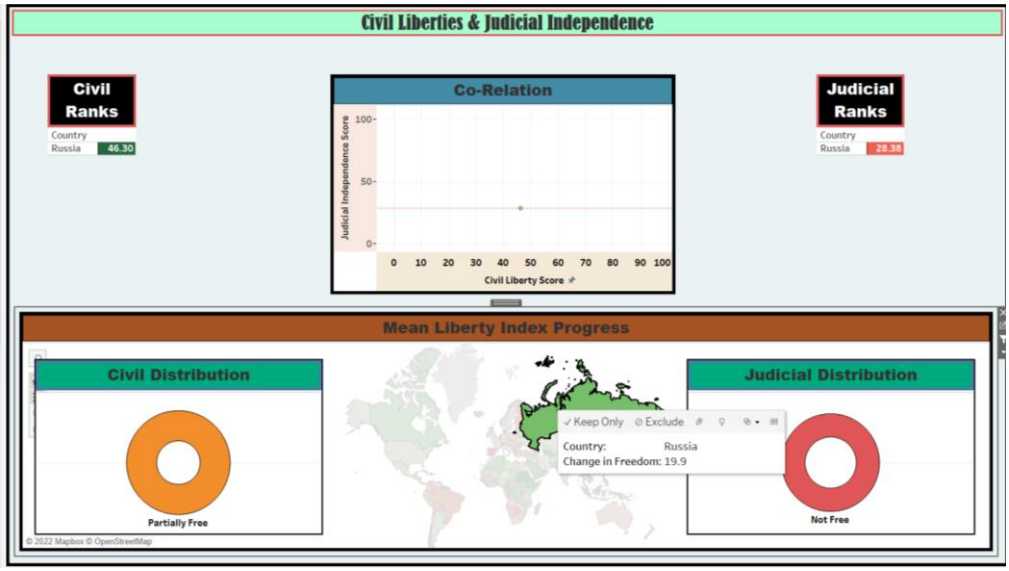
Data visualisation is the final aspect of the data science project. To facilitate visualization in my project, I have made use of tableau software. Here is the created dashboard:



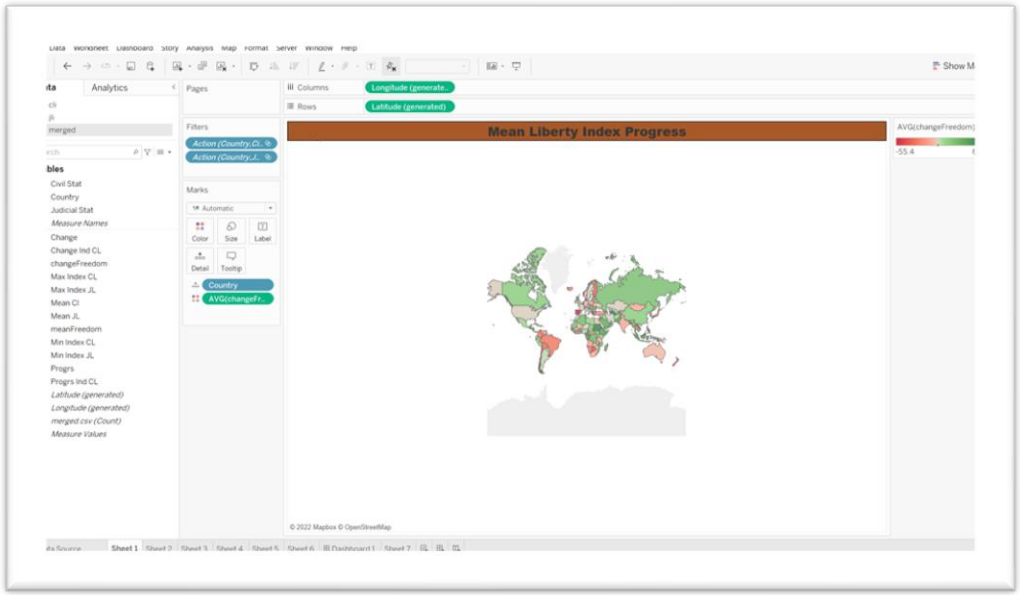
In the centre on the top row I have shown the graph of correlation between mean civil liberty and mean judicial index. To the left and right are highlight tables in which nations are ranked on the maximum score achieved by them in these rankings. The bottom row consists of a map which shows the progress of different nations in 2019 compared to the mean of all years from 2019 to 1990. Beside the map are 2 pie-charts depicting the proportion of categories the nations have been divided into.

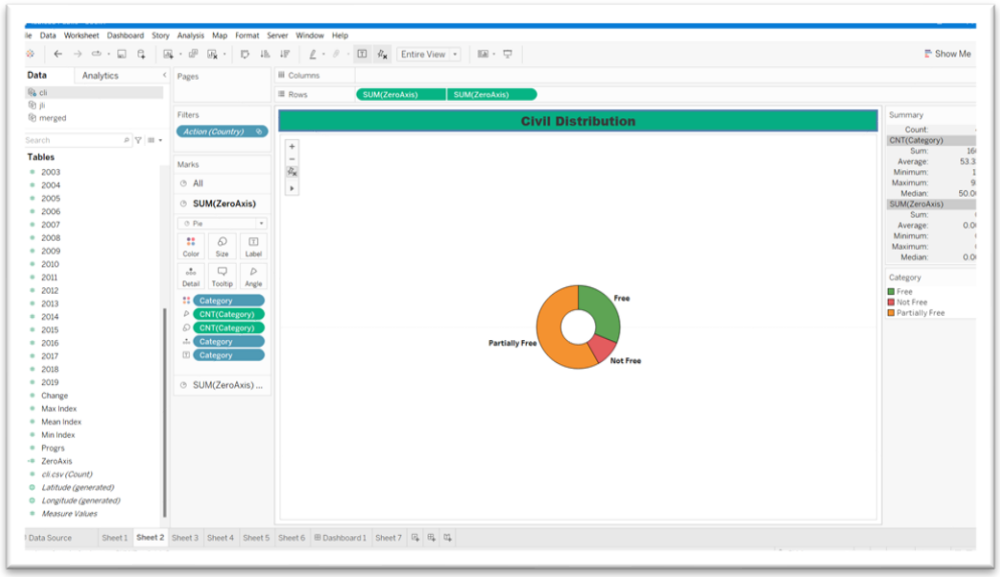
Filters have been used to provide capability to select a country from the map or the graph to bring up all details relevant to that country.

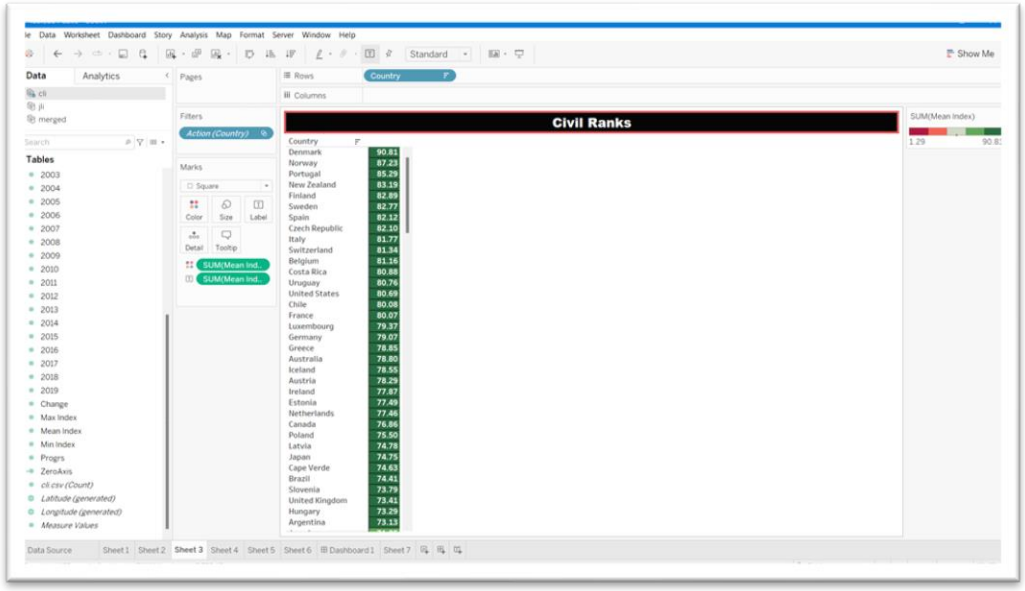


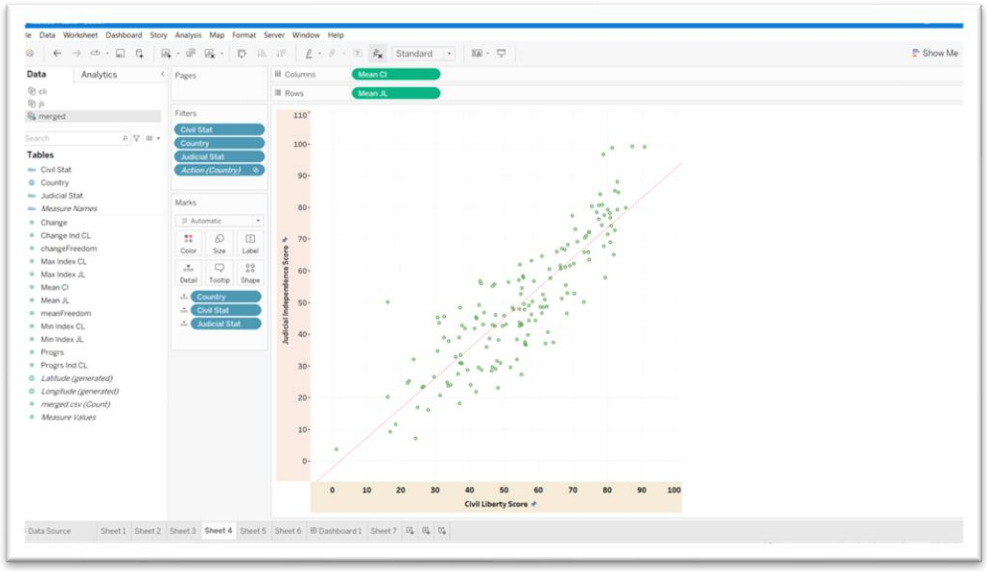


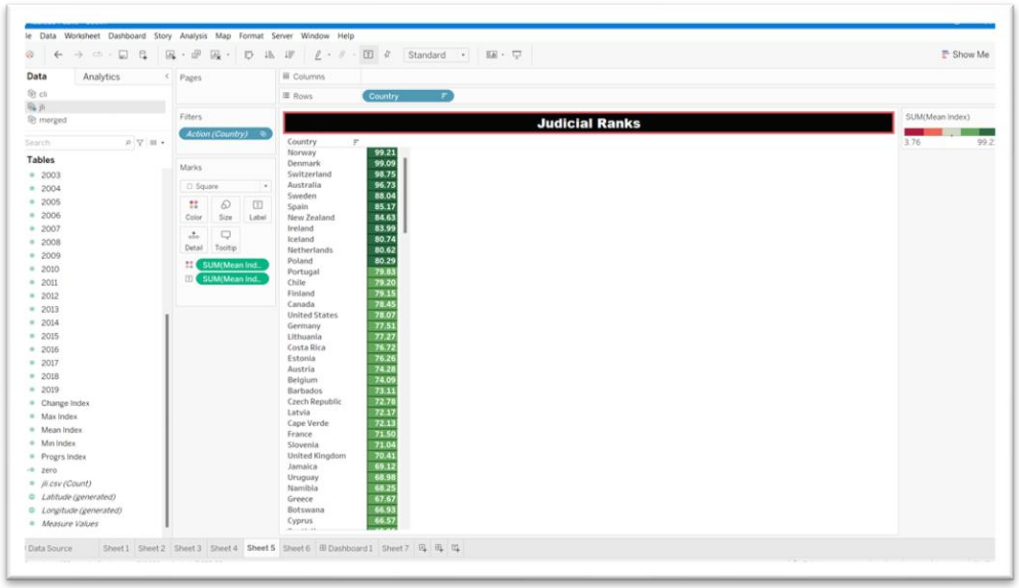
Here are the component worksheets:

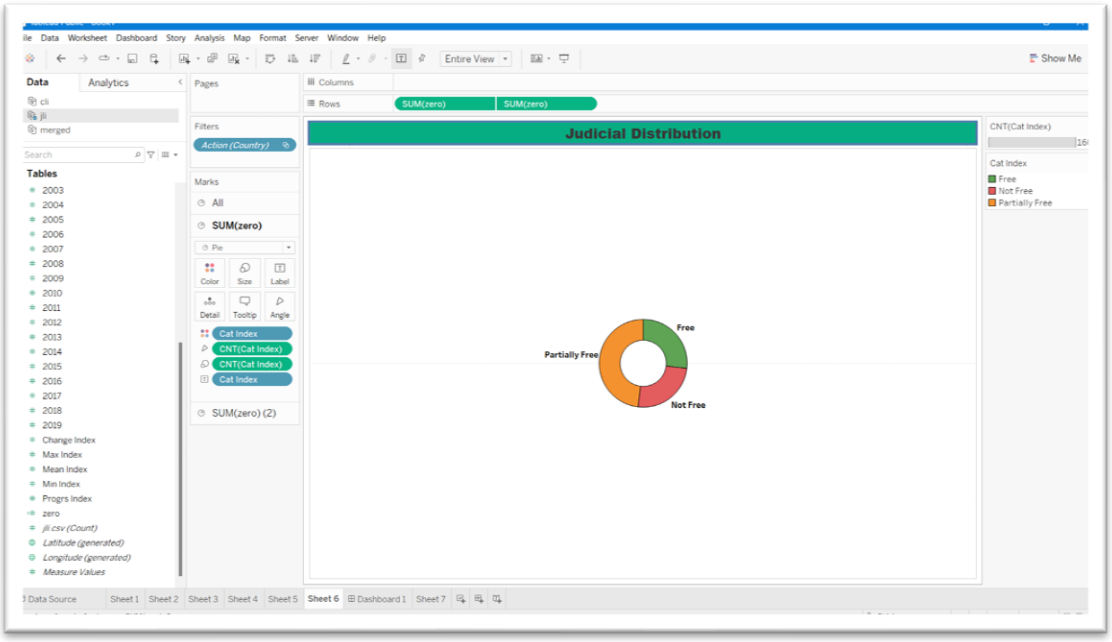



Civil Distribution
Free
Partially Free
Not Free


Civil Ranks
Country
Denmark 90.81
Norway 87.23
Portugal 85.29
New Zealand 83.19
Finland 82.89
Sweden 82.77
Spain 82.12
Czech Republic 82.10
Italy 81.77
Switzerland 81.34
Belgium 81.16
Costa Rica 80.88
Uruguay 80.76
United States 80.69
Chile 80.08
France 80.07
Luxembourg 79.37
Germany 79.07
Greece 78.85
Australia 78.80
Iceland 78.55
Austria 78.29
Ireland 77.87
Estonia 77.49
Netherlands 77.46
Canada 76.86
Poland 75.50
Latvia 74.78
Japan 74.75
Cape Verde 74.63
Brazil 74.41
Slovenia 73.79
United Kingdom 73.41
Hungary 73.29
Argentina 73.13


Judicial Independence Score
Civil Liberty Score


Judicial Ranks
Country
Norway 99.21
Denmark 99.09
Switzerland 98.75
Australia 96.73
Sweden 88.04
Spain 85.17
New Zealand 84.63
Ireland 83.99
Iceland 80.74
Netherlands 80.62
Poland 80.29
Portugal 79.83
Chile 79.20
Finland 79.15
Canada 78.45
United States 78.07
Germany 77.51
Lithuania 77.27
Costa Rica 76.72
Estonia 76.26
Austria 74.28
Belgium 74.09
Barbados 73.11
Czech Republic 72.78
Latvia 72.17
Cape Verde 72.13
France 71.50
Slovenia 71.04
United Kingdom 70.41
Jamaica 69.12
Uruguay 68.98
Namibia 68.25
Greece 67.67
Botswana 66.93
Cyprus 66.57


Judicial Distribution
Free
Partially Free
Not Free
Cat Index
Free
Not Free
Partially Free