```
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
from sklearn.impute import SimpleImputer


df=pd.read_csv('/content/visualizing_global_co2_data.csv')
```

READ FILE

```
df.head()
```

	country	year	iso_code	population	gdp	cement_co2	cement_co2_per_capita	co2	co2_growth_abs	co2_growth_prct	...	share_glol
0	Afghanistan	1850	AFG	3752993.0	NaN	NaN	NaN	NaN	NaN	NaN	...	
1	Afghanistan	1851	AFG	3767956.0	NaN	NaN	NaN	NaN	NaN	NaN	...	
2	Afghanistan	1852	AFG	3783940.0	NaN	NaN	NaN	NaN	NaN	NaN	...	
3	Afghanistan	1853	AFG	3800954.0	NaN	NaN	NaN	NaN	NaN	NaN	...	
4	Afghanistan	1854	AFG	3818038.0	NaN	NaN	NaN	NaN	NaN	NaN	...	

5 rows × 79 columns

```
df.columns
```

```
Index(['country', 'year', 'iso_code', 'population', 'gdp', 'cement_co2',
       'cement_co2_per_capita', 'co2', 'co2_growth_abs', 'co2_growth_prct',
       'co2_including_luc', 'co2_including_luc_growth_abs',
       'co2_including_luc_growth_prct', 'co2_including_luc_per_capita',
       'co2_including_luc_per_gdp', 'co2_including_luc_per_unit_energy',
       'co2_per_capita', 'co2_per_gdp', 'co2_per_unit_energy', 'coal_co2',
       'coal_co2_per_capita', 'consumption_co2', 'consumption_co2_per_capita',
       'consumption_co2_per_gdp', 'cumulative_cement_co2', 'cumulative_co2',
       'cumulative_co2_including_luc', 'cumulative_coal_co2',
       'cumulative_flaring_co2', 'cumulative_gas_co2', 'cumulative_luc_co2',
       'cumulative_oil_co2', 'cumulative_other_co2', 'energy_per_capita',
       'energy_per_gdp', 'flaring_co2', 'flaring_co2_per_capita', 'gas_co2',
       'gas_co2_per_capita', 'ghg_excluding_lucf_per_capita', 'ghg_per_capita',
       'land_use_change_co2', 'land_use_change_co2_per_capita', 'methane',
       'methane_per_capita', 'nitrous_oxide', 'nitrous_oxide_per_capita',
       'oil_co2', 'oil_co2_per_capita', 'other_co2_per_capita',
       'other_industry_co2', 'primary_energy_consumption',
       'share_global_cement_co2', 'share_global_co2',
       'share_global_co2_including_luc', 'share_global_coal_co2',
       'share_global_cumulative_cement_co2', 'share_global_cumulative_co2',
       'share_global_cumulative_co2_including_luc',
       'share_global_cumulative_coal_co2',
       'share_global_cumulative_flaring_co2',
       'share_global_cumulative_gas_co2', 'share_global_cumulative_luc_co2',
       'share_global_cumulative_oil_co2', 'share_global_cumulative_other_co2',
       'share_global_flaring_co2', 'share_global_gas_co2',
       'share_global_luc_co2', 'share_global_oil_co2',
       'share_global_other_co2', 'share_of_temperature_change_from_ghg',
       'temperature_change_from_ch4', 'temperature_change_from_co2',
       'temperature_change_from_ghg', 'temperature_change_from_n2o',
       'total_ghg', 'total_ghg_excluding_lucf', 'trade_co2',
       'trade_co2_share'],
      dtype='object')
```

```
df.shape
```

```
(50598, 79)
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50598 entries, 0 to 50597
Data columns (total 79 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   country                               50598 non-null  object
 1   year                                  50598 non-null  int64
```

```
2   iso_code                              42142 non-null  object
3   population                            40008 non-null  float64
4   gdp                                   14564 non-null  float64
5   cement_co2                            24974 non-null  float64
6   cement_co2_per_capita                 22714 non-null  float64
7   co2                                   31349 non-null  float64
8   co2_growth_abs                        29010 non-null  float64
9   co2_growth_prct                       25032 non-null  float64
10  co2_including_luc                     24218 non-null  float64
11  co2_including_luc_growth_abs          23917 non-null  float64
12  co2_including_luc_growth_prct         24123 non-null  float64
13  co2_including_luc_per_capita          23696 non-null  float64
14  co2_including_luc_per_gdp             15701 non-null  float64
15  co2_including_luc_per_unit_energy     9722 non-null   float64
16  co2_per_capita                        26915 non-null  float64
17  co2_per_gdp                           16291 non-null  float64
18  co2_per_unit_energy                   10276 non-null  float64
19  coal_co2                              25069 non-null  float64
20  coal_co2_per_capita                   24421 non-null  float64
21  consumption_co2                       4600 non-null   float64
22  consumption_co2_per_capita            4227 non-null   float64
23  consumption_co2_per_gdp               3898 non-null   float64
24  cumulative_cement_co2                 24873 non-null  float64
25  cumulative_co2                        29356 non-null  float64
26  cumulative_co2_including_luc          24218 non-null  float64
27  cumulative_coal_co2                   24968 non-null  float64
28  cumulative_flaring_co2                24774 non-null  float64
29  cumulative_gas_co2                    24842 non-null  float64
30  cumulative_luc_co2                    39388 non-null  float64
31  cumulative_oil_co2                    24941 non-null  float64
32  cumulative_other_co2                  2386 non-null   float64
33  energy_per_capita                     10320 non-null  float64
34  energy_per_gdp                        7159 non-null   float64
35  flaring_co2                           24875 non-null  float64
36  flaring_co2_per_capita                24272 non-null  float64
37  gas_co2                               24943 non-null  float64
38  gas_co2_per_capita                    24340 non-null  float64
39  ghg_excluding_lucf_per_capita         6149 non-null   float64
40  ghg_per_capita                        6149 non-null   float64
41  land_use_change_co2                   39388 non-null  float64
42  land_use_change_co2_per_capita        36026 non-null  float64
43  methane                               6150 non-null   float64
44  methane_per_capita                    6150 non-null   float64
45  nitrous_oxide                         6150 non-null   float64
46  nitrous_oxide_per_capita              6150 non-null   float64
47  oil_co2                               25042 non-null  float64
48  oil_co2_per_capita                    24367 non-null  float64
49  other_co2_per_capita                  2386 non-null   float64
50  other_industry_co2                    2386 non-null   float64
51  primary_energy_consumption            10391 non-null  float64
```

```
df.isnull().sum()
```

```
country                            0
year                               0
iso_code                        8456
population                     10590
gdp                            36034
                               ...
temperature_change_from_n2o    12978
total_ghg                      44449
total_ghg_excluding_lucf       44449
trade_co2                      46339
trade_co2_share                46340
Length: 79, dtype: int64
```

HANDLING NULL VALUES

```
num_cols=df.select_dtypes(include=np.number).columns.tolist()
```

```
num_cols
```

```
['year',
 'population',
 'gdp',
 'cement_co2',
 'cement_co2_per_capita',
 'co2',
 'co2_growth_abs',
 'co2_growth_prct',
 'co2_including_luc',
```

```
        'co2_including_luc_growth_abs',
        'co2_including_luc_growth_prct',
        'co2_including_luc_per_capita',
        'co2_including_luc_per_gdp',
        'co2_including_luc_per_unit_energy',
        'co2_per_capita',
        'co2_per_gdp',
        'co2_per_unit_energy',
        'coal_co2',
        'coal_co2_per_capita',
        'consumption_co2',
        'consumption_co2_per_capita',
        'consumption_co2_per_gdp',
        'cumulative_cement_co2',
        'cumulative_co2',
        'cumulative_co2_including_luc',
        'cumulative_coal_co2',
        'cumulative_flaring_co2',
        'cumulative_gas_co2',
        'cumulative_luc_co2',
        'cumulative_oil_co2',
        'cumulative_other_co2',
        'energy_per_capita',
        'energy_per_gdp',
        'flaring_co2',
        'flaring_co2_per_capita',
        'gas_co2',
        'gas_co2_per_capita',
        'ghg_excluding_lucf_per_capita',
        'ghg_per_capita',
        'land_use_change_co2',
        'land_use_change_co2_per_capita',
        'methane',
        'methane_per_capita',
        'nitrous_oxide',
        'nitrous_oxide_per_capita',
        'oil_co2',
        'oil_co2_per_capita',
        'other_co2_per_capita',
        'other_industry_co2',
        'primary_energy_consumption',
        'share_global_cement_co2',
        'share_global_co2',
        'share_global_co2_including_luc',
        'share_global_coal_co2',
        'share_global_cumulative_cement_co2',
        'share_global_cumulative_co2',
        'share_global_cumulative_co2_including_luc',
        'share_global_cumulative_coal_co2',
```

```
imputer= SimpleImputer(strategy='mean')
df[num_cols]=imputer.fit_transform(df[num_cols])
```

```
df.isnull().sum()
```

```
country                          0
year                             0
iso_code                      8456
population                       0
gdp                              0
                                ...
temperature_change_from_n2o      0
total_ghg                        0
total_ghg_excluding_lucf         0
trade_co2                        0
trade_co2_share                  0
Length: 79, dtype: int64
```

```
df.dropna(inplace=True)
```

```
df.isnull().sum()
```

```
country                       0
year                          0
iso_code                      0
population                    0
gdp                           0
                             ..
temperature_change_from_n2o   0
total_ghg                     0
total_ghg_excluding_lucf      0
```

```
trade_co2                   0
trade_co2_share             0
Length: 79, dtype: int64
```

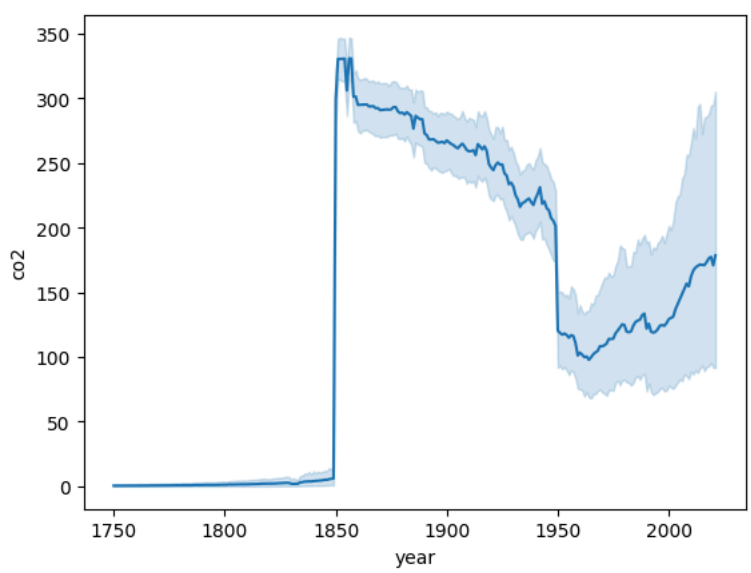
```
df.head(3)
```

population	gdp	cement_co2	cement_co2_per_capita	co2	co2_growth_abs
3752993.0	2.677586e+11	8.392491	0.063958	380.237549	5.703529
3767956.0	2.677586e+11	8.392491	0.063958	380.237549	5.703529
3783940.0	2.677586e+11	8.392491	0.063958	380.237549	5.703529

EDA ANALYSIS

1. How have global emissions of carbon dioxide (CO2) changed over time?

```
sns.lineplot(x='year',y='co2',data=df)
plt.show()
```



2. Who emits the most CO2 each year?

```
df1=df.groupby('country')['co2'].sum().reset_index().sort_values(by='co2',ascending=False).head(5)
sns.barplot(x='country',y='co2',data=df1)
plt.show()
```

3. Where in the world does the average person emit the most carbon dioxide (CO2) each year?

```
# prompt:

df2=df.groupby('country')['co2_per_capita'].mean().reset_index().sort_values(by='co2_per_capita',ascending=False).head(3)
sns.barplot(x='country',y='co2_per_capita',data=df2)
plt.show()
```