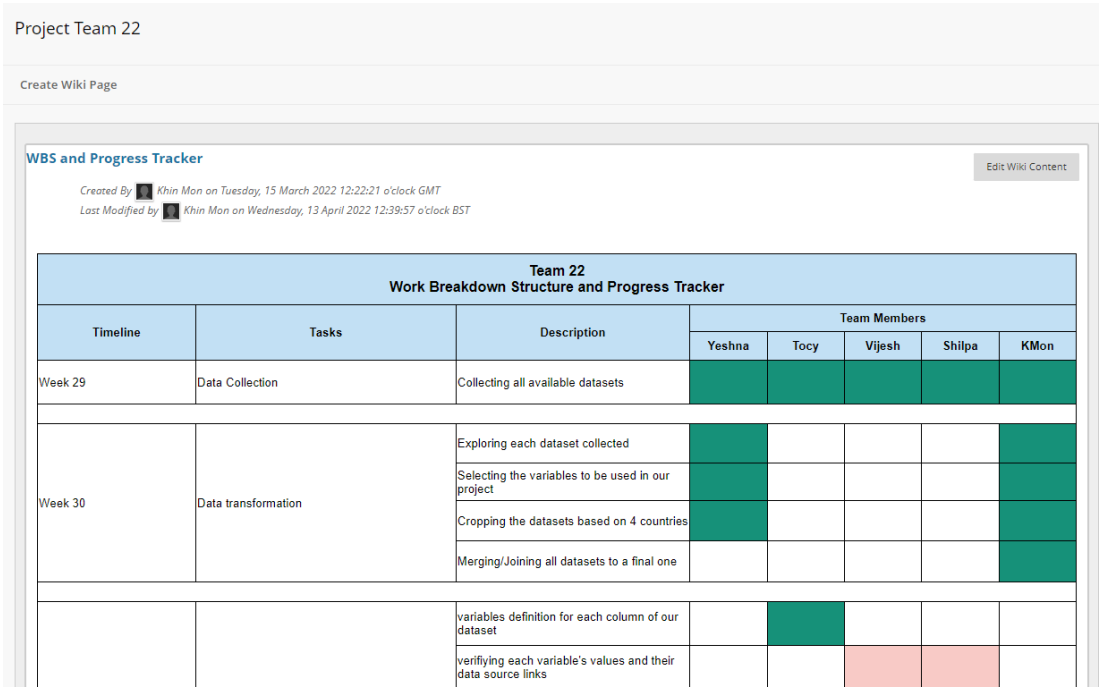
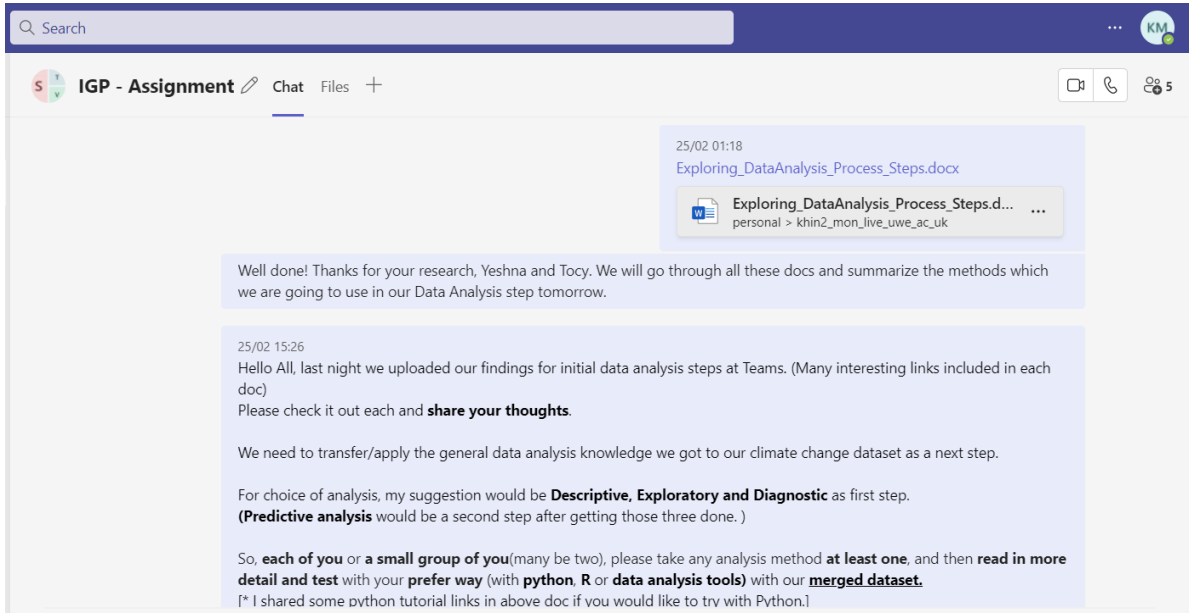




PROJECT MANAGEMENT AND TEAM COMMUNICATION

Skill sets used - python, R programming, Machine Learning algorithms and data science tools (Tableau)

About the project

- Type – regression problem(multiple LR)

Analyse the specific relationships between the independent variables and the dependent one. [credit to ML lecture note]

Predict the value of a dependent variable from the value of independent variable [credit to ML lecture note]

What are the impacts of climate change caused by carbon dioxide emission and its consequences on financial stability such as GDP per capita?

No

Dependent variable – Gdp per capita
Independent variable – Co2 and others

Micro detail ..
Impact of Co2 on GDP per capita in USA?
Yes/no/may be

Impact of Co2 on GDP per capita in India?
Yes/no/may be

Correlation ?
gas, oil, coal => Co2 ....impacts on ... Gdp => consumption co2, consumption co2 per capita

DATASET

#	Variables	Type	Units	File Name	From Year	To Year	Data Source Links	Website	Last Updated Date
1	co2		kiloton	API_EN.ATM.CO2E.KT_DS2_en_csv_v2_3731329.csv	1960	2018	https://data.worldbank.org/indicator/EN.ATM.CO2E.KT	WorldBank	2/15/2022
2	co2_per_capita	Per capita	metric tons per capita	API_EN.ATM.CO2E.PC_DS2_en_csv_v2_3731558.csv	1960	2018	https://data.worldbank.org/indicator/EN.ATM.CO2E.PC	WorldBank	2/15/2022
3	consumption_co2	Annual	tonnes(billion t)	consumption-co2-emissions.csv	1990	2019	https://ourworldindata.org/grapher/consumption-co2-emissions?tab=chart&time=earliest..latest&country=IND~	Ourworldindata	
4	consumption_co2_per_capita	Per capita	tonnes per person per year	consumption-co2-per-capita.csv	1990	2019	https://ourworldindata.org/grapher/consumption-co2-per-capita?tab=chart&country=IND~JPN~RUS~USA	Ourworldindata	
5	cumulative_co2	Cumulative	tonnes(billion t)	cumulative-co-emissions.csv	1960	2020	https://ourworldindata.org/grapher/cumulative-co-emissions?tab=chart&time=1960..2020&country=IND~JPN~F	Ourworldindata	
6	coal_co2	Annual	tonnes(million/billion t)	annual-co2-coal.csv	1960	2020	https://ourworldindata.org/grapher/annual-co2-coal?tab=chart&time=1960..latest&country=IND~JPN~RUS~USA	Ourworldindata	
7	cement_co2	Annual	tonnes(million/billion t)	annual-co2-cement.csv	1960	2020	https://ourworldindata.org/grapher/annual-co2-cement?tab=chart&time=1960..latest&country=IND~JPN~RUS~	Ourworldindata	
8	flaring_co2	Annual	tonnes(million/billion t)	annual-co2-flaring.csv	1960	2020	https://ourworldindata.org/grapher/annual-co2-flaring?tab=chart&time=1960..latest&country=IND~JPN~RUS~U	Ourworldindata	
9	gas_co2	Annual	tonnes(million/billion t)	annual-co2-gas.csv	1960	2020	https://ourworldindata.org/grapher/annual-co2-gas?tab=chart&time=1960..latest&country=IND~JPN~RUS~USA	Ourworldindata	
10	oil_co2	Annual	tonnes(million/billion t)	annual-co2-oil.csv	1960	2020	https://ourworldindata.org/grapher/annual-co2-oil?tab=chart&time=1960..latest&country=IND~JPN~RUS~USA	Ourworldindata	
11	coal_co2_per_capita	Per capita	tonnes per person per year	per-capita-co2-coal.csv	1960	2020	https://ourworldindata.org/grapher/per-capita-co2-coal?tab=chart&time=1960..latest&country=IND~JPN~RUS~	Ourworldindata	
12	cement_co2_per_capita	Per capita	tonnes per person per year	per-capita-co2-cement.csv	1960	2020	https://ourworldindata.org/grapher/per-capita-co2-cement?tab=chart&time=1960..latest&country=IND~JPN~R	Ourworldindata	
13	flaring_co2_per_capita	Per capita	tonnes per person per year	per-capita-co2-flaring.csv	1960	2020	https://ourworldindata.org/grapher/per-capita-co2-flaring?tab=chart&time=1960..latest&country=IND~JPN~RU	Ourworldindata	
14	gas_co2_per_capita	Per capita	tonnes per person per year	per-capita-co2-gas.csv	1960	2020	https://ourworldindata.org/grapher/per-capita-co2-gas?tab=chart&time=1960..latest&country=IND~JPN~RUS~	Ourworldindata	
15	oil_co2_per_capita	Per capita	tonnes per person per year	per-capita-co2-oil.csv	1960	2020	https://ourworldindata.org/grapher/per-capita-co2-oil?tab=chart&time=1960..latest&country=IND~JPN~RUS~U	Ourworldindata	
16	cumulative_coal_co2	Cumulative	tonnes(billion t)	cumulative-co2-coal.csv	1960	2020	https://ourworldindata.org/grapher/cumulative-co2-coal?tab=chart&time=1960..latest&country=IND~JPN~RUS	Ourworldindata	
17	cumulative_cement_co2	Cumulative	tonnes(million/billion t)	cumulative-co2-cement.csv	1960	2020	https://ourworldindata.org/grapher/cumulative-co2-cement?tab=chart&time=1960..latest&country=IND~JPN~F	Ourworldindata	
18	cumulative_flaring_co2	Cumulative	tonnes(million/billion t)	cumulative-co2-flaring.csv	1960	2020	https://ourworldindata.org/grapher/cumulative-co2-flaring?tab=chart&time=1960..latest&country=IND~JPN~RU	Ourworldindata	
19	cumulative_gas_co2	Cumulative	tonnes(billion t)	cumulative-co2-gas.csv	1960	2020	https://ourworldindata.org/grapher/cumulative-co2-gas?tab=chart&time=1960..latest&country=IND~JPN~RUS~	Ourworldindata	
20	cumulative_oil_co2	Cumulative	tonnes(billion t)	cumulative-co2-oil.csv	1960	2020	https://ourworldindata.org/grapher/cumulative-co2-oil?tab=chart&time=1960..latest&country=IND~JPN~RUS~U	Ourworldindata	
21	gdp (current US$)		US$	API_NY.GDP.MKTP.CD_DS2_en_csv_v2_3731268.csv	1960	2020	https://data.worldbank.org/indicator/Ny.Gdp.Mktp.Cd	WorldBank	2/15/2022
22	co2_per_gdp (production-based)	Annual	(Intl $)	co2.csv	1960	2018	https://ourworldindata.org/explorers/co2?time=1960..latest&facet=none&country=USA~IND~RUS~JPN&Gas=C0	Ourworldindata	
23	consumption_co2_per_gdp	Annual	GDP(Intl $)	consumption-based-carbon-intensity.csv	1990	2018	https://ourworldindata.org/grapher/consumption-based-carbon-intensity?tab=chart&country=IND~JPN~RUS~US	Ourworldindata	
24	US$))	Per capita	US$	API_NY.GDP.PCAP.CD_DS2_en_csv_v2_3731360.csv	1960	2020	https://data.worldbank.org/indicator/NY.GDP.PCAP.CD	WorldBank	2/15/2022
25	annual_co2_fossil_cement	Annual	tonnes(billion t)	annual-co2-emissions-per-country.csv	1960	2020	https://ourworldindata.org/grapher/annual-co2-emissions-per-country?tab=chart&time=1960..latest&country=I	Ourworldindata	
	Country, Year + 25 variables								

LITERATURE REVIEW

We did literature review on 20 different lectures.

All data are gathered from trustful sources such as Ourworldindata and Worldbank.
Each dataset has been verified for the data integrity
Data are filtered based on 4 different countries.
We analysis data based on 60 years historical data.

About dataset

Total – 27 variables

Used (4) - no outliers

- Country
- Year
- Co2
- Gdp_per_capita

Less Used (5)

For Co2,

- Coal_co2
- Gas_co2
- Oil_co2

For Gdp per capita,

- Consumption_co2
- Consumption_co2_per_capita

*** If feature engineering/selection with Ordinary least squares algorithm is done in advance, Consumption_co2 and Consumption_co2_per_capita might not be included for further analysis process.

Not Used (18)

- ~~Co2_per_capita 8~~
- ~~Gdp 28~~
- ~~Cement_co2 0~~
- ~~Flaring_co2 27~~
- ~~Flaring_co2_per_capita 27~~
- ~~Coal_co2_per_capita 0~~
- ~~Cement_co2_per_capita 0~~
- ~~Gas_co2_per_capita 1~~
- ~~Oil_co2_per_capita 0~~
- ~~Cumulative_co2 0~~
- ~~Cumulative_coal_co2 0~~
- ~~Cumulative_cement_co2 0~~
- ~~Cumulative_flaring_co2 27~~
- ~~Cumulative_gas_co2 1~~
- ~~Cumulative_oil_co2 0~~
- ~~Consumption_co2_per_gdp 128~~
- ~~Annual_co2_fossil_cement 0~~
- ~~Co2_per_gdp 8~~

Variables	Outliers	Null
Country	0	0
Year	0	0
Co2	0	8
Gdp_per_capita	0	28
Coal_co2	0	0
Gas_co2	0	1
Oil_co2	50	0
Consumption_co2	16	124
Consumption_co2_per_capita	0	124

Timeframe

1960 ??? -2020

Missing values

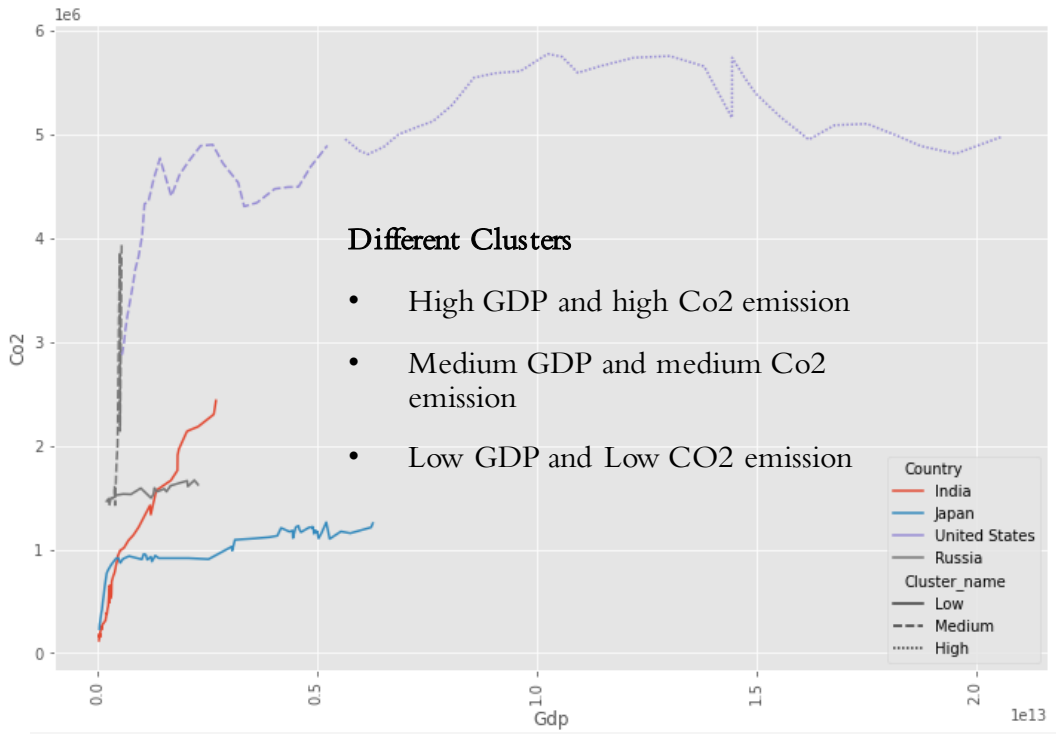
Imputed or not?

Outliers?

Removed or not

All Data Analysis Steps	Why is it required?
Descriptive analysis	basic exploration
Provide Basic Descriptions	basic exploration
Missing value imputation	
Identify Significant Correlations	co2 => oil, gas, coal .....gdp pc => consumption co2, co2 pc
Find relationships between variables	double confirmed!
Spot Outliers in the Dataset	any outliers?
Identifying of anomalies	extreme/abnormal values in study variables?
Hypothesis testing	observations are within the avg range?
Correlation and causation	GDP per captica's changes(ups and downs) is correlated with Co2 emission?
Regression analysis	how much talk about co2 by gas, oil and coal ....Tell about gdp pc by consumption co2 and co2 pc ...
Clustering analysis	unsupervised ML method, identifying and grouping similar data points ... To be easily understand and manipulate data
Filtering	to know the most important variables in the dataset which can tell much information about project question .. consumption variables were not selected ..
Outliers Analysis	to know how to handle outliers in dataset ... transformed data for regression?
Time-series analysis	after diagnostic, we answer the research question... then we want to know what will happen in future based on current data ..
~~Probability theory~~	

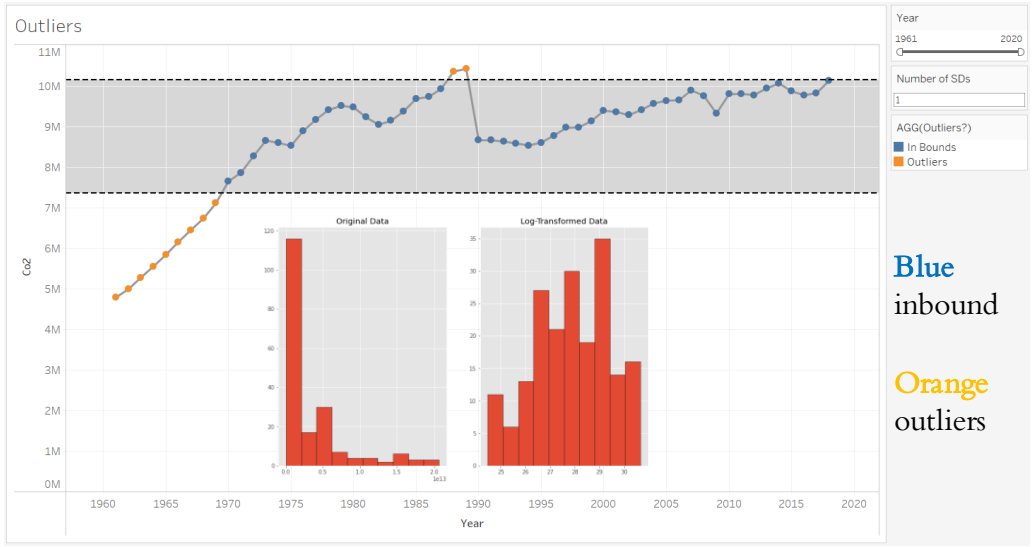
Cluster Analysis Source Code



Different Clusters

- High GDP and high Co2 emission
- Medium GDP and medium Co2 emission
- Low GDP and Low CO2 emission

Outlier Analysis Source Code



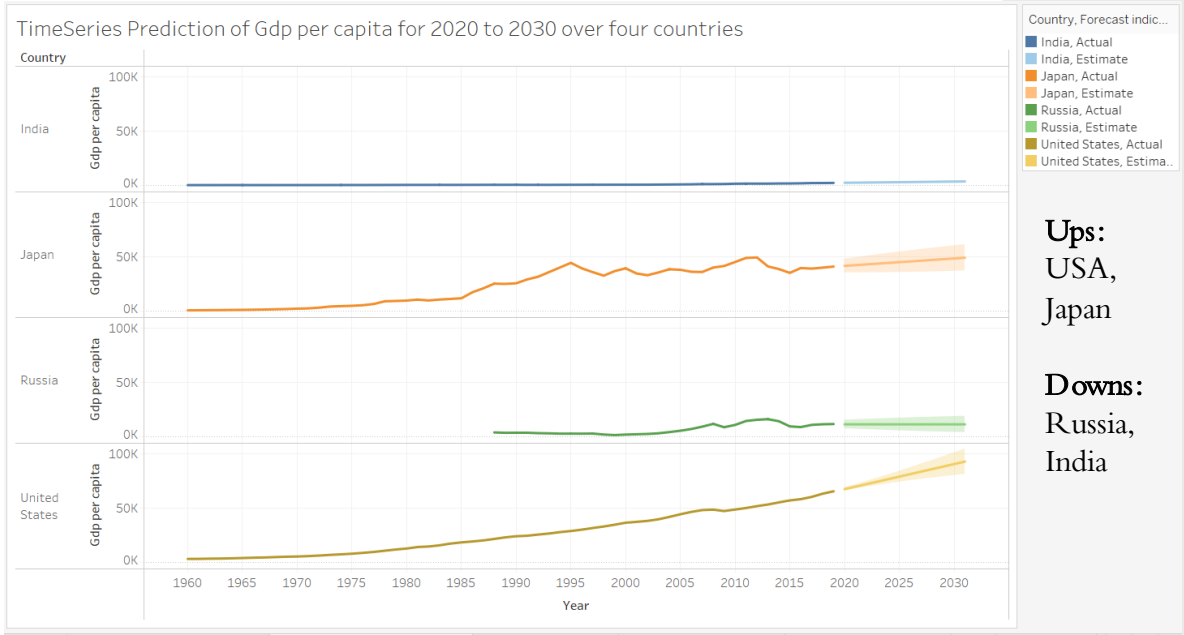
Blue inbound

Orange outliers

Solution for outliers: Log transformation?

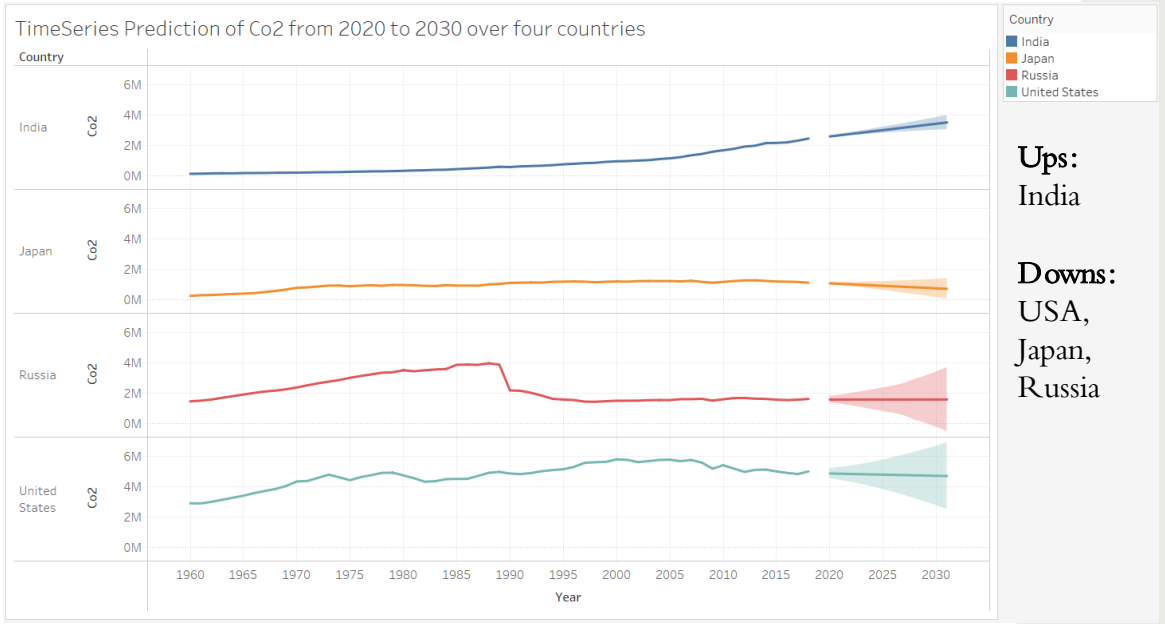
What will be the impacts of Co2 emission on GDP per capita?

Predictive Analysis / Forecasting Source Code



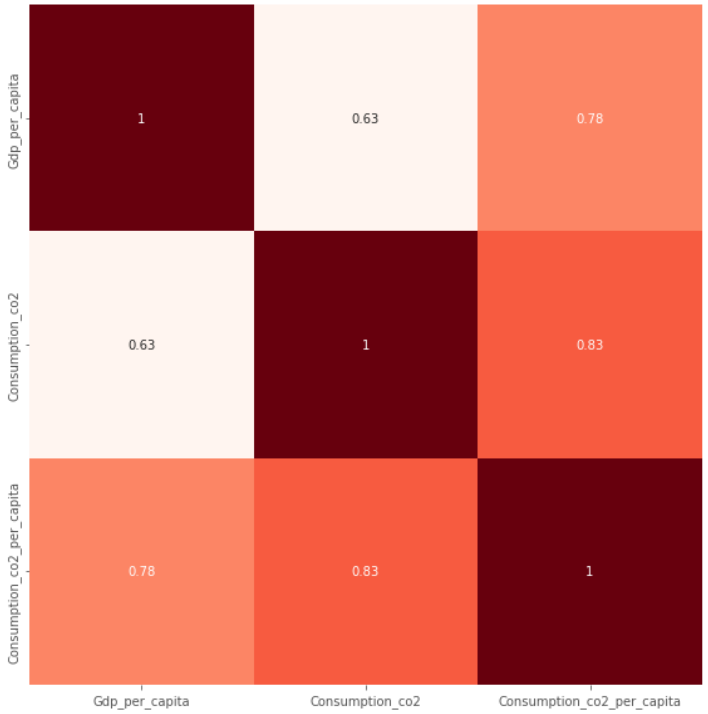
Ups: USA, Japan

Downs: Russia, India



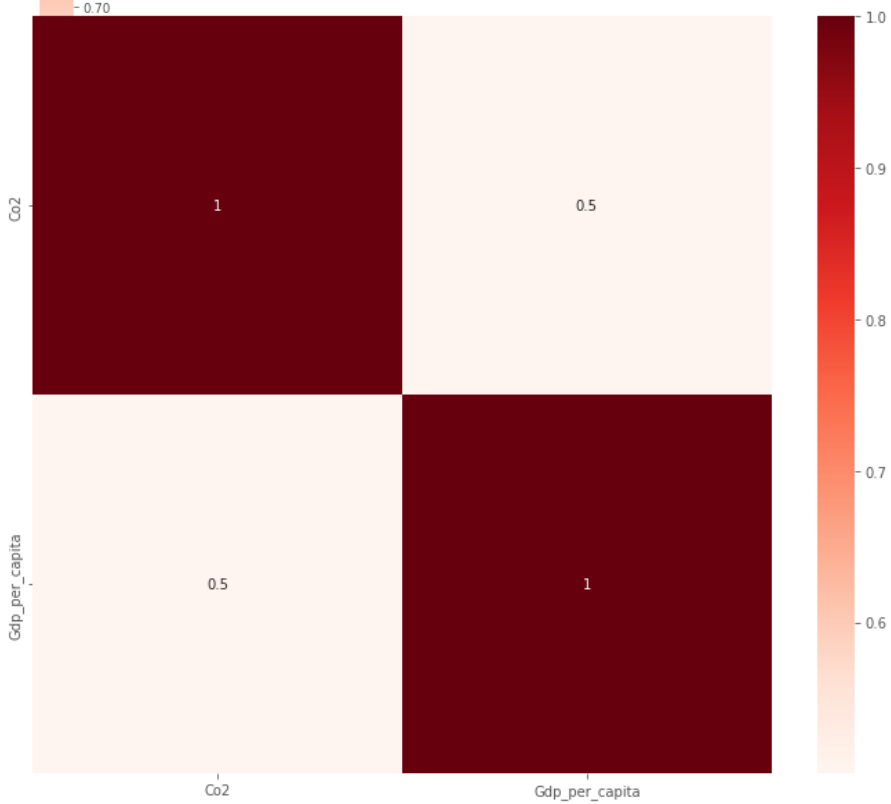
Ups: India

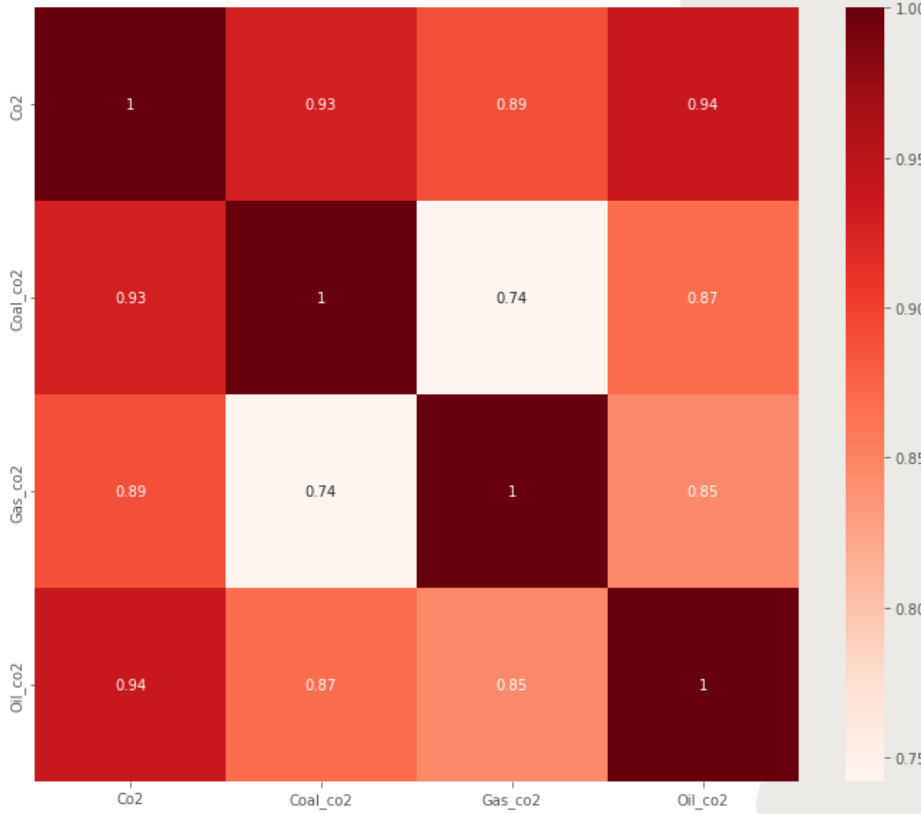
Downs: USA, Japan, Russia



Correlation coefficient, Co2 vs 2 other variables

Correlation coefficient, Co2 vs Gdp pc





Correlation coefficient, Gdp pc vs 3 other variables

Evaluation: LR
More correlated, higher R2?

DATA ANALYSIS RESULTS

In the past, did carbon dioxide emission and its consequences impact on GDP per capita?

No strong evidence

If yes, what are the impacts?
Agriculture suffers losses due to excessive CO2 and temperature…so the impacts is ***less GDP per capita***

CONCLUSION

- Based on the data, we found that there were high Co2 emissions in the developing countries, especially in UAS, compared with the developing countries.
- In USA, after reaching to its peak in 2000, the amount of Co2 emission dropped down significantly till 2020, it was one of the highest Co2 emission countries in the past 1960. It could be because of the government good policy, people awareness on environmental issues and technological advancements which purposely reduce the Co2 emissions. Based on the 60 years historical data of Co2 emission, when we predict the Co2 emission for 2030 with the help of Machine Learning algorithms, the statistical figures show us the reduced amount of Co2 emission.
- When we study the Co2 emission and its consequences on GDP, we did not see any direct relationship between those two variables statistically. But the developed with higher income looks getting aware of the risks of Co2 emission and they are apparently trying to reduce the amount of Co2 emission year after year. Compared it with low-income countries, the amount of Co2 emission are increased gradually and based on our predictions, it will increase in coming years.