Programming with Data

Introduction to research space

Aims and Objective:

To study Carbon emissions in Singapore and how it has impacted the climate in Singapore.

Research Summary:

Climate change is a crucial topic to be discussed. In recent years, many countries are actively discussing and finding ways to battle this global crisis. In the 27th session of the Conference of the Parties to the United Nations climate convention, COP27 that took place, the United Nations Secretary General, Antonio Guterres has reiterated how we are on an accelerated path to "Climate Hell ". ('A Reason to Act Faster': World Leaders Meet on Climate Amid Other Crises, 2022)

Climate change is a very relevant topic in Singapore. Singapore has been experiencing extreme weathers in the last decades. There is an impending threat of extreme rise in sea levels. Singapore being a coastal city surreounded by the sea in all directions, it is an ever growing threat. It is extremely pivotal that we study the carbon emissions in Singapore and see how it affects the weather events in Singapore. (Mse.gov.sg, 2018)

In Singapore, climate change has been a key pillar that the government and policy makers primarily focus on. Across the various years, Singapore has played its part in managing climate change. In the COP27, Singapore along with many other countries have pledged to reduce greenhouse gas emissions from fossil fuels.

In April 2022, the government launched the Singapore Green Plan 2030 where it has roped in more than 1,500 public members and stakeholders to work on Singapore's climate ambition. By 2050, Singapore aims to achieve net zero emissions. This shows that climate change is a relevant and important topic for Singapore. (Nccs.gov.sg, 2015)

Relevance of this Field:

In the context of Singapore pledging to reduce greenhouse gas emissions from fossil fuels, I wanted to conduct research to study how Carbon Dioxide emissions, a key greenhouse gas, has impacted Singapore's climate over the last decade. There are many factors that can be used to measure and study climate change. I have picked the factor of rainfall. Precipitation rates are a common dataset that climate activists and researchers use to see the changes in climates. Even though, theoretically, it is known that increased emissions will lead to increased precipitation, I want to study if the emissions have decreased in Singapore and if that have lead to low rainfall.

Singapore has involved itself regularly in climate change protocols over the years, naming:

- Kyoto Protocol in 2015
- UN Climate Change Conference in Doha, 2012
- The Paris Agreement in COP-12 Paris
- United Nations Framework Convention on Climate Change, 2021

Taking all these key events into account, the study will focus on the effects of Carbon Emissions on Singapore's Rainfall amount.

Hypothesis to prove:

I have chosen two hypothesis to prove in this report:

1. Increase in carbon dioxide emission in Singapore has led to increased rainfalls.
2. Rainfall is a good metric that can be used to study climate change.

Key Questions to address

1. Will Singapore taking part in Climate Change protocols cause a drop in carbon emissions in the last decade?
2. What affects rainfall in Singapore other than Carbon emissions?
3. Is Rainfall a good metric to study climate change trends?

Scope of Work:

To conduct this study, I will be working on two datasets, Global Carbon Dioxide Emissions and Singapore Climate Details. I will be extracting the carbon dioxide emissions of Singapore using from the global dataset. From the Singapore Climate Details dataset, I will be extracting the rainfall details. Using these 2 extracted datasets, I will compare them to study the trends across the years from 2022-2022. I will also be looking at the key events stated above in the "Relevance of this Field" section to see if there are any improvements after the protocols.

Data Processing Plan:

Below are the break down on how the two datasets will be processed

First Dataset

- Step 1: Importing the first dataset on global Carbon dioxide emissions.
- Step 2: Cleaning the dataset by removing unwanted field
- Step 3: Extract the Dataset for Singapore from the global dataset and storing it in a new table

Second Dataset

- Step 1: Importing the dataset on Singapore Weather
- Step 2: Cleaning the dataset by removing unwanted field
- Step 3: Extracting Total Rainfall and Number of Rainy Days data and storing it in a new annual Singapore weather dataset.
- Step 4: Merging the first and second dataset into one dataset
- Step 5: Plotting and Analysis of the Data

Data Visualisation

1. Analysing carbon dioxide emissions in Singapore across the last decade to study the trend in the release of emissions
2. Analysing total rainfall in Singapore across the last decade to study the trend in the rainfall amounts
3. Analysing average number of rainy days across the years in Singapore to study the treidn in the rainfall patterns
4. Analysing the relationship between the carbon dioxide emissions and rainfall across the last decade in Singapore

First dataset on finding carbon dioxide emissions in Singapore

Why this Dataset?

This data set contains all the Carbon Emissions from countries all around the world from the year 1750 to 2020. With columns ranging from code (country code) to the total carbon emissions in the individual years, this dataset allows me to extract the dataset for a specific country, in my case, Singapore to study the trends of its emissions across the various years. This dataset is update annually with the carbon emissions and it contains the period of years that I want to study in this report.

This data is sufficiently big with a huge dataset across multiple decades. This dataset also has Singapore data in it with no missing data of any years. This allows me to generate a complete and fair data especially when comparing with other datasets.

Dataset Details

I have acquired this dataset from Kaggle and it is sourced from url={https://www.kaggle.com/dsv/4202243} (https://www.kaggle.com/dsv/4202243%7D). I downloaded it from the site as an excel file for data analysis on Jupyter Notebook.

Importing Pandas, Numpy and Matplotlib

In [79]:

```
#Will be using pandas for data cleansing, merges and joins and for statistical analysis
import pandas as pd

#will be using numpy to do mathematical functions
import numpy as np

#will be using matplot to plot my dataset for data visualisation
import matplotlib.pyplot as plt
%matplotlib inline
```

In [80]:

```
#Importing the Dataset
Carbon_emissions_df = pd.read_excel("Carbon Emissions.xlsx")

#Display the header and tail for this dataset
display(Carbon_emissions_df.head())
display(Carbon_emissions_df.tail())
```

	Country	Code	Calling Code	Year	CO2 emission (Tons)	Population(2022)	Area	% of World	Density(km2)
0	Afghanistan	AF	93	1750	0.0	41128771.0	652230.0	0.40%	63/km²
1	Afghanistan	AF	93	1751	0.0	41128771.0	652230.0	0.40%	63/km²
2	Afghanistan	AF	93	1752	0.0	41128771.0	652230.0	0.40%	63/km²
3	Afghanistan	AF	93	1753	0.0	41128771.0	652230.0	0.40%	63/km²
4	Afghanistan	AF	93	1754	0.0	41128771.0	652230.0	0.40%	63/km²

	Country	Code	Calling Code	Year	CO2 emission (Tons)	Population(2022)	Area	% of World	Density(km2)
59615	Zimbabwe	ZW	263	2016	736467042.0	16320537.0	390757.0	0.30%	42/km²
59616	Zimbabwe	ZW	263	2017	746048675.0	16320537.0	390757.0	0.30%	42/km²
59617	Zimbabwe	ZW	263	2018	757903042.0	16320537.0	390757.0	0.30%	42/km²
59618	Zimbabwe	ZW	263	2019	768852126.0	16320537.0	390757.0	0.30%	42/km²
59619	Zimbabwe	ZW	263	2020	779383468.0	16320537.0	390757.0	0.30%	42/km²

Cleaning the Dataset by removing the unwanted fields

I will be needing the following fields

- Country
- Code
- Calling Code
- Year
- CO2 emission (Tons)

I will not be needing the following fields.

- Population
- Area
- % of World
- Density

Hence, I will be removing these unwanted fields to clean the dataset. To do so, I will be utilising the drop function. In the drop function, I will be inserting the fields to be removed and set axis = 1 as I am removing columns.

In [81]:

```
#Using drop to remove the fields
Carbon_emissions_df.drop(["Population(2022)", "Area","% of World", "Density(km2)"], axis=1, inplace = True)

#Display the header to check if the fields have been dropped
display(Carbon_emissions_df.head())
display(Carbon_emissions_df.tail())
```

	Country	Code	Calling Code	Year	CO2 emission (Tons)
0	Afghanistan	AF	93	1750	0.0
1	Afghanistan	AF	93	1751	0.0
2	Afghanistan	AF	93	1752	0.0
3	Afghanistan	AF	93	1753	0.0
4	Afghanistan	AF	93	1754	0.0

	Country	Code	Calling Code	Year	CO2 emission (Tons)
59615	Zimbabwe	ZW	263	2016	736467042.0
59616	Zimbabwe	ZW	263	2017	746048675.0
59617	Zimbabwe	ZW	263	2018	757903042.0
59618	Zimbabwe	ZW	263	2019	768852126.0
59619	Zimbabwe	ZW	263	2020	779383468.0

Extracting Singapore Data from the Global Dataset

In this section, I will ve extracting Singapore data from the global dataset. To achieve that, I can use the country code unique to each country to extract the data of any specific country. In my case, I will be using the Code "SG" and setting a condition for the years (From year 2000) to extract the data for Singapore and store it in a new table called Singapore_Carbon_Emissions.

By setting Carbon_emissions_df["Code"] == "SG" and Carbon_emissions_df["Year"] > 1999, I will be able to extract the Singapore carbon emissions data from year 2000.

In [82]:

```
#Setting conditions and extracting Singapore Dataset and storing it in a new dataset
Singapore_Carbon_emissions = Carbon_emissions_df[(Carbon_emissions_df["Code"] == "SG") & (Carbon_emissions_df["Year"] > 1999)]

#Display the header to check if the dataset has been extracted
display(Singapore_Carbon_emissions.head())
display(Singapore_Carbon_emissions.tail())
```

	Country	Code	Calling Code	Year	CO2 emission (Tons)
48217	Singapore	SG	65	2000	1.166442e+09
48218	Singapore	SG	65	2001	1.215642e+09
48219	Singapore	SG	65	2002	1.262736e+09
48220	Singapore	SG	65	2003	1.312174e+09
48221	Singapore	SG	65	2004	1.359033e+09

	Country	Code	Calling Code	Year	CO2 emission (Tons)
48233	Singapore	SG	65	2016	1.980751e+09
48234	Singapore	SG	65	2017	2.019817e+09
48235	Singapore	SG	65	2018	2.065819e+09
48236	Singapore	SG	65	2019	2.111524e+09
48237	Singapore	SG	65	2020	2.157028e+09

View the Singapore Carbon Emissions dataset

In [83]:

```
#Display the whole dataset
print(Singapore_Carbon_emissions)

#This was done to check the final output of Singapore_Carbon_emissions dataframe
```

```
         Country Code Calling Code  Year  CO2 emission (Tons)
48217  Singapore   SG           65  2000         1.166442e+09
48218  Singapore   SG           65  2001         1.215642e+09
48219  Singapore   SG           65  2002         1.262736e+09
48220  Singapore   SG           65  2003         1.312174e+09
48221  Singapore   SG           65  2004         1.359033e+09
48222  Singapore   SG           65  2005         1.395501e+09
48223  Singapore   SG           65  2006         1.438545e+09
48224  Singapore   SG           65  2007         1.486064e+09
48225  Singapore   SG           65  2008         1.543636e+09
48226  Singapore   SG           65  2009         1.633734e+09
48227  Singapore   SG           65  2010         1.690354e+09
48228  Singapore   SG           65  2011         1.720626e+09
48229  Singapore   SG           65  2012         1.768771e+09
48230  Singapore   SG           65  2013         1.823137e+09
48231  Singapore   SG           65  2014         1.878346e+09
48232  Singapore   SG           65  2015         1.940477e+09
48233  Singapore   SG           65  2016         1.980751e+09
48234  Singapore   SG           65  2017         2.019817e+09
48235  Singapore   SG           65  2018         2.065819e+09
48236  Singapore   SG           65  2019         2.111524e+09
48237  Singapore   SG           65  2020         2.157028e+09
```

Second Dataset on Singapore Weather

Why this dataset

This dataset contains the monthly weather details of Singapore such as Air Temperature and Sunshine, Relative Humidity and Rainfall. For my report, I will be utilising the rainfall data. This dataset is from Singstat and the source is from National Environment Agency which is a reliable source to get the dataset from. (Singstat.gov.sg, 2023)

The dataset is complete with no missing fields. The dataset gives me monthly data across the years ranging from 1975 to 2022 September. The dataset is constantly updated and is well detailed with multiple fields. The dataset gives me multiple fields related to rainfall such as Total Rainfall (Millimetre), Highest Daily Rainfall Total (Millimetre) and Number Of Rainy Days (Number). This allows me to explore more deeply into the rainfall patterns in Singapore. Hence, this dataset is appropriate for the research question that I have chosen.

Dataset Details

I have downloaded the data as a CSV file from URL = {https://tablebuilder.singstat.gov.sg/table/TS/M890191} (https://tablebuilder.singstat.gov.sg/table/TS/M890191%7D).

In [84]:

```
#Importing the csv file
weather_details_df = pd.read_csv("Weather Details.csv")

#Printing the head and tail
display(weather_details_df.head())
display(weather_details_df.tail())
```

	Theme: Society	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	...	Unnamed: 564	Unnamed: 565	Unnamec 56
0	Subject: Environment	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	Nal
1	Topic: Air Quality and Climate	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	Nal
2	Table Title: Air Temperature And Sunshine, Rel...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	Nal
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	Nal
4	Data last updated: 27/10/2022	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	Nal

5 rows × 574 columns

	Theme: Society	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	...	Unnamed: 564	Unnamed: 565	I
30	Values are shown in Singapore dollars (unless ...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
31	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
32	Generated by: SingStat Table Builder	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
33	Date generated: 15/11/2022	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
34	Contact: info@singstat.gov.sg	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	

5 rows × 574 columns

Removing Header and Footers in the Dataset

The original dataset imported has headers and footers that are unwanted during the data visualisation part. I will be setting a range in the drop function to remove the rows that are not needed.

In the drop function, I can set the range of rows that I want to drop without having to drop them row by row. By simply setting a range, it makes the code more optimised. By setting the axis = 0 in the drop function, I will be ensuring that rows are dropped and not any columns.

After dropping, I will be using the iloc function to set the new header with the months as the new header of the table.

In [85]:

```python
#Dropping the headers
weather_details_df.drop(labels=range(0,6), axis=0, inplace = True)

#Dropping the footers
weather_details_df.drop(labels=range(17,35), axis=0, inplace = True)

#Making the Data Series and months as the headers for the table
weather_details_df.columns = weather_details_df.iloc[0]
weather_details_df = weather_details_df[1:]

#Printing the head and tail to check the table
display(weather_details_df.head())
display(weather_details_df.tail())
```

6	Data Series	2022 Sep	2022 Aug	2022 Jul	2022 Jun	2022 May	2022 Apr	2022 Mar	2022 Feb	2022 Jan	...	1975 Oct	1975 Sep	1975 Aug	1975 Jul	1975 Jun	1975 May	1975 Apr	1975 Mar	1975 Feb	1975 Jan
7	Air Temperature Means Daily Maximum (Degree Ce...	31.1	31.2	31.7	31.1	32.9	32.1	32.7	31.2	31.6	...	30.5	30.3	30.3	30	30.6	31.4	31.1	31.1	30.4	30.5
8	Air Temperature Means Daily Minimum (Degree Ce...	25.4	25.8	25.8	25.3	26.2	25.1	25.1	24.8	24.9	...	23.3	23.5	23.6	22.9	23.7	24	23.8	23.1	23	22.8
9	Air Temperature Absolute Extremes Maximum (Deg...	32.6	33.2	32.9	33.6	35.4	34	34.2	33.6	33.1	...	32.6	32.1	31.7	31.7	32.2	33.1	32.8	32.5	32.9	31.8
10	Air Temperature Absolute Extremes Minimum (Deg...	22.4	23.7	22.2	22.9	23.5	23.1	23	22.6	23.2	...	21.5	21.7	20.5	20.3	21.5	22.2	22.1	21.8	21.4	20
11	Total Rainfall (Millimetre)	121	141.4	145.2	211.8	102	239.6	163.2	175	99.8	...	169.8	146.5	177	192.5	212.8	104.6	79.7	197.9	98.6	121

5 rows × 574 columns

6	Data Series	2022 Sep	2022 Aug	2022 Jul	2022 Jun	2022 May	2022 Apr	2022 Mar	2022 Feb	2022 Jan	...	1975 Oct	1975 Sep	1975 Aug	1975 Jul	1975 Jun	1975 May	1975 Apr	1975 Mar	1975 Feb	1975 Jan
12	Highest Daily Rainfall Total (Millimetre)	38.2	32.4	27.8	59.4	44.6	58.8	36.2	29.6	50	...	52.1	45.1	62	54.2	71	49	16.7	40	38	41
13	Number Of Rainy Days (Number)	14	15	16	17	12	18	21	16	9	...	20	13	17	15	14	13	17	18	13	13
14	Bright Sunshine Daily Mean (Hour)	3.8	5.5	6.2	4.7	6.7	5.2	6.4	5.9	8.3	...	4.4	5.5	6	6.1	4.9	5.6	5.7	5.8	4.8	6.2
15	Minimum Relative Humidity (Per Cent)	52	50	53	54	41	51	48	51	49	...	na	na	na	na	na	na	na	na	na	na
16	24 Hours Mean Relative Humidity (Per Cent)	76.9	77.4	76.8	79.1	76.5	80	79.2	80.3	77.4	...	86	86.8	86.3	86.5	86.5	86.1	87.9	87.8	86.2	85.8

5 rows × 574 columns

Dropping the fields that will not be used

For this survey, I will be focusing on two main fields:

- Total Rainfall
- Total number of rainy days

These two fields will allow me to study the general trend in rainfall amounts in Singapore across the decade. Anaylising total number of rainy days allows me to study the rainfall patterns and spot any irregularities.

I will be dropping every other fields by using the drop function. In the drop function, I have keyed in the index of the fields that will not be used and have set the axis = 0 to drop those specific rows.

In [86]:

```
#Dropping fields using the index
weather_details_df.drop([7, 8, 9, 10, 12, 14, 15, 16], axis = 0, inplace = True)


#Printing header to check
weather_details_df.head()
```

Out[86]:

6	Data Series	2022 Sep	2022 Aug	2022 Jul	2022 Jun	2022 May	2022 Apr	2022 Mar	2022 Feb	2022 Jan	...	1975 Oct	1975 Sep	1975 Aug	1975 Jul	1975 Jun	1975 May	1975 Apr	1975 Mar	1975 Feb	1975 Jan
11	Total Rainfall (Millimetre)	121	141.4	145.2	211.8	102	239.6	163.2	175	99.8	...	169.8	146.5	177	192.5	212.8	104.6	79.7	197.9	98.6	121
13	Number Of Rainy Days (Number)	14	15	16	17	12	18	21	16	9	...	20	13	17	15	14	13	17	18	13	13

2 rows × 574 columns

Converting values from String to Float

The rainfall data in the original dataset acquired is stored in strings. However, for me to carry out calculations, I need to values to be stored as floats.

In [87]:

```
headers = list(weather_details_df.columns.values)
weatherYears = headers[1:]
weather_details_df[weatherYears] = weather_details_df[weatherYears].astype("float64")
```

Creating a new table to Store the Annual Weather Data

As one can see in the table above, I have the monthly rainfall data. For the report, I need the annual data of every year. Before calculating the annual data, I will be creating a new table to store the annual weather data. For that, I have created a new dataframe named Annual_weather_df

In [88]:

```
#Creating a new dataset, will be keeping the headers
Annual_weather_df = pd.DataFrame(weather_details_df[headers[0]].copy())

#Check the headers
display(Annual_weather_df)
```

	Data Series
11	Total Rainfall (Millimetre)
13	Number Of Rainy Days (Number)

Calculating Annual Data

The dataset acquired has the rainfall data in months. For this report, I need the annual data. For that, I have created a for loop to calculate the annual data. The for loop sums up the months in a year and calculates the mean to produce the annual data. Since a year consists of 12 months, a for loop can be created to find the total mean of every 12 months in repeated cycles.

For example, 2020 Annual Data = mean of (Jan 2020 + Feb 2020 + ... + Dec 2020) and then moves on the next 12 months which will be the one of 2019. These data till be stored in the new Annual_weather_df that I have created above. One exception will be for the year 2022 as the data set only has the rainfall data till september.

Calculating annual data as such is a more optimised way instead of manually adding the months and finding its mean from years 1975 to 2022.

In [89]:

```
#finding mean for year 2022
year_C = 1

#For year 2022, its set to 10 as the dataset is only till sept
year2022 = headers[1:10]

Annual_weather_df[headers[1][:4]] = weather_details_df[year2022].mean(axis = 1)

#Year increases by one from 1975 to 2022
year_C += 1

#Since every year has the data for 12 months, I can do a for loop to add in cycles of 12
index = range(10, len(headers), 12)
for idx in index:
    currentYear = headers[idx:idx+12]
    Annual_weather_df[headers[idx][:4]] = weather_details_df[currentYear].mean(axis = 1)
    year_C += 1


#Display the table to check
display(Annual_weather_df.head())
```

	Data Series	2022	2021	2020	2019	2018	2017	2016	2015	2014	...	1984	1983
11	Total Rainfall (Millimetre)	155.444444	234.133333	157.216667	113.958333	142.350000	170.466667	162.975000	105.566667	128.200000	...	223.891667	173.400000
13	Number Of Rainy Days (Number)	15.333333	15.166667	14.833333	10.000000	15.666667	17.000000	14.916667	10.416667	12.666667	...	15.916667	12.083333

2 rows × 49 columns

Removing unwanted years

I will only be needing the years from 2000 to 2020. Now that I have concised the table to the years and not months, I can easily remove the years that are not needed.

I will also be cleaning the table by renaming the column from "Data Series" to "Years" and also resetting the index of the table.

In [90]:

```
headers = list(Annual_weather_df.columns.values)

#Dropping years to get dataset from year 2000
index = headers.index('1999')
Annual_weather_df = Annual_weather_df[headers[:index]]

#Dropping years 2022 and 2021
Annual_weather_df.drop(Annual_weather_df.columns[[1, 2]], axis=1, inplace=True)

#Renaming Data Series to Year
Annual_weather_df = Annual_weather_df.rename(columns={"Data Series": "Year"})

#Resetting the index
Annual_weather_df.reset_index(drop = True, inplace = True)
Annual_weather_df.head()

#Displaying the dataset to check
display(Annual_weather_df)
```

	Year	2020	2019	2018	2017	2016	2015	2014	2013	2012	...	2009	2008
0	Total Rainfall (Millimetre)	157.216667	113.958333	142.350000	170.466667	162.975000	105.566667	128.200000	229.033333	179.991667	...	160.075000	193.758333
1	Number Of Rainy Days (Number)	14.833333	10.000000	15.666667	17.000000	14.916667	10.416667	12.666667	17.166667	15.916667	...	13.833333	15.000000

2 rows × 22 columns

Changing the colums and rows by flipping them

For plotting, I will be making the Total Rainfall and Number of Rainy Days as the columns and the Years as the rows.

In [91]:

```
Annual_weather_df.info()

#to flip the rows and columns
Annual_weather_df = Annual_weather_df.T
Annual_weather_df.reset_index(inplace = True)

#to reset the index
Annual_weather_df.columns = Annual_weather_df.iloc[0]
Annual_weather_df = Annual_weather_df[1:]
Annual_weather_df.reset_index(drop = True, inplace = True)

#to ensure data is stored as the right data type
Annual_weather_df["Year"] = Annual_weather_df["Year"].astype("int64")
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2 entries, 0 to 1
Data columns (total 22 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Year    2 non-null      object
 1   2020    2 non-null      float64
 2   2019    2 non-null      float64
 3   2018    2 non-null      float64
 4   2017    2 non-null      float64
 5   2016    2 non-null      float64
 6   2015    2 non-null      float64
 7   2014    2 non-null      float64
 8   2013    2 non-null      float64
 9   2012    2 non-null      float64
 10  2011    2 non-null      float64
 11  2010    2 non-null      float64
 12  2009    2 non-null      float64
 13  2008    2 non-null      float64
 14  2007    2 non-null      float64
 15  2006    2 non-null      float64
 16  2005    2 non-null      float64
 17  2004    2 non-null      float64
 18  2003    2 non-null      float64
 19  2002    2 non-null      float64
 20  2001    2 non-null      float64
 21  2000    2 non-null      float64
dtypes: float64(21), object(1)
memory usage: 480.0+ bytes
```

Merging Dataset

Now that all the dataframes are cleaned and processed, I can now merge all the data into one dataframe for easier analysis. I will be merging both my datasets into a singular dataset named Singapore to view all the information under on dataset. With the merge, I can bring Singapore_Carbon_emissions and Annual_weather_df into one dataframe named Singapore_df.

In [92]:

```
#Merging Singapore_Carbon_emissions and Annual_weather_df into one dataframe named Singapore_df
Singapore_df = pd.merge(Singapore_Carbon_emissions, Annual_weather_df, on = "Year", how = "outer")

#displaying the header and tail to check the merge
display(Singapore_df.head())
display(Singapore_df.tail())
```

	Country	Code	Calling Code	Year	CO2 emission (Tons)	Total Rainfall (Millimetre)	Number Of Rainy Days (Number)
0	Singapore	SG	65	2000	1.166442e+09	197.541667	15.666667
1	Singapore	SG	65	2001	1.215642e+09	231.925	15.333333
2	Singapore	SG	65	2002	1.262736e+09	145.741667	11.583333
3	Singapore	SG	65	2003	1.312174e+09	199.266667	15.5
4	Singapore	SG	65	2004	1.359033e+09	178.033333	12.333333

	Country	Code	Calling Code	Year	CO2 emission (Tons)	Total Rainfall (Millimetre)	Number Of Rainy Days (Number)
16	Singapore	SG	65	2016	1.980751e+09	162.975	14.916667
17	Singapore	SG	65	2017	2.019817e+09	170.466667	17.0
18	Singapore	SG	65	2018	2.065819e+09	142.35	15.666667
19	Singapore	SG	65	2019	2.111524e+09	113.958333	10.0
20	Singapore	SG	65	2020	2.157028e+09	157.216667	14.833333

Data Visualisation

In this section, I will be presenting the dataset in different visualisations to study the results and analyse the proposed hypothesis.

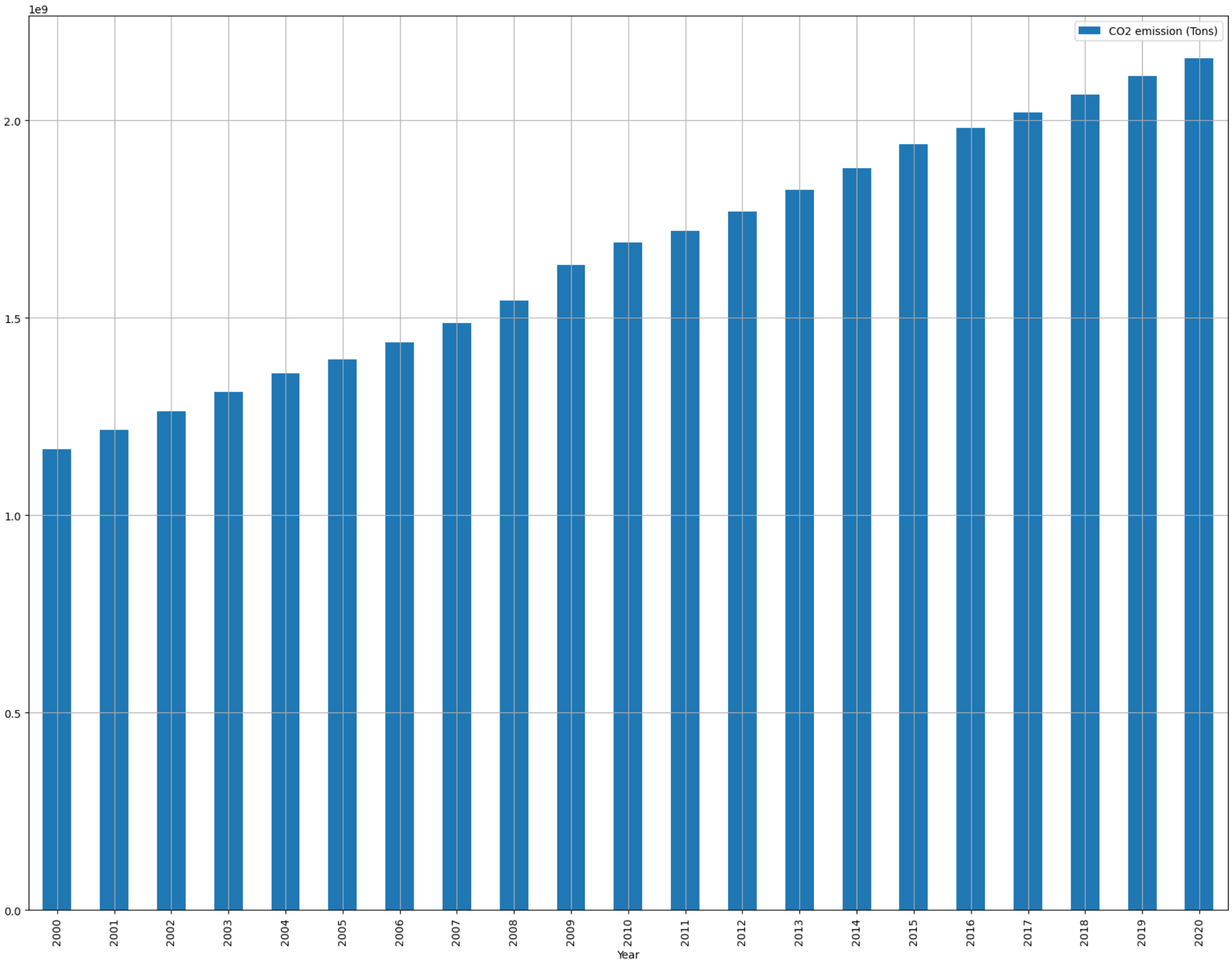
Analysing the carbon dioxide emissions across the last decade in Singapore

In [93]:

```
#carbon emission across the years
# x axis will be Year
# y axis will be Carbon dioxide emission (Tons)

Singapore_df.plot.bar(x = 'Year', y = 'CO2 emission (Tons)', figsize = (20, 15), grid = True)
```

Out[93]:

```
<AxesSubplot:xlabel='Year'>
```



Discussion of analysis

Generally, there has been an increase in carbon dioxide emissions across the years. Though there have been active climate change actions, the emissions are contradictory to that. This could be due to Singapore's heavy reliance on burning fossil fuels to meet its energy needs. (Nccs.gov.sg, 2018)

In the earlier section of the report where I have pointed out key milestone climate change protocols that Singapore took part in. I expected to see a decrease, at least a mild dip, as such protocols might have had an impact on the carbon emissions. However, it is clearly evident that these protocols work on a longer term basis with countries trying to adapt their mode of operation.

Many ways Singapore could reduce its carbon emissions is to use alternative natural energy sources like solar power. Singapore generally has been facing limited renewable energy such as limited hydro resources, limited wind speeds and a lack of geothermal energy. Over the last decade, solar energy has not been economically viable for Singapore. (Ema.gov.sg, 2022) But with more developments and economic stability, It is currently under work in the Singapore Green Plan 2030 where Singapore is planning to deploy solar energy to at least two gigawatt-peak.

This shows how even countries like Singapore are actively taking part in climate change actions, that does not mean it will lead to a change in carbon emissions over a span of a decade. It may take multiple decades and in this case Singapore is using 2030 as its benchmark year.

This answers Question 1 on Key Questions to Address under Introduction to Research space.

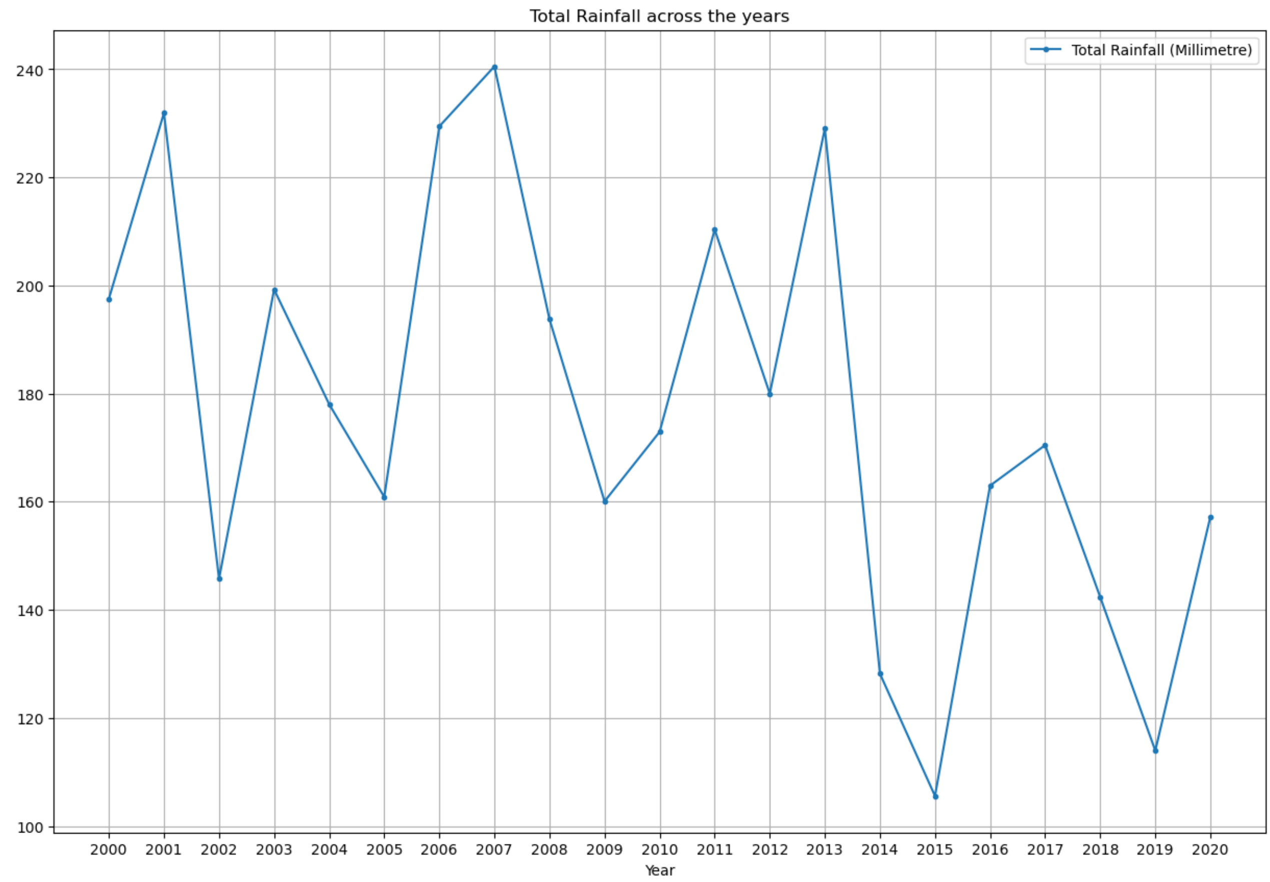
Analysing Total Rainfall across the last decade in Singapore

In [94]:

```
# presenting total rainfall across the years
# x axis will be Year
# y axis will be "Total Rainfall (Millimetre)

years = Singapore_df["Year"].tolist()
Singapore_df.plot(title = "Total Rainfall across the years", kind = "line", figsize = (15, 10), \
                                                x = "Year", y = "Total Rainfall (Millimetre)", xticks = years, \
                                               marker = ".", grid = True)
```

Out[94]:

```
<AxesSubplot:title={'center':'Total Rainfall across the years'}, xlabel='Year'>
```



Analysis

According the graph above, there is no significant trend in the rainfall across the decade. This irregularity in weather might be due to climate change as extreme changes in weather causes extreme changes in rainfall amounts.

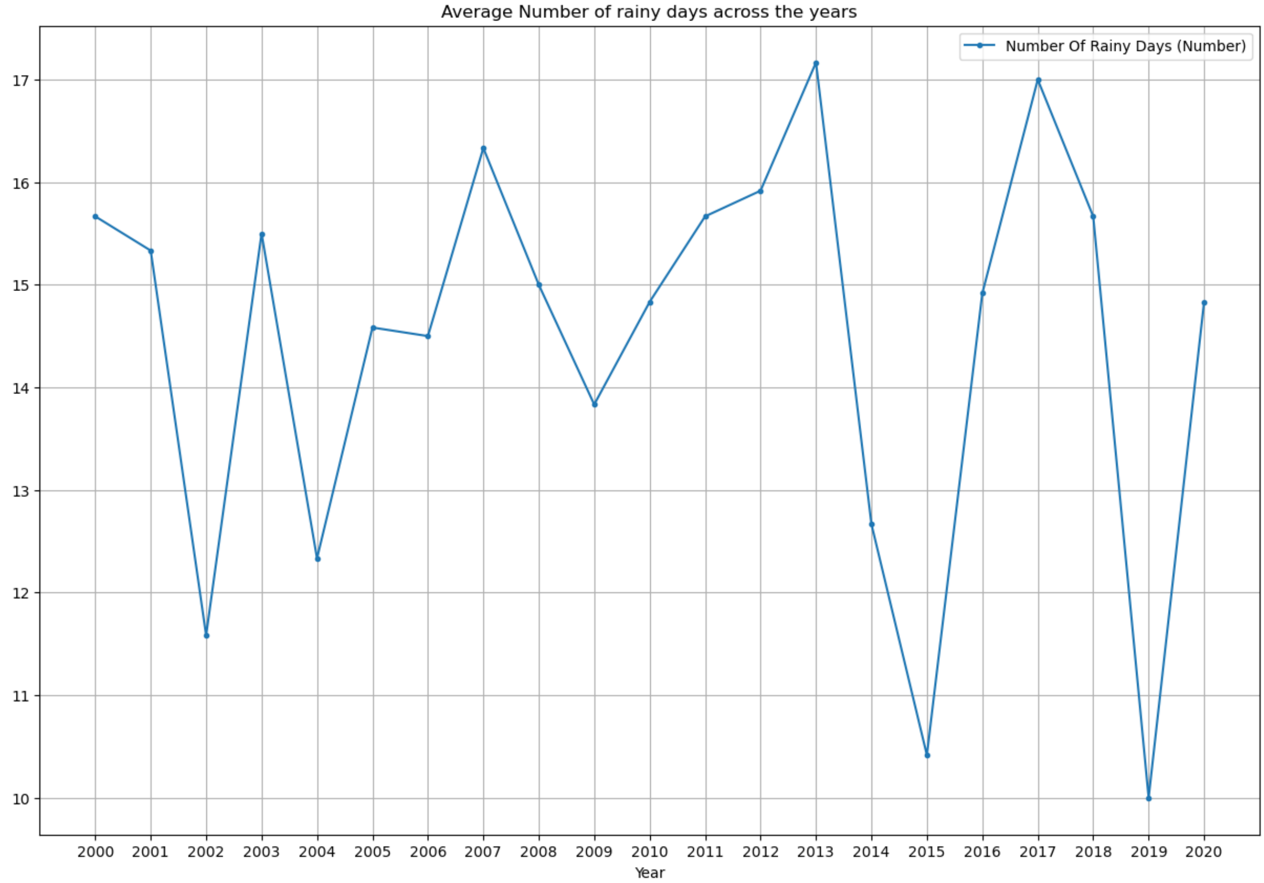
Analysing Average Number of Rainy days across the years

In [95]:

```
# x axis will be Year
# y axis will be Number Of Rainy Days (Number)

years = Singapore_df["Year"].tolist()
# presenting number of rainy days across the years
Singapore_df.plot(title = "Average Number of rainy days across the years", kind = "line", figsize = (15, 10), \
                                                    x = "Year", y = "Number Of Rainy Days (Number)", xticks = years, \
                                                   marker = ".", grid = True)
```

Out[95]:

```
<AxesSubplot:title={'center':'Average Number of rainy days across the years'}, xlabel='Year'>
```



Analysis and Explanation

From the graph, it evident that Singapore generally has a high number of average number of rainy days. There can be two possible reasons for this.

1. Singapore's rainfall is affected by the northeast and southwest monsoon causing Singapore to receive high precipitation across the years.
2. The carbon emission can be contributing to the rainfall (which be explored in the next section)

Analysing the relationship between carbon dioxide emissions and Rainfall Across the decade in Singapore

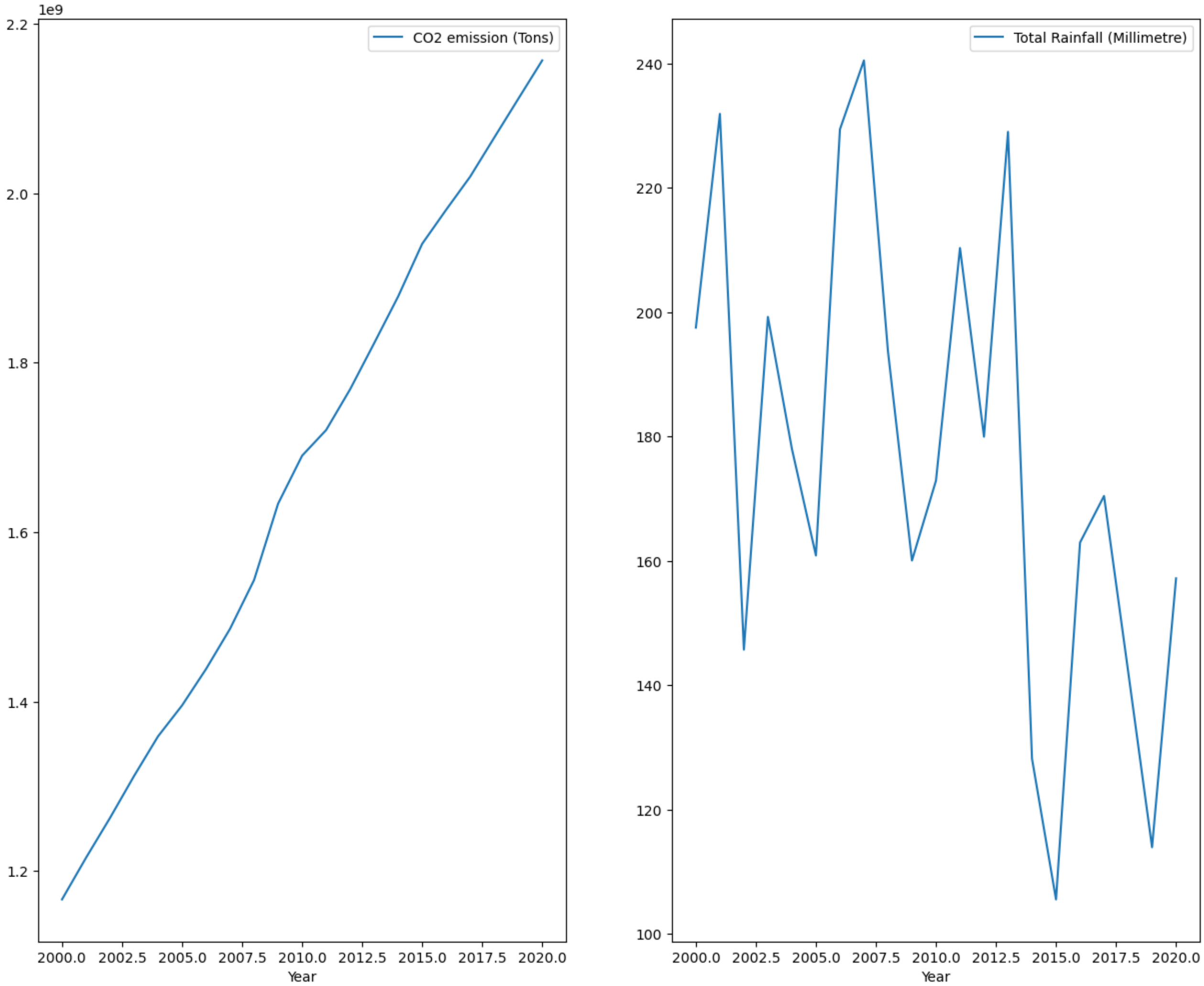
For this plot, I will be using the subplot function to plot the carbon dioxide emissions across the years on the left side graph and the Rainfall Across the decade in Singapore on the right side graph. Plotting these two on the same graph will not be possible as the emissions are in tons and rainfall data is in range below 300 millimetre. Hence, to compare their corelation, I have to plot them seperately but beside each other to study the relationship.

In [96]:

```
fig = plt.figure(figsize=(15,12))

#carbon dioxide emissions across the years
axes1 = fig.add_subplot(121)

#Rainfall Across the decade in Singapore
axes2 = fig.add_subplot(122)

Singapore_df.plot(x='Year', y='CO2 emission (Tons)', ax=axes1)
Singapore_df.plot(x='Year', y='Total Rainfall (Millimetre)', ax=axes2)
```

Out[96]:

```
<AxesSubplot:xlabel='Year'>
```



Analysis and Explanation

For this section, I have plotted the two graphs, Carbon dioxide emissions across the years and Total Rainfall across the years beside each other. By doing so, it is easier to study and compare the trends between the two factors. From this plot, it is evident that there is no relation between total rainfall and carbon dioxide emission. Even though carbon dioxide emissions have been on the rise over the last decade, there is no corelation to the rainfall.

For example, in the year 2020, the rainfall is relatively low even though it has a high carbon dioxide emission that year. On the other end of the spectrum, the year 2007 has the highest recorded rainfall in the last decade but the carbon dioxide emission was significantly lesser compared to 2020.

Overall Conclusion

Hypothesis 1: Increase in Carbon Dioxide emission in Singapore has led to increased Rainfalls.

After analysing the data, it is clearly evident that the carbon dioxide emissions has been on the rise over the decade. However, the rainfall patterns over the decade has been varied with significant drops and rises in certain years with little to no clear visible trends in increase in rainfalls. Hence, there is no evidence from the data to support this hypothesis.

Hypothesis 2: Rainfall is a good metric that can be used to study climate change in Singapore

Generally, increased rainfall is a metric used by climate scientists to study climate change as increased rainfall is a common effect of climate change. However, looking at the rainfall patterns in Singapore over the last decade, it states otherwise. There is a no significant growing trend to prove this hypothesis.

Conclusion

With Singapore actively taking part in climate change protocol, it will take more than a decade for Singapore to change from its current reliance on fossil fuels for electricity to cleaner energy sources. With the government using 2030 as the benchmark year and currently allocating more budget for cleaner energy alternatives, it is expected to see an improvement in the carbon dioxide emissions.

As for rainfall, climate change not only causes increase in rainfalls, it also creates drastic anomalies in the climate which can also lead to severe dip in rainfall as much as it may also cause severe floods due to high rainfall. In the visualisations, it is evident that Singapore is experiencing significant anomalies in the rainfall. With Singapore being in the tropical region affected by the monsoons and naturally being prone to receiving high precipitations, it might be necessary to consider other metrics like surface temperatures.

Ethics of usage of data

For both my datasets, I have taken from Kaggle and another from Singapore Department of Statistics. Both websites strictly only allows their data to be used for the purpose of a study research. The data from their sites must be utilised properly with no commercial purposes. For my case, I have used these datasets for my school report and not for any commercial reasons.

The report also clearly demonstrates all the data processing that has been done using the dataframes. It reduces the potential for any personally indentifiable distinctions and all the data is readily accessible in the notebook.

Hence, the datasets has been used ethically.

Considerations about usage of this data

This report that I have written can be used for a wide range of purposes. Students can use this report to further continue their study on Singapore climate change using other metrics as stated in the conclusion section to study if carbon emissions have impacted other weather elements.

Policy makers can also use this report to study and discuss how their efforts in the last decade has had any impacts in the local climate. They can use this extrapolate and predict the future climate change and see what can be done to tackle this issue. This report functions as a foundation to start many discussions on past and future efforts to be taken.

Potential biasness/problems of this dataset ¶

This has been addressed in the analysis and conclusion section of this report.

References

1. ‘A Reason to Act Faster’: World Leaders Meet on Climate Amid Other Crises. (2022). The New York Times. [online] Available at: https://www.nytimes.com/2022/11/07/climate/climate-change-crisis-cop27.html (https://www.nytimes.com/2022/11/07/climate/climate-change-crisis-cop27.html) [Accessed 17 Nov. 2022].

2. Nccs.gov.sg. (2015). Singapore And International Efforts. [online] Available at: https://www.nccs.gov.sg/singapores-climate-action/singapore-and-international-efforts/#:~:text=Singapore%20also%20signed%20the%20Joint,Singapore%20Cooperation%20Programme%20(SCP) (https://www.nccs.gov.sg/singapores-climate-action/singapore-and-international-efforts/#:~:text=Singapore%20also%20signed%20the%20Joint,Singapore%20Cooperation%20Programme%20(SCP)). [Accessed 17 Nov. 2022].

3. Nccs.gov.sg. (2018). Mitigation Action. [online] Available at: https://www.nccs.gov.sg/faqs/mitigation-action/ (https://www.nccs.gov.sg/faqs/mitigation-action/) [Accessed 29 Nov. 2022].

4. Ema.gov.sg. (2022). EMA | Renewable Energy - Overview. [online] Available at: https://www.ema.gov.sg/Renewable_Energy_Overview.aspx#:~:text=In%20Singapore%2C%20we%20have%20limited,when%20it%20becomes%20commercially%20 (https://www.ema.gov.sg/Renewable_Energy_Overview.aspx#:~:text=In%20Singapore%2C%20we%20have%20limited,when%20it%20becomes%20commercially%2 [Accessed 29 Nov. 2022].

5. Moazzim Ali Bhatti (2022). CO2 Emission by countries Year wise (1750-2022). [online] Kaggle.com. Available at: https://www.kaggle.com/datasets/moazzimalibhatti/co2-emission-by-countries-year-wise-17502022 (https://www.kaggle.com/datasets/moazzimalibhatti/co2-emission-by-countries-year-wise-17502022) [Accessed 1 Jan. 2023].

6. Singstat.gov.sg. (2023). [online] Available at: https://tablebuilder.singstat.gov.sg/table/TS/M890191 (https://tablebuilder.singstat.gov.sg/table/TS/M890191) [Accessed 1 Jan. 2023].

7. Mse.gov.sg. (2018). Climate Change. [online] Available at: https://www.mse.gov.sg/policies/climate-change#:~:text=Protecting%20Against%20Sea%20Level%20Rise&text=Singapore%20is%20particularly%20vulnerable%20to,about%201%20metre%20by%202100 (https://www.mse.gov.sg/policies/climate-change#:~:text=Protecting%20Against%20Sea%20Level%20Rise&text=Singapore%20is%20particularly%20vulnerable%20to,about%201%20metre%20by%202100). [Accessed 5 Jan. 2023].

Report by Logesh Kiran

UOL Student Number: 210201255

SIM Student ID: 10215008