STREAMERS

Climate Change:
An analysis of World temperature and CO2 emissions.

Temitope Adeniyi
Feven Belay
Neca Bryan

OVERVIEW

This is an analysis of the Earth's surface temperature from 1900-2013 and CO2 data from 1990-2018 to determine the trends and to build models to predict the variability of the World future temperatures

OBJECTIVES

- Affirms the consensus on global warming.

- Highlight temperature and Carbon dioxide trends across the major industrial countries of the world.

- To use the dataset to build machine learning model to predict the variation of the world`s future temperatures.

PROJECT QUESTIONS

Climate Change

- What is the trends of average temperature of the world?
- What is the trends Co2 emissions in the world ?
- To determine most appropriate machine learning models for the datasets
- To calculate the accuracy and precision of the models?

CONTENTS

The areas of focus for this analysis will be

- **Technologies** – Python, Jupyter Notebook, PostgreSQL, Tableau
- **Database** for creating tables and ascending data.
- **Machine Learning** to predict variations from world temperature, trends of temperature, predict whether dataset above or below the world mean temperatures
- **Summary of Results** of Data exploratory Phase, Data Analysis and Machine learning

TECHNOLOGIES

Python - Sqlalchemy- to load into the PostgreSQL

Jupyter Notebook – preprocessing data and to run ML

PostgreSQL – to store the 'Climate_Change DB' and create new tables

Tableau Public - visualizing data

DATA EXPLORATION PHASE

World temp on Jupyter Notebook

```
world_temp_byCountry
```

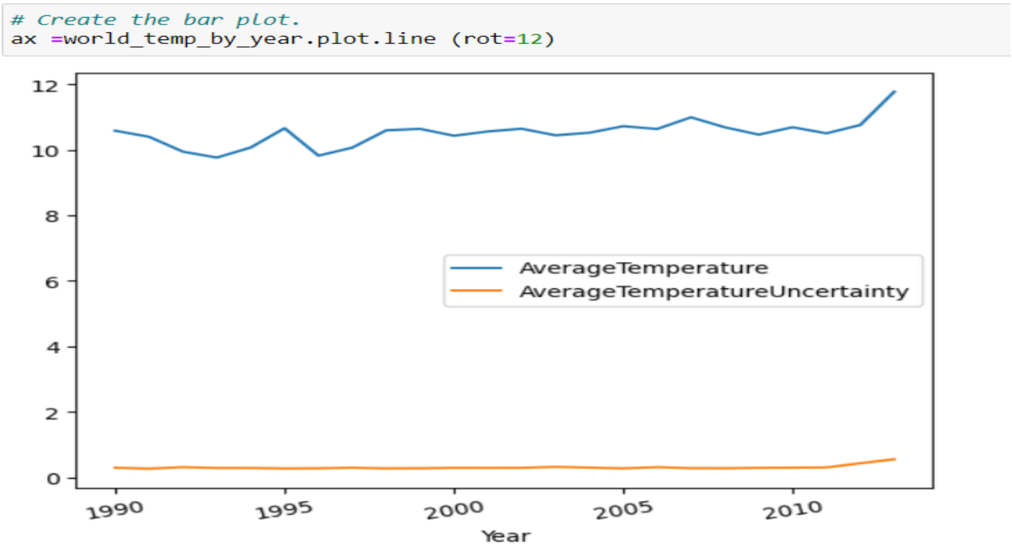
Country	AverageTemperature	AverageTemperatureUncertainty
Australia	18.488940	0.218123
Brazil	24.649577	0.372301
Canada	-0.504399	0.325735
China	12.322680	0.262353
India	22.934573	0.364369
Russia	2.655988	0.334668
United States	11.672875	0.217623

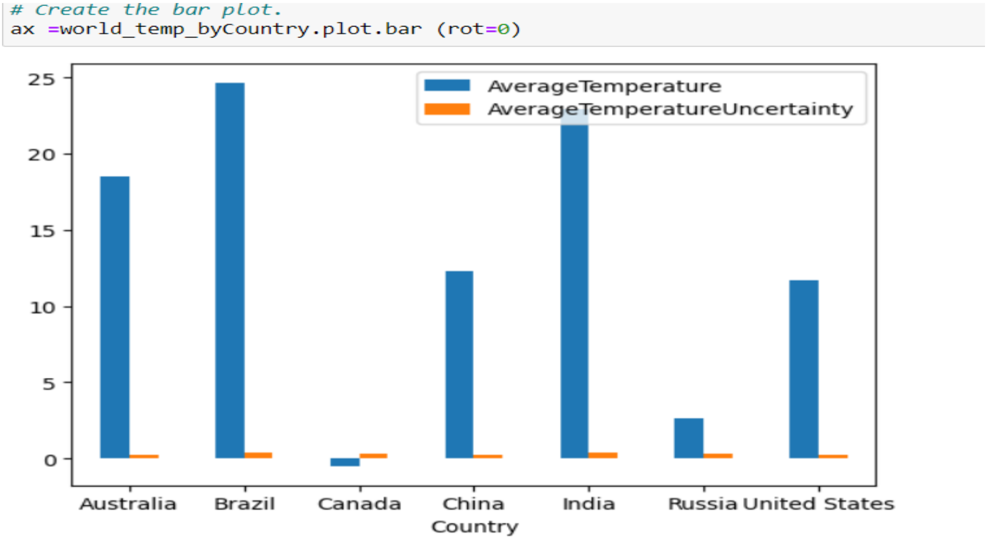
```
Global_temp.describe().transpose()
```

	count	mean	std	min	25%	50%	75%	max
AverageTemperature	328784.0	9.767931	13.931064	-45.389	0.399	12.196	20.91625	36.339
AverageTemperatureUncertainty	328784.0	0.417751	0.319615	0.036	0.238	0.330	0.48600	7.638
Year	328784.0	1956.344652	32.819143	1900.000	1928.000	1956.000	1985.00000	2013.000

Data Exploration Phase

of world temp on Jupyter Notebook

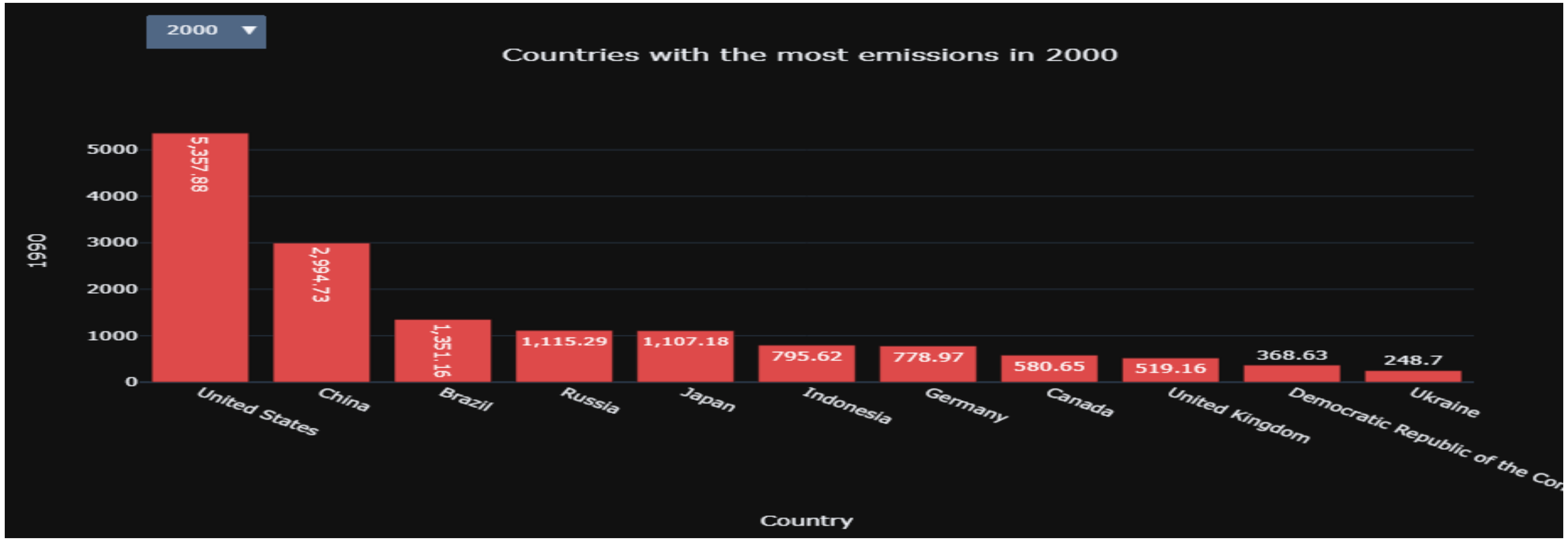


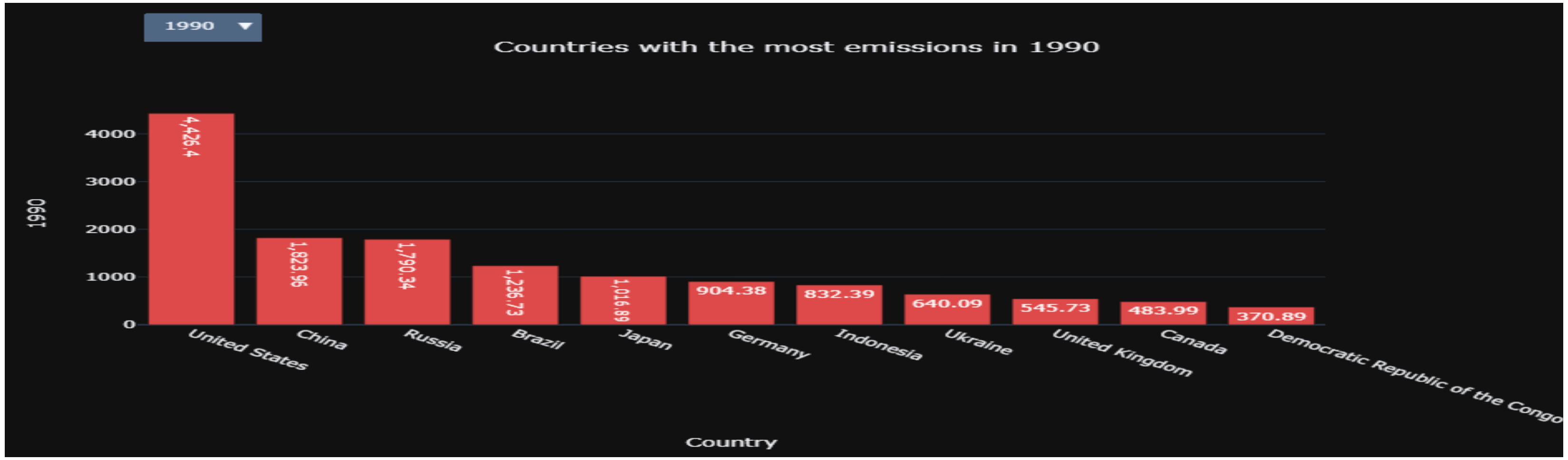


Data Exploration Phase

of world CO_2 emissions on Jupyter Notebook

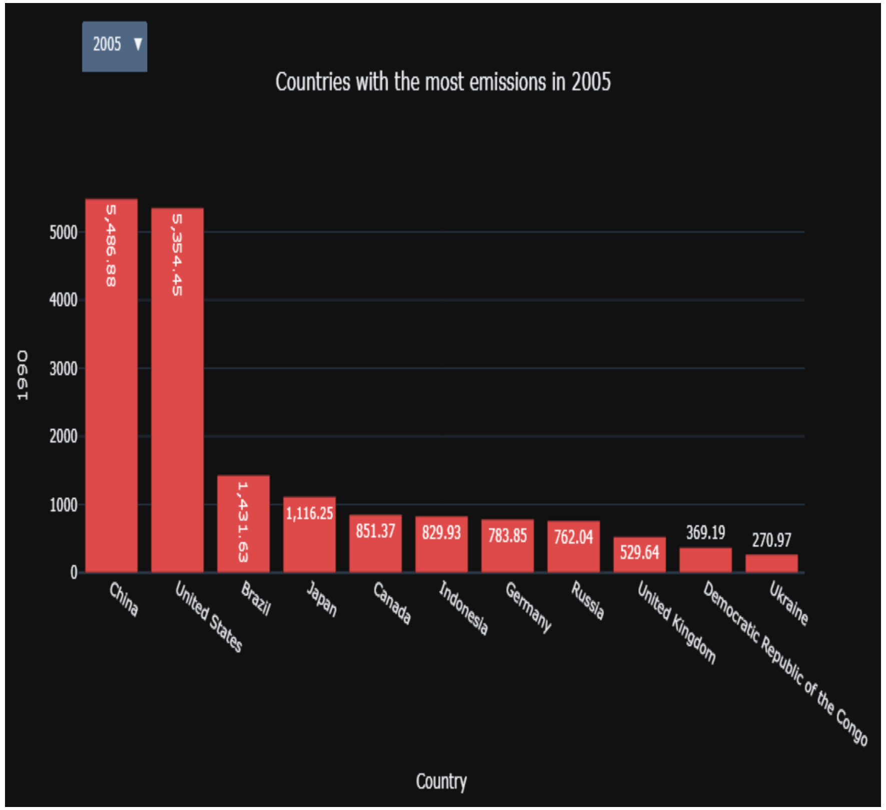
Before 2005 USA was main contributor CO_2 emissions

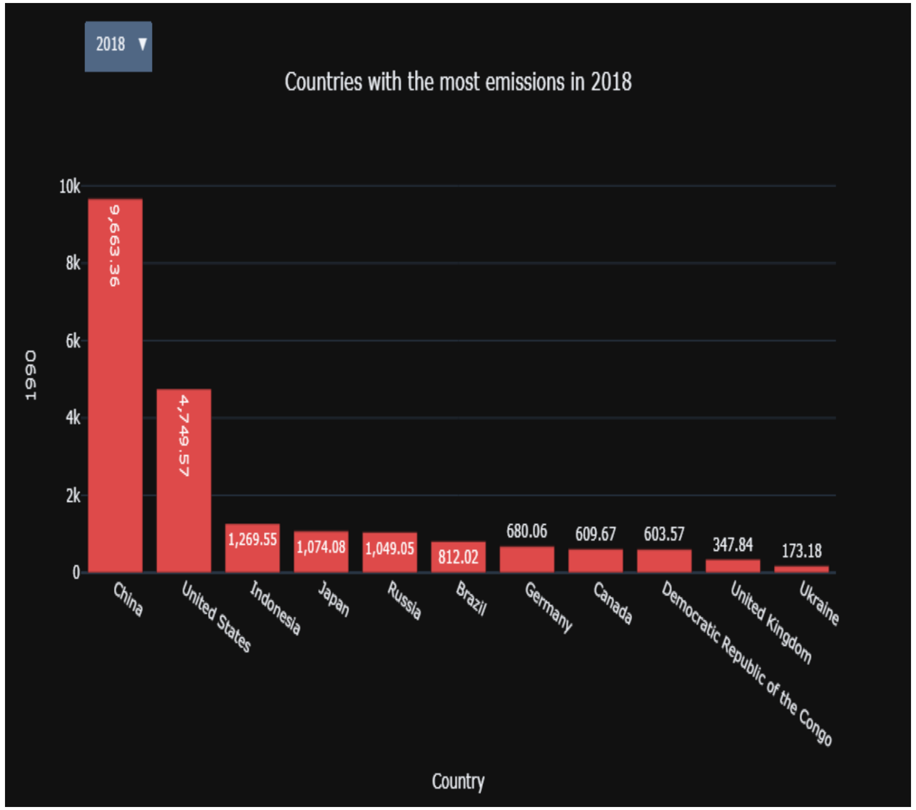




Data Exploration Phase

of world Co2 emissions on Jupyter Notebook





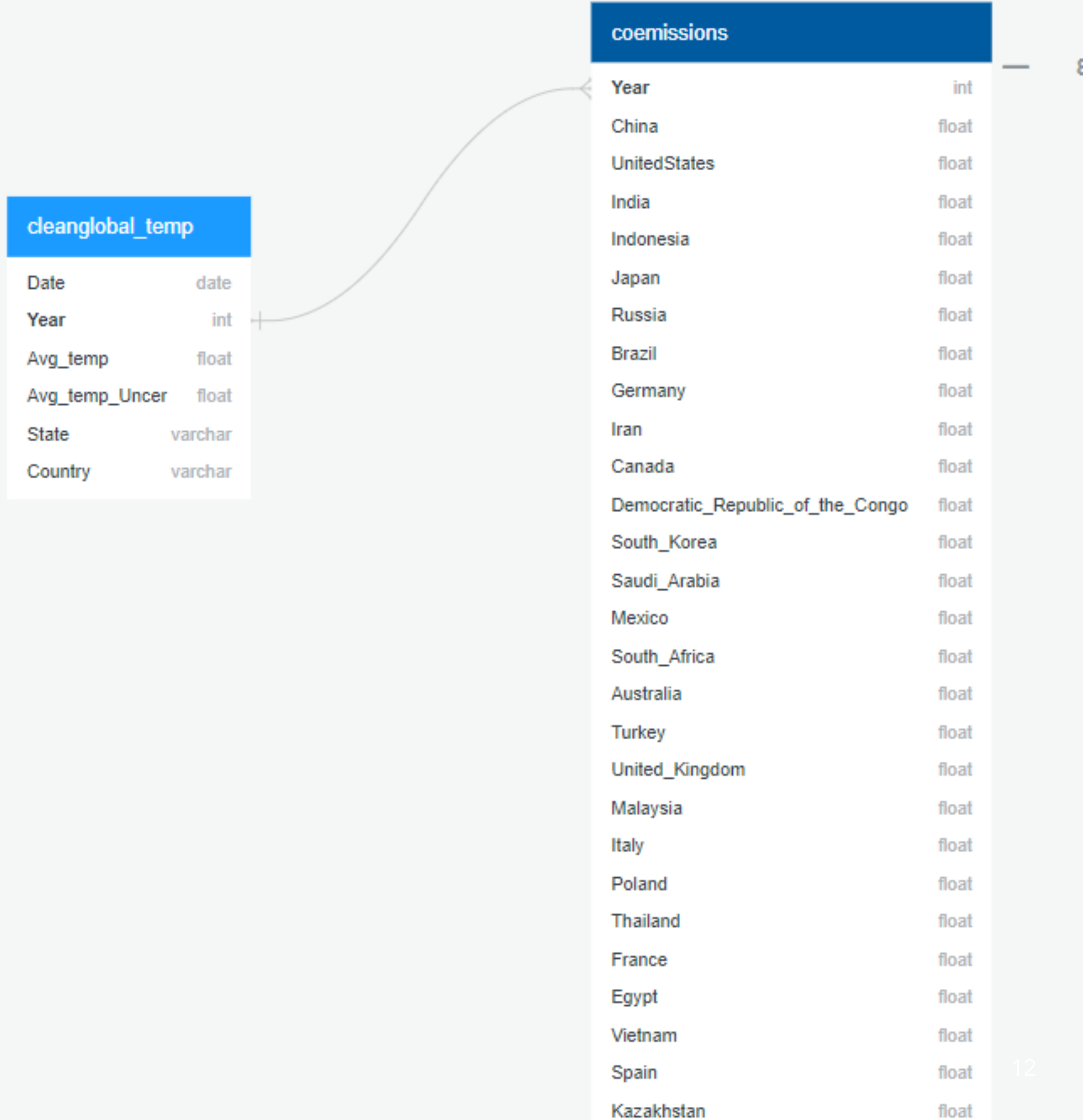
China overtook USA as main contributor Co2 emissions by 2005 and more than double USA emissions by 2018

DATABASE

PostgreSQL

Entity

Relationship Diagram



Climate_temp table

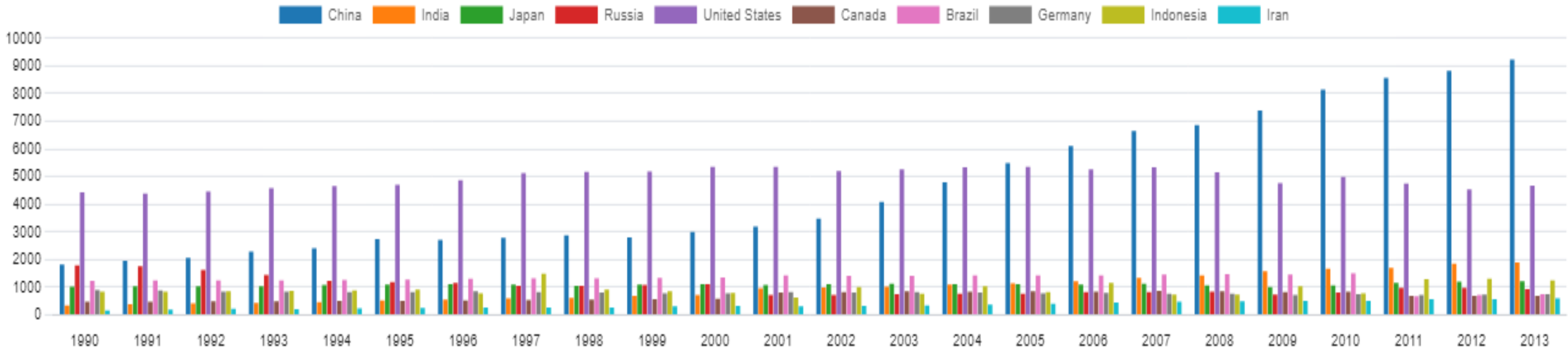
	Year bigint	avg_temp double precision	avg_temp_uncer double precision
1	1990	10.5799118257261	0.29483921161825
2	1991	10.3948243430152	0.26655290456431
3	1992	9.94019536652836	0.31140041493775
4	1993	9.76020401106502	0.28625587828492
5	1994	10.0674706085754	0.28581466113416
6	1995	10.656667704011	0.27156811894882
7	1996	9.81872302904563	0.27696092669432
8	1997	10.0640255878285	0.29461721991701
9	1998	10.592178769018	0.27458195020746
10	1999	10.63740802213	0.27794847856154
11	2000	10.4269713001383	0.29134439834024
12	2001	10.5575892116183	0.28895954356846
13	2002	10.6417130013831	0.29070089903181
14	2003	10.440566735823	0.32096887966805
15	2004	10.518698824343	0.29718948824343
16	2005	10.7185425311203	0.27509508990318
17	2006	10.6332994467496	0.31069778699861
18	2007	10.9899398340249	0.28091390041493
19	2008	10.6851791147994	0.27875414937759
20	2009	10.4615373443984	0.29079495159059
21	2010	10.6866614799447	0.29522337482710
22	2011	10.5007320193638	0.30258333333333
23	2012	10.7555881742739	0.43144260027662
24	2013	11.7711151911469	0.55289436619718

```
SELECT cg."Year", avg(cg."Avg_temp") AS
Avg_temp,avg(cg."Avg_temp_Uncer") AS
Avg_temp_Uncer
INTO climate_temp
FROM cleanglobal_temp AS cg
Group BY cg."Year"
Order by cg."Year"ASC;
```

Global_Climate table

```
SELECT
ct."Year",ct.avg_temp,ct
.avg_temp_uncer,
co."China",co."United
States",co."India",
co."Indonesia",co."Japa
n",co."Russia",co."Brazi
l",co."Germany",co."Ira
n",co."Canada"
INTO global_climate
FROM climate_temp AS
ct
LEFT JOIN coemissions
As co ON co."Year" =
ct."Year";
```

	Year bigint	avg_temp double precision	avg_temp_uncer double precision	China double precision	United States double precision	India double precision	Indonesia double precision	Japan double precision	Russia double precision	Brazil double precision	Germany double precision	Iran double precision	Canada double precision
1	1990	10.5799118257261	0.29483921161825	1823.96	4426.4	341.32	832.39	1016.89	1790.34	1236.73	904.38	158.83	483.99
2	1991	10.3948243430152	0.26655290456431	1952.78	4389.5	386.17	846.23	1029.99	1766.89	1243.65	881.44	197.26	477.81
3	1992	9.94019536652836	0.31140041493775	2068.77	4461.62	409.09	855.3	1041.79	1630.98	1247.13	841.83	212.99	490.88
4	1993	9.76020401106502	0.28625587828492	2294.12	4581.76	431.31	869.88	1033.99	1447.33	1253.64	836.08	210.86	487.56
5	1994	10.0674706085754	0.28581466113416	2414.5	4654.52	466.79	881.31	1087.22	1238.41	1261.19	824.87	240.77	503.69
6	1995	10.656667704011	0.27156811894882	2735.48	4708.31	519.98	909.21	1097.1	1190.39	1279.97	822.15	251.99	514.62
7	1996	9.81872302904563	0.27696092669432	2715.5	4864.46	555.6	772.38	1110.51	1159.83	1301.37	851.78	259.83	528.32
8	1997	10.0640255878285	0.29461721991701	2779.27	5129.29	600.38	1484.56	1098.26	1052.5	1319.99	819.93	271	543.72
9	1998	10.592178769018	0.27458195020746	2882.75	5172.06	618.73	907.96	1055.5	1043.96	1329.24	812.12	272.42	552.1
10	1999	10.63740802213	0.27794847856154	2799.84	5191.66	683	849.42	1090.14	1081.87	1339.1	782.33	306	560.63
11	2000	10.4269713001383	0.29134439834024	2994.73	5357.88	719.07	795.62	1107.18	1115.29	1351.16	778.97	320.34	580.65
12	2001	10.5575892116183	0.28895954356846	3194.5	5347.79	958.86	631.16	1073.16	721.85	1420.6	829.38	306.74	808.34
13	2002	10.6417130013831	0.29070089903181	3476.08	5192.68	992.51	1001.39	1109.5	723.05	1418.33	815.01	324.45	826.9
14	2003	10.440566735823	0.32096887966805	4081.05	5258.12	1020.87	761.93	1118.29	753.44	1410.4	818.31	346.42	846.79
15	2004	10.518698824343	0.29718948824343	4789.59	5338.46	1092.16	1031.49	1113.6	765.36	1427.88	802.5	375.44	838.71
16	2005	10.7185425311203	0.27509508990318	5486.88	5354.45	1142.71	829.93	1116.25	762.04	1431.63	783.85	406.93	851.37
17	2006	10.6332994467496	0.31069778699861	6099.67	5254.87	1219.23	1152.83	1092.63	821.2	1435.1	795.83	446.5	842.42
18	2007	10.9899398340249	0.28091390041493	6655.98	5338.52	1341.07	738.57	1128.57	822.52	1452.63	765.28	479.47	872.59
19	2008	10.6851791147994	0.27875414937759	6862.78	5161.51	1428.85	731.76	1061.73	844.09	1473.28	771.22	489.82	851.45
20	2009	10.4615373443984	0.29079495159059	7382.89	4757.65	1573.51	1036.71	1003.19	726.06	1451.95	716.32	507.16	822.04
21	2010	10.6866614799447	0.29522337482710	8138.34	4990.96	1670.29	773.92	1057.96	814.11	1498.84	754.58	506.63	836.28
22	2011	10.5007320193638	0.30258333333333	8568.09	4753.12	1695.97	1285.69	1162.67	974.12	679.1	717.17	561.73	682.88
23	2012	10.7555881742739	0.43144260027662	8823.05	4531.16	1843.74	1303.87	1202.66	979.37	713.75	730.81	571.12	682.31
24	2013	11.7711151911469	0.55289436619718	9226.51	4670.34	1901.98	1250.27	1211.27	936.52	744.85	748.42	592.48	691.53



Global Emissions

China and the United States has the most CO2 emissions.

Global Emissions

China and the United States has the most CO2 emissions.

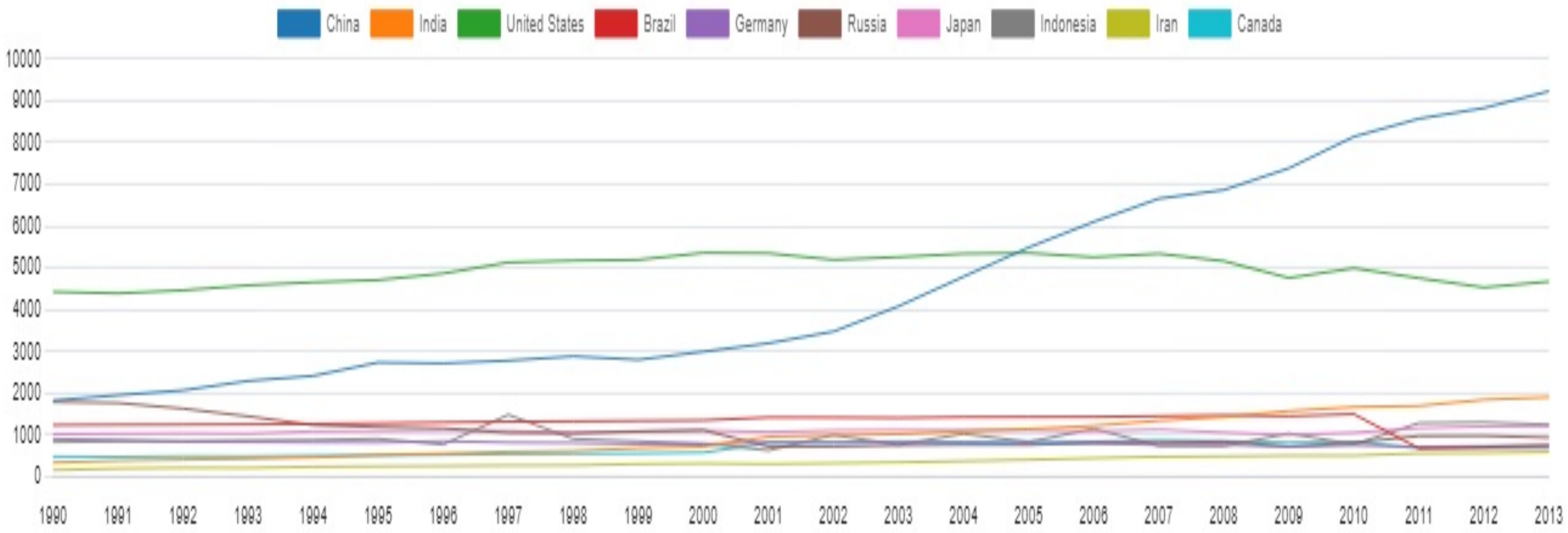
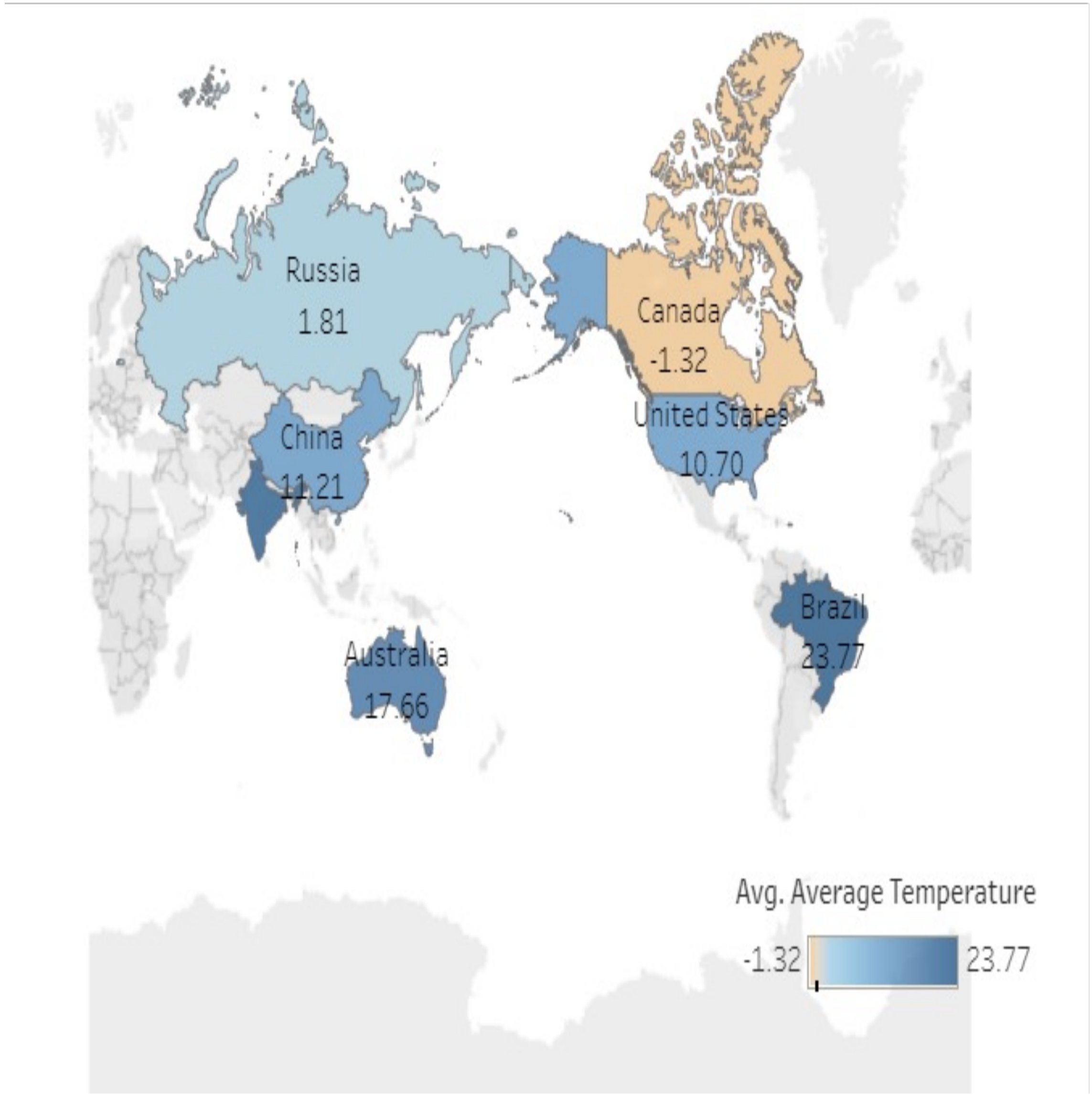


Tableau Public



RESULTS

- https://public.tableau.com/views/NetflixBestMovies/ClimateChangestory?:language=en-US&:display_count=n&:origin=viz_share_link

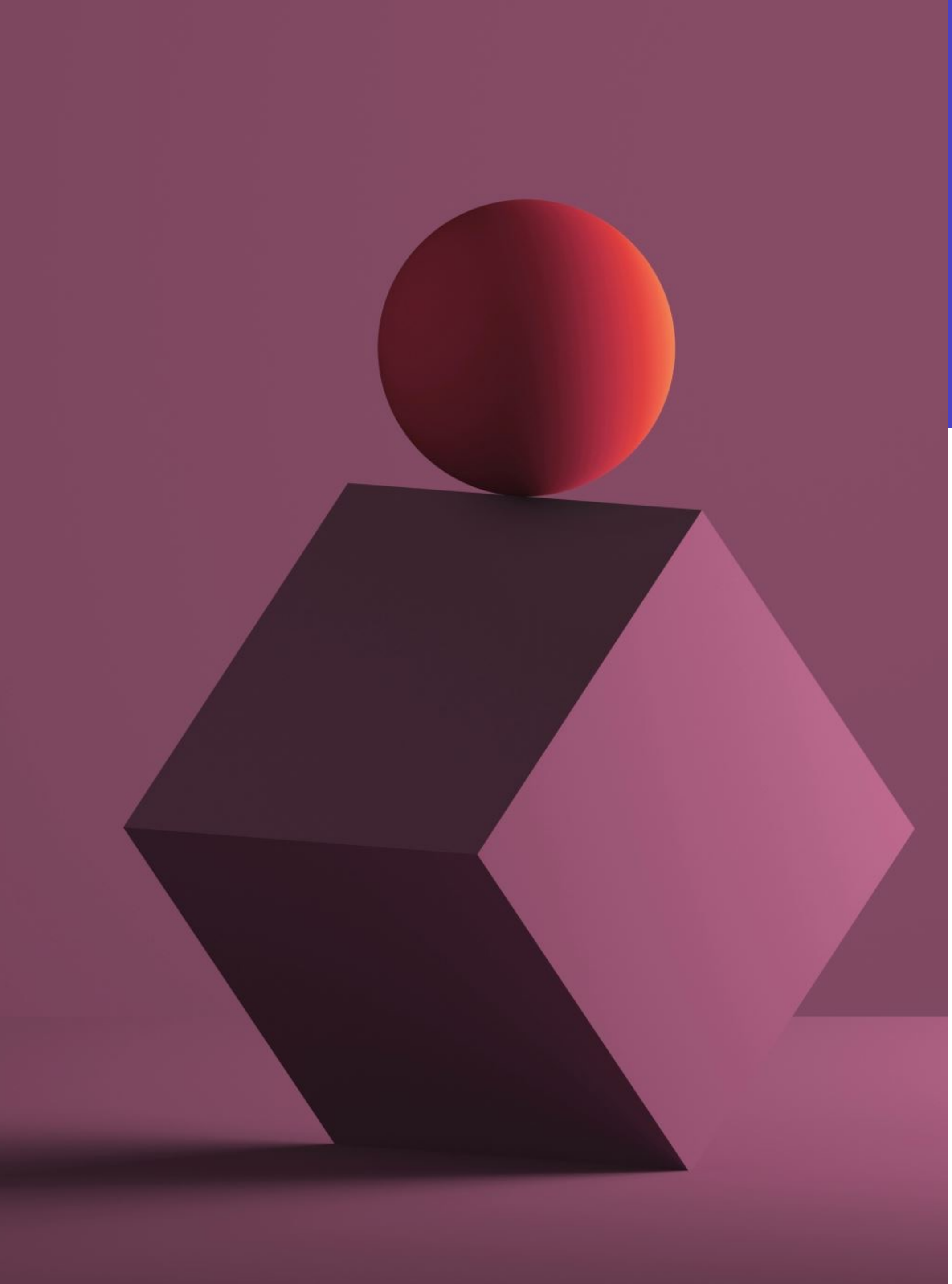
Machine Learning

Decision Tree Model (DT)
Random Forest Model (RF)

A mix of classification and and regression techniques to address imbalance data.

BENEFITS

- Robust against overfitting
- Rank importance of input variables
- Robust to outliers & non -linear data
- Efficient for large datasets



PROCESS

- Extract cleaned, renamed & formatted data
- Preprocessing
- Split the data into training & testing
- Instantiate model
- Fit the model
- Making predictions using the testing data
- Calculate the balanced accuracy score

RESULTS

Decision Tree

- Accuracy Score – 68%
- Classification Report

	Precision	Recall	F1Score
0	0.64	0.64	0.64
1	0.72	0.72	0.72

- Confusion Matrix

	Predicted0	Predicted1
Actual0	23024	12983
Actual1	13052	33137

Random Forest

- Accuracy Score – 71%.
- Classification Report

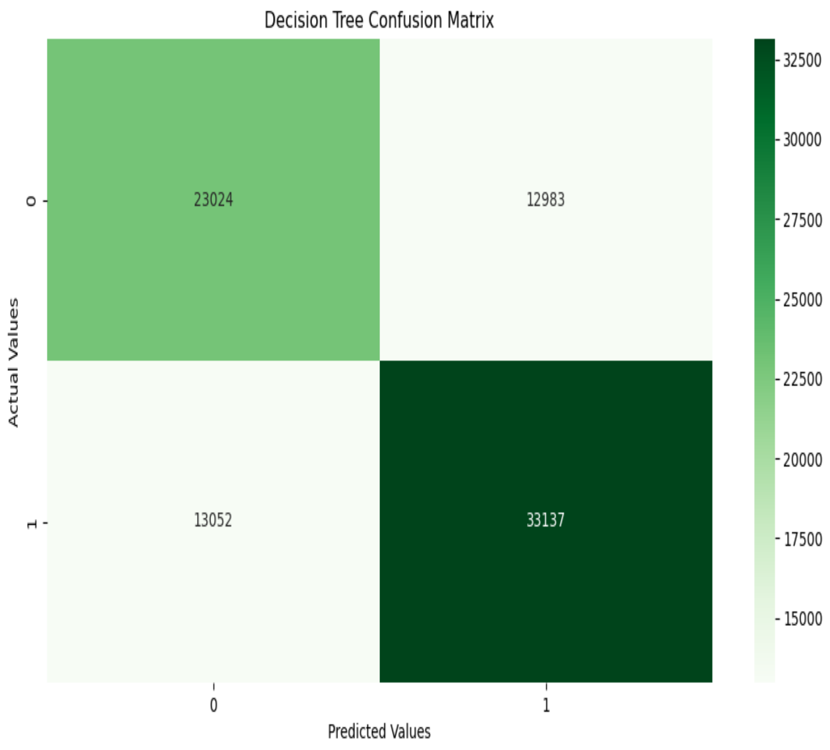
Precision Recall F1 Score 0
0.66 0.70 0.68 1
0.75 0.72 0.73

- Confusion Matrix

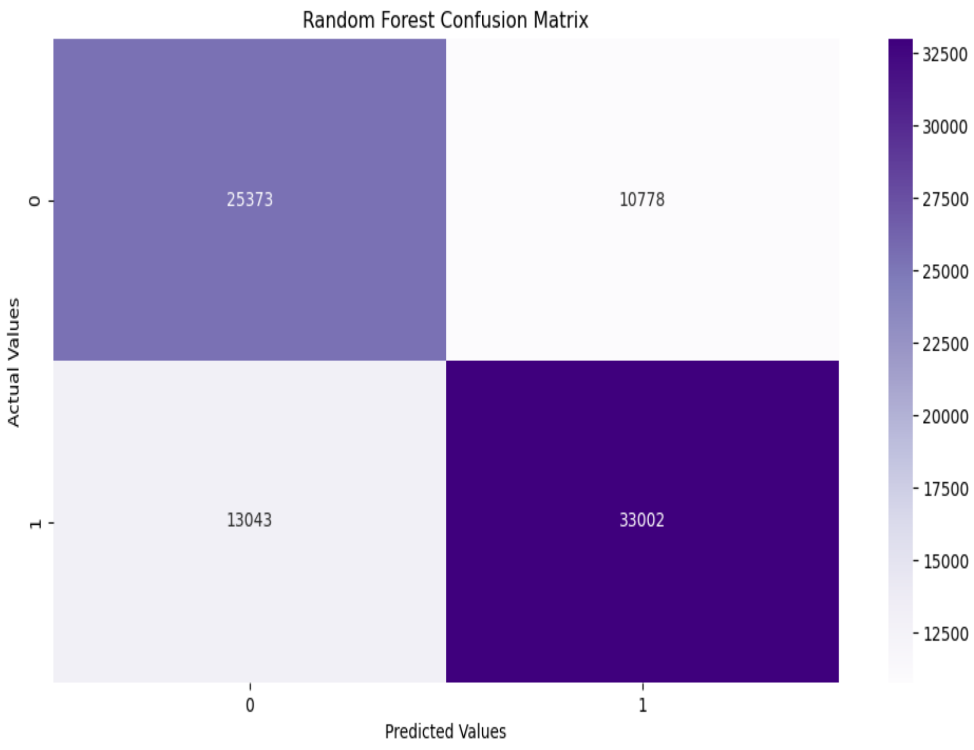
	Predicted0	Predicted1
Actual 0	25373	10778
Actual1	13043	33002

RESULTS

Decision Tree



Random Forest

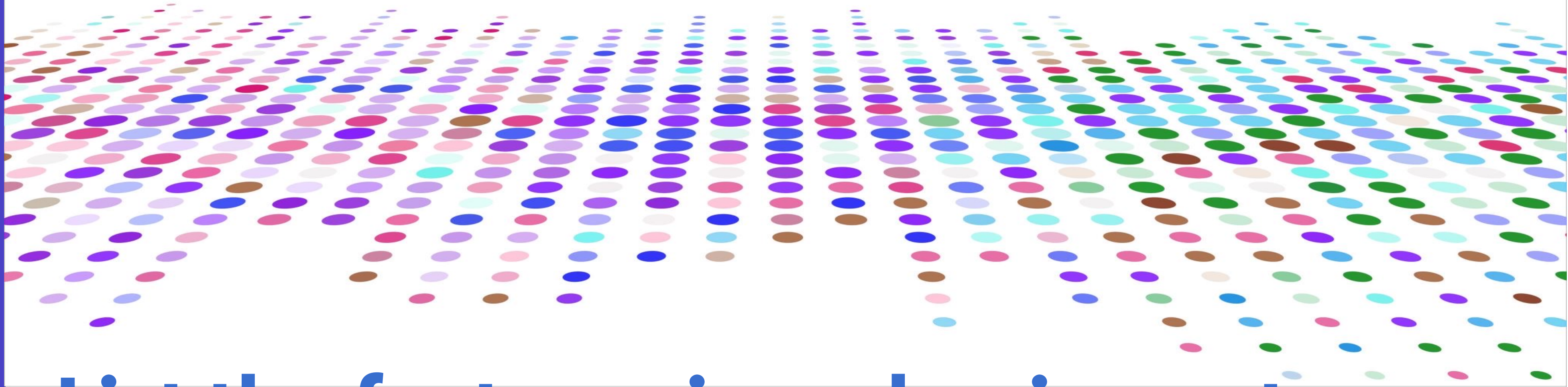


Random

Forest

ML

Results



List the features in order importance

- AverageTemperatureUncertainty
- (0.36085115372033477)
- Year: (0.21958864244658743)
- Country: (0.20989706071406417)
- State: (0.20966314311901363)

SUMMARY

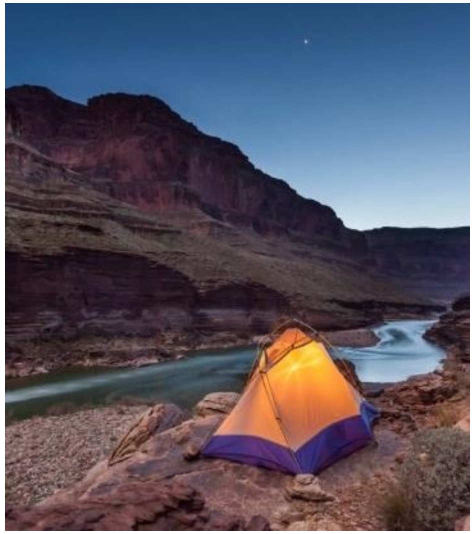
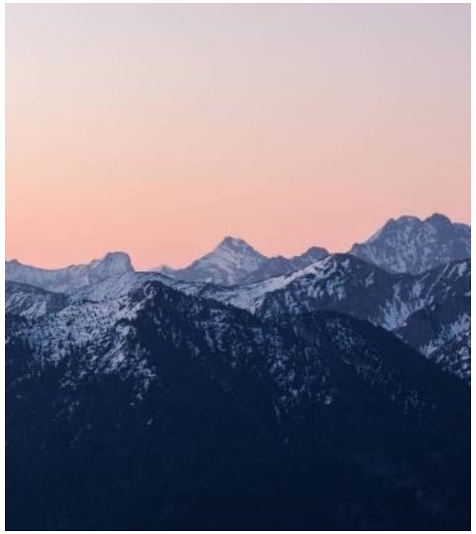
- The world avg temperature rose by 3.3point from 1900-2013.

- China and USA are the major contributors of CO_2 emissions.

Recommendations

- Doing the analysis with better temperature data

- Doing Neural Network Model

RESOURCES

- Berkeley Earth, affiliated Lawrence Berkeley National Laboratory
- Climate Watch Data