Data Processing Project

Import useful python libraries such as Numpy for numerical computation , Pandas for data analysis , Matplotlib & Seaborn for Data Visualization

In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

1. understand the data

we have First understand the data that means how much rows, columns present in it. Than we have to understand maximum, minimum value of data, standard deviation, mean, mode.

In [2]:
```
co2 = pd.read_csv('D:\machine learning project\co2_emission.csv')
co2
```

Out[2]:

	Entity	Code	Year	Annual CO_2 emissions (tonnes)
0	Afghanistan	AFG	1949	14656.00
1	Afghanistan	AFG	1950	84272.00
2	Afghanistan	AFG	1951	91600.00
3	Afghanistan	AFG	1952	91600.00
4	Afghanistan	AFG	1953	106256.00
5	Afghanistan	AFG	1954	106256.00
6	Afghanistan	AFG	1955	153888.00
7	Afghanistan	AFG	1956	183200.00
8	Afghanistan	AFG	1957	293120.00
9	Afghanistan	AFG	1958	329760.00
10	Afghanistan	AFG	1959	384571.42
11	Afghanistan	AFG	1960	413883.42
12	Afghanistan	AFG	1961	490797.70
13	Afghanistan	AFG	1962	688594.27
14	Afghanistan	AFG	1963	706735.98
15	Afghanistan	AFG	1964	838550.83
16	Afghanistan	AFG	1965	1006916.53
17	Afghanistan	AFG	1966	1091158.82
18	Afghanistan	AFG	1967	1281865.11
19	Afghanistan	AFG	1968	1223389.69
20	Afghanistan	AFG	1969	941231.98
21	Afghanistan	AFG	1970	1670397.69
22	Afghanistan	AFG	1971	1893931.41
23	Afghanistan	AFG	1972	1531165.69
24	Afghanistan	AFG	1973	1637243.40
25	Afghanistan	AFG	1974	1915677.68
26	Afghanistan	AFG	1975	2124525.68
27	Afghanistan	AFG	1976	1984225.38
28	Afghanistan	AFG	1977	2386745.67
29	Afghanistan	AFG	1978	2155419.66
...	...	...	...	...
20823	Zimbabwe	ZWE	1988	16029460.93
20824	Zimbabwe	ZWE	1989	16113653.54
20825	Zimbabwe	ZWE	1990	15569294.66
20826	Zimbabwe	ZWE	1991	15865346.05
20827	Zimbabwe	ZWE	1992	16917418.67
20828	Zimbabwe	ZWE	1993	16259101.52
20829	Zimbabwe	ZWE	1994	17684463.69
20830	Zimbabwe	ZWE	1995	15029425.61
20831	Zimbabwe	ZWE	1996	14890560.82
20832	Zimbabwe	ZWE	1997	14294211.06
20833	Zimbabwe	ZWE	1998	14139636.04
20834	Zimbabwe	ZWE	1999	15727748.13
20835	Zimbabwe	ZWE	2000	13818181.23
20836	Zimbabwe	ZWE	2001	12509195.38
20837	Zimbabwe	ZWE	2002	11897006.10
20838	Zimbabwe	ZWE	2003	10608340.13
20839	Zimbabwe	ZWE	2004	9428548.18
20840	Zimbabwe	ZWE	2005	10699642.16
20841	Zimbabwe	ZWE	2006	10364869.23
20842	Zimbabwe	ZWE	2007	9834014.17
20843	Zimbabwe	ZWE	2008	7718874.78
20844	Zimbabwe	ZWE	2009	5519508.33
20845	Zimbabwe	ZWE	2010	7681010.98
20846	Zimbabwe	ZWE	2011	9439850.98
20847	Zimbabwe	ZWE	2012	7614600.24
20848	Zimbabwe	ZWE	2013	11536239.29
20849	Zimbabwe	ZWE	2014	11866348.41
20850	Zimbabwe	ZWE	2015	10907603.94
20851	Zimbabwe	ZWE	2016	9932649.88
20852	Zimbabwe	ZWE	2017	10397718.47

20853 rows × 4 columns

In [3]:
```
co2.shape
```

Out[3]:
```
(20853, 4)
```

In [4]:
```
co2.head()
```

Out[4]:

	Entity	Code	Year	Annual CO_2 emissions (tonnes)
0	Afghanistan	AFG	1949	14656.0
1	Afghanistan	AFG	1950	84272.0
2	Afghanistan	AFG	1951	91600.0
3	Afghanistan	AFG	1952	91600.0
4	Afghanistan	AFG	1953	106256.0

In [5]:
```
co2.tail()
```

Out[5]:

	Entity	Code	Year	Annual CO_2 emissions (tonnes)
20848	Zimbabwe	ZWE	2013	11536239.29
20849	Zimbabwe	ZWE	2014	11866348.41
20850	Zimbabwe	ZWE	2015	10907603.94
20851	Zimbabwe	ZWE	2016	9932649.88
20852	Zimbabwe	ZWE	2017	10397718.47

In [6]:
```
co2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20853 entries, 0 to 20852
Data columns (total 4 columns):
Entity                            20853 non-null object
Code                              18646 non-null object
Year                              20853 non-null int64
Annual CO₂ emissions (tonnes )    20853 non-null float64
dtypes: float64(1), int64(1), object(2)
memory usage: 651.7+ KB
```

In [7]:
```
co2.describe()
```

Out[7]:

	Year	Annual CO_2 emissions (tonnes)
count	20853.000000	2.085300e+04
mean	1953.339424	1.930517e+08
std	57.903089	1.345143e+09
min	1751.000000	-6.255223e+08
25%	1932.000000	3.187680e+05
50%	1971.000000	3.828880e+06
75%	1995.000000	3.706898e+07
max	2017.000000	3.615326e+10

In [8]:
```
co2.columns
```

Out[8]:
```
Index(['Entity', 'Code', 'Year', 'Annual CO₂ emissions (tonnes )'], dtype='object')
```

In [9]:
```
co2.nunique()
```

Out[9]:
```
Entity                              233
Code                                222
Year                                267
Annual CO₂ emissions (tonnes )    13092
dtype: int64
```

In [10]:
```
co2['Entity'].unique()
```

Out[10]:
```
array(['Afghanistan', 'Africa', 'Albania', 'Algeria', 'Americas (other)',
       'Andorra', 'Angola', 'Anguilla', 'Antarctic Fisheries',
       'Antigua and Barbuda', 'Argentina', 'Armenia', 'Aruba',
       'Asia and Pacific (other)', 'Australia', 'Austria', 'Azerbaijan',
       'Bahamas', 'Bahrain', 'Bangladesh', 'Barbados', 'Belarus',
       'Belgium', 'Belize', 'Benin', 'Bermuda', 'Bhutan', 'Bolivia',
       'Bonaire Sint Eustatius and Saba', 'Bosnia and Herzegovina',
       'Botswana', 'Brazil', 'British Virgin Islands', 'Brunei',
       'Bulgaria', 'Burkina Faso', 'Burundi', 'Cambodia', 'Cameroon',
       'Canada', 'Cape Verde', 'Cayman Islands',
       'Central African Republic', 'Chad', 'Chile', 'China',
       'Christmas Island', 'Colombia', 'Comoros', 'Republic of the Congo',
       'Cook Islands', 'Costa Rica', "Cote d'Ivoire", 'Croatia', 'Cuba',
       'Curacao', 'Cyprus', 'Czech Republic', 'Czechoslovakia',
       'Democratic Republic of Republic of the Congo', 'Denmark',
       'Djibouti', 'Dominica', 'Dominican Republic', 'EU-28', 'Ecuador',
       'Egypt', 'El Salvador', 'Equatorial Guinea', 'Eritrea', 'Estonia',
       'Ethiopia', 'Europe (other)', 'Faeroe Islands', 'Falkland Islands',
       'Fiji', 'Finland', 'France', 'French Guiana', 'French Polynesia',
       'Gabon', 'Gambia', 'Georgia', 'Germany', 'Ghana', 'Gibraltar',
       'Greece', 'Greenland', 'Grenada', 'Guadeloupe', 'Guatemala',
       'Guinea', 'Guinea-Bissau', 'Guyana', 'Haiti', 'Honduras',
       'Hong Kong', 'Hungary', 'Iceland', 'India', 'Indonesia',
       'International transport', 'Iran', 'Iraq', 'Ireland', 'Israel',
       'Italy', 'Jamaica', 'Japan', 'Jordan', 'Kazakhstan', 'Kenya',
       'Kiribati', 'Kuwait', 'Kyrgysztan', 'Kyrgyzstan', 'Laos', 'Latvia',
       'Lebanon', 'Lesotho', 'Liberia', 'Libya', 'Liechtenstein',
       'Lithuania', 'Luxembourg', 'Macao', 'Macedonia', 'Madagascar',
       'Malawi', 'Malaysia', 'Maldives', 'Mali', 'Malta',
       'Marshall Islands', 'Martinique', 'Mauritania', 'Mauritius',
       'Mexico', 'Micronesia (country)', 'Middle East', 'Moldova',
       'Mongolia', 'Montenegro', 'Montserrat', 'Morocco', 'Mozambique',
       'Myanmar', 'Namibia', 'Nauru', 'Nepal', 'Netherlands',
       'New Caledonia', 'New Zealand', 'Nicaragua', 'Niger', 'Nigeria',
       'Niue', 'North Korea', 'Norway', 'Oman', 'Pakistan', 'Palau',
       'Palestine', 'Panama', 'Papua New Guinea', 'Paraguay', 'Peru',
       'Philippines', 'Poland', 'Portugal', 'Qatar', 'Reunion', 'Romania',
       'Russia', 'Rwanda', 'Saint Helena', 'Saint Kitts and Nevis',
       'Saint Lucia', 'Saint Pierre and Miquelon',
       'Saint Vincent and the Grenadines', 'Samoa',
       'Sao Tome and Principe', 'Saudi Arabia', 'Senegal', 'Serbia',
       'Seychelles', 'Sierra Leone', 'Singapore',
       'Sint Maarten (Dutch part)', 'Slovakia', 'Slovenia',
       'Solomon Islands', 'Somalia', 'South Africa', 'South Korea',
       'South Sudan', 'Spain', 'Sri Lanka', 'Statistical differences',
       'Sudan', 'Suriname', 'Swaziland', 'Sweden', 'Switzerland', 'Syria',
       'Taiwan', 'Tajikistan', 'Tanzania', 'Thailand', 'Timor', 'Togo',
       'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey',
       'Turkmenistan', 'Turks and Caicos Islands', 'Tuvalu', 'Uganda',
       'Ukraine', 'United Arab Emirates', 'United Kingdom',
       'United States', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela',
       'Vietnam', 'Wallis and Futuna Islands', 'World', 'Yemen', 'Zambia',
       'Zimbabwe'], dtype=object)
```

In [11]:
```
co2.Year.unique()
```

Out[11]:
```
array([1949, 1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959,
       1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970,
       1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981,
       1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992,
       1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003,
       2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014,
       2015, 2016, 2017, 1751, 1752, 1753, 1754, 1755, 1756, 1757, 1758,
       1759, 1760, 1761, 1762, 1763, 1764, 1765, 1766, 1767, 1768, 1769,
       1770, 1771, 1772, 1773, 1774, 1775, 1776, 1777, 1778, 1779, 1780,
       1781, 1782, 1783, 1784, 1785, 1786, 1787, 1788, 1789, 1790, 1791,
       1792, 1793, 1794, 1795, 1796, 1797, 1798, 1799, 1800, 1801, 1802,
       1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813,
       1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824,
       1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835,
       1836, 1837, 1838, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846,
       1847, 1848, 1849, 1850, 1851, 1852, 1853, 1854, 1855, 1856, 1857,
       1858, 1859, 1860, 1861, 1862, 1863, 1864, 1865, 1866, 1867, 1868,
       1869, 1870, 1871, 1872, 1873, 1874, 1875, 1876, 1877, 1878, 1879,
       1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889, 1890,
       1891, 1892, 1893, 1894, 1895, 1896, 1897, 1898, 1899, 1900, 1901,
       1902, 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912,
       1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923,
       1924, 1925, 1926, 1927, 1928, 1929, 1930, 1931, 1932, 1933, 1934,
       1935, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1943, 1944, 1945,
       1946, 1947, 1948], dtype=int64)
```

In [12]:
```
co2.Year.min()
```

Out[12]:
```
1751
```

In [13]:
```
co2.Year.max()
```

Out[13]:
```
2017
```

In [14]:
```
co2['Annual CO₂ emissions (tonnes )'].max()
```

Out[14]:
```
36153261645.0
```

In [15]:
```
co2['Annual CO₂ emissions (tonnes )'].min()
```

Out[15]:
```
-625522256.7
```

In [16]:
```
condition=co2['Annual CO₂ emissions (tonnes )']>=0
df=co2[condition]
```

In [17]:
```
df['Annual CO₂ emissions (tonnes )'].min()
```

Out[17]:
```
0.0
```

In [18]:
```
df['Annual CO₂ emissions (tonnes )'].max()
```

Out[18]:
```
36153261645.0
```

In [19]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 20813 entries, 0 to 20852
Data columns (total 4 columns):
Entity                            20813 non-null object
Code                              18646 non-null object
Year                              20813 non-null int64
Annual CO₂ emissions (tonnes )    20813 non-null float64
dtypes: float64(1), int64(1), object(2)
memory usage: 813.0+ KB
```

2. Data Cleaning

Data cleaning is the time consuming process. This is the most important steps of data science life cycle. Here we have to perform function on missing value.

In [20]:
```
co2.isnull().sum()
```

Out[20]:
```
Entity                               0
Code                              2207
Year                                 0
Annual CO₂ emissions (tonnes )       0
dtype: int64
```

In [21]:
```
drop_col = co2.drop(['Code'], axis=1)
```

In [22]:
```
drop_col.head()
```

Out[22]:

	Entity	Year	Annual CO_2 emissions (tonnes)
0	Afghanistan	1949	14656.0
1	Afghanistan	1950	84272.0
2	Afghanistan	1951	91600.0
3	Afghanistan	1952	91600.0
4	Afghanistan	1953	106256.0

In [23]:
```
drop_col.isnull().sum()
```

Out[23]:
```
Entity                            0
Year                              0
Annual CO₂ emissions (tonnes )    0
dtype: int64
```

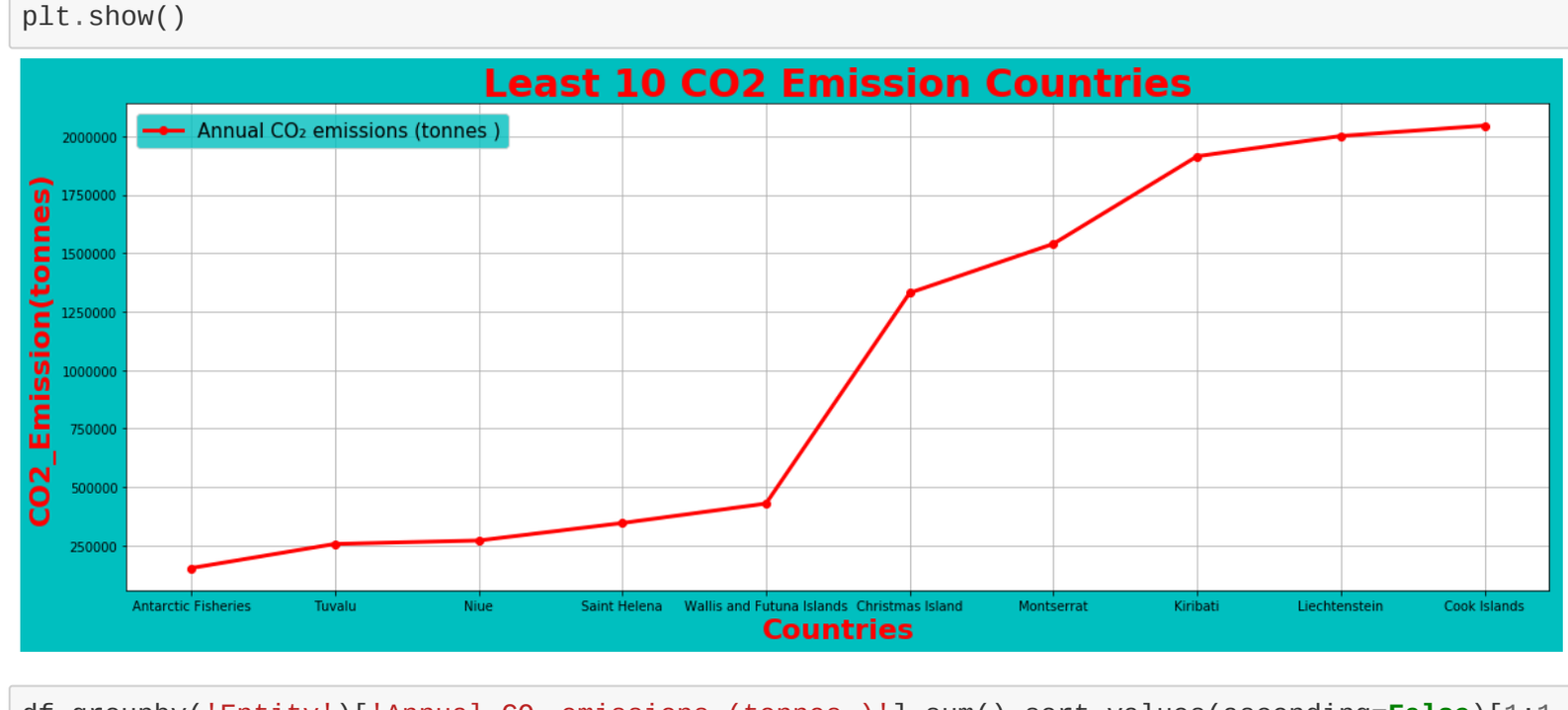
3. Visualisation

Data visualisation is the process through which we can give data to meaning full insight.

In [24]:
```
df.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=True)[1:11
]
```

Out[24]:
```
Entity
Antarctic Fisheries           153888.00
Tuvalu                        256922.58
Niue                          271578.58
Saint Helena                  346308.79
Wallis and Futuna Islands     429573.15
Christmas Island             1330032.00
Montserrat                   1539754.12
Kiribati                     1913894.59
Liechtenstein                2000665.44
Cook Islands                 2044872.31
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

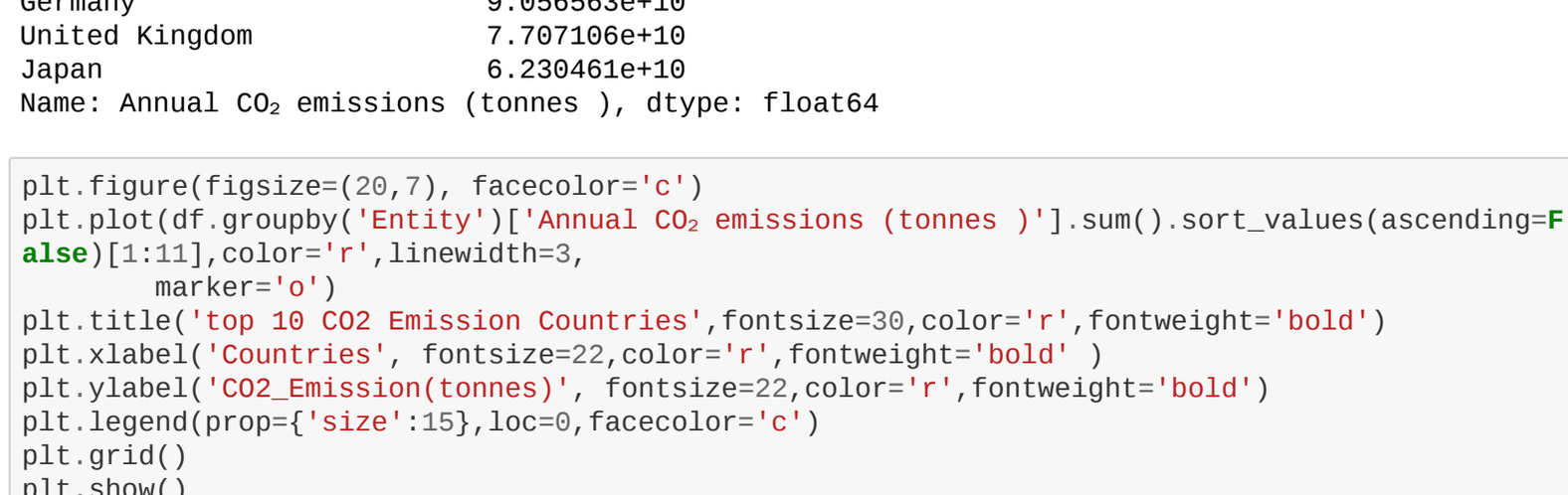
In [25]:
```
plt.figure(figsize=(20,7), facecolor='c')
plt.plot(df.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=True)[1:11],color='r',linewidth=3,
        marker='o')
plt.title('Least 10 CO2 Emission Countries',fontsize=30,color='r',fontweight='bold')
plt.xlabel('Countries', fontsize=22,color='r',fontweight='bold' )
plt.ylabel('CO2_Emission(tonnes)', fontsize=22,color='r',fontweight='bold')
plt.legend(prop={'size':15},loc=0,facecolor='c')
plt.grid()
plt.show()
```



In [26]:
```
df.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[26]:
```
Entity
United States               3.993783e+11
EU-28                       3.528583e+11
China                       2.001365e+11
Asia and Pacific (other)    1.780573e+11
Europe (other)              1.575402e+11
Russia                      1.005891e+11
Americas (other)            9.864116e+10
Germany                     9.056563e+10
United Kingdom              7.707106e+10
Japan                       6.230461e+10
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [27]:
```
plt.figure(figsize=(20,7), facecolor='c')
plt.plot(df.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:11],color='r',linewidth=3,
        marker='o')
plt.title('top 10 CO2 Emission Countries',fontsize=30,color='r',fontweight='bold')
plt.xlabel('Countries', fontsize=22,color='r',fontweight='bold' )
plt.ylabel('CO2_Emission(tonnes)', fontsize=22,color='r',fontweight='bold')
plt.legend(prop={'size':15},loc=0,facecolor='c')
plt.grid()
plt.show()
```



Group the year in ratio 50

In [28]:
```
x1 = df['Year']<=1800
p1 = df[x1]
```

In [29]:
```
x2 = df['Year']<=1850
y2 = df['Year']>1800
p2 = df[x2&y2]
```

In [30]:
```
x3 = df['Year']<=1900
y3 = df['Year']>1850
p3 = df[x3&y3]
```

In [31]:
```
x4 = df['Year']<=1950
y4 = df['Year']>1900
p4 = df[x4&y4]
```

In [32]:
```
x5 = df['Year']<=2000
y5 = df['Year']>1950
p5 = df[x5&y5]
```

In [33]:
```
x6 = df['Year']>2000
p6 = df[x6]
```

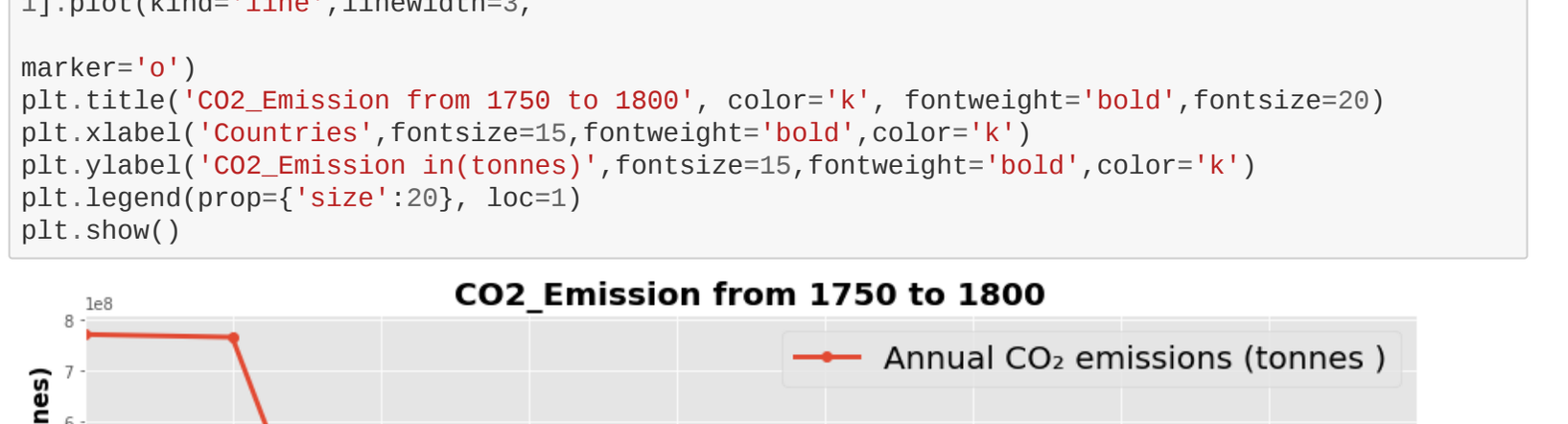
1. From year 1750 to 1800

In [34]:
```
p1.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[34]:
```
Entity
EU-28                      771850912.0
United Kingdom             766523456.0
Germany                      4920752.0
Poland                        406704.0
United States                 252816.0
Canada                         58624.0
Americas (other)               58624.0
Statistical differences            0.0
Middle East                        0.0
International transport            0.0
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [35]:
```
plt.style.use('ggplot')
plt.figure(figsize=(15,5))
p1.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1].plot(kind='line',linewidth=3,

marker='o')
plt.title('CO2_Emission from 1750 to 1800', color='k', fontweight='bold',fontsize=20)
plt.xlabel('Countries',fontsize=15,fontweight='bold',color='k')
plt.ylabel('CO2_Emission in(tonnes)',fontsize=15,fontweight='bold',color='k')
plt.legend(prop={'size':20}, loc=1)
plt.show()
```



2. From year 1800 to 1850

In [36]:
```
p2.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[36]:
```
Entity
EU-28             3.786469e+09
United Kingdom    2.927455e+09
France            3.404222e+08
Germany           2.497309e+08
United States     1.897219e+08
Belgium           1.504951e+08
Poland            6.295485e+07
Austria           3.064936e+07
Netherlands       1.685440e+07
Spain             2.740672e+06
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [37]:
```
plt.style.use('ggplot')
plt.figure(figsize=(15,5))
p2.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1].plot(kind='line',marker='o',

linewidth=3)
plt.title('CO2_Emission from 1800 to 1850', color='k', fontweight='bold',fontsize=20)
plt.xlabel('Countries',fontsize=15,fontweight='bold',color='k')
plt.ylabel('CO2_Emission in(tonnes)',fontsize=15,fontweight='bold',color='k')
plt.legend(prop={'size':20}, loc=1)
plt.show()
```

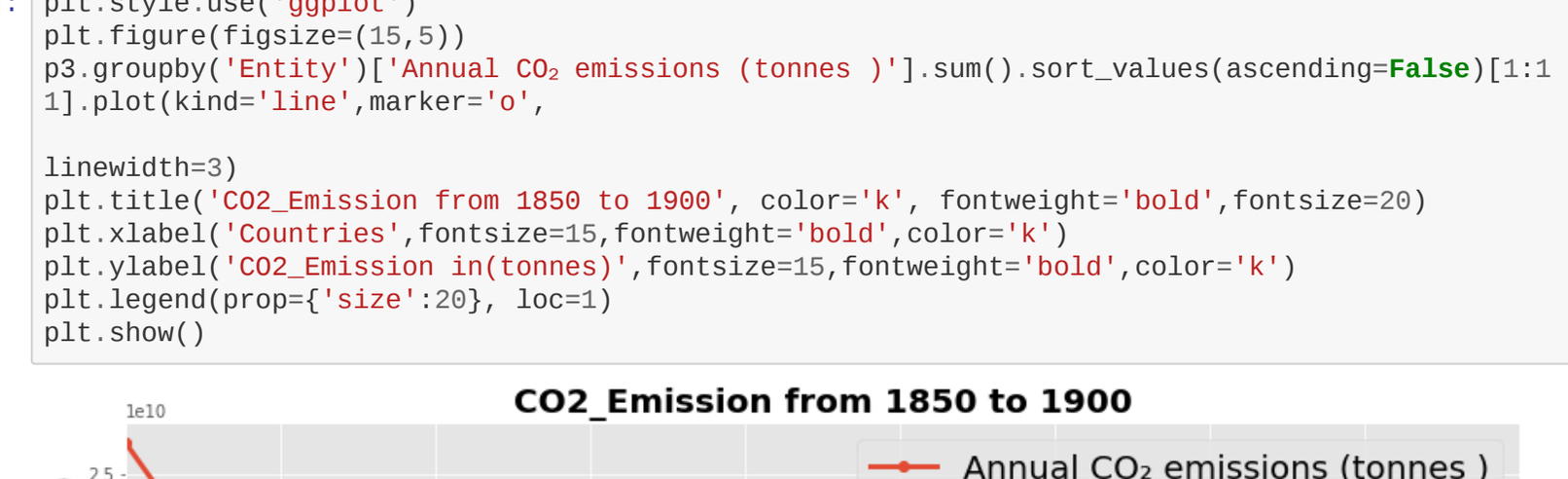
3. From the year 1850 to 1900

In [38]:
```
p3.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[38]:
```
Entity
EU-28             2.701552e+10
United Kingdom    1.304018e+10
United States     1.047446e+10
Germany           6.136203e+09
France            3.370041e+09
Europe (other)    1.776490e+09
Belgium           1.357424e+09
Poland            1.271983e+09
Czechoslovakia    8.433759e+08
Austria           5.286383e+08
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [39]:
```
plt.style.use('ggplot')
plt.figure(figsize=(15,5))
p3.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1].plot(kind='line',marker='o',

linewidth=3)
plt.title('CO2_Emission from 1850 to 1900', color='k', fontweight='bold',fontsize=20)
plt.xlabel('Countries',fontsize=15,fontweight='bold',color='k')
plt.ylabel('CO2_Emission in(tonnes)',fontsize=15,fontweight='bold',color='k')
plt.legend(prop={'size':20}, loc=1)
plt.show()
```



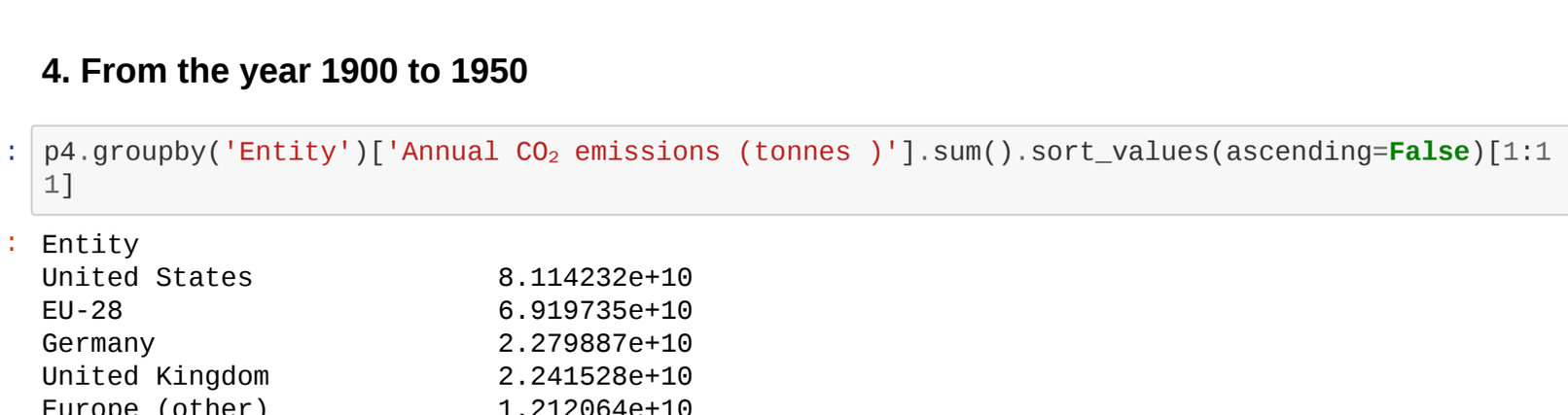
4. From the year 1900 to 1950

In [40]:
```
p4.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[40]:
```
Entity
United States               8.114232e+10
EU-28                       6.919735e+10
Germany                     2.279887e+10
United Kingdom              2.241528e+10
Europe (other)              1.212064e+10
Americas (other)            8.595019e+09
France                      8.174043e+09
Asia and Pacific (other)    6.615554e+09
Poland                      4.231121e+09
Canada                      4.224995e+09
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [41]:
```
plt.style.use('ggplot')
plt.figure(figsize=(15,5))
p4.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1].plot(kind='line',marker='o',

linewidth=3)
plt.title('CO2_Emission from 1900 to 1950', color='k', fontweight='bold',fontsize=20)
plt.xlabel('Countries',fontsize=15,fontweight='bold',color='k')
plt.ylabel('CO2_Emission in(tonnes)',fontsize=15,fontweight='bold',color='k')
plt.legend(prop={'size':20}, loc=1)
plt.show()
```



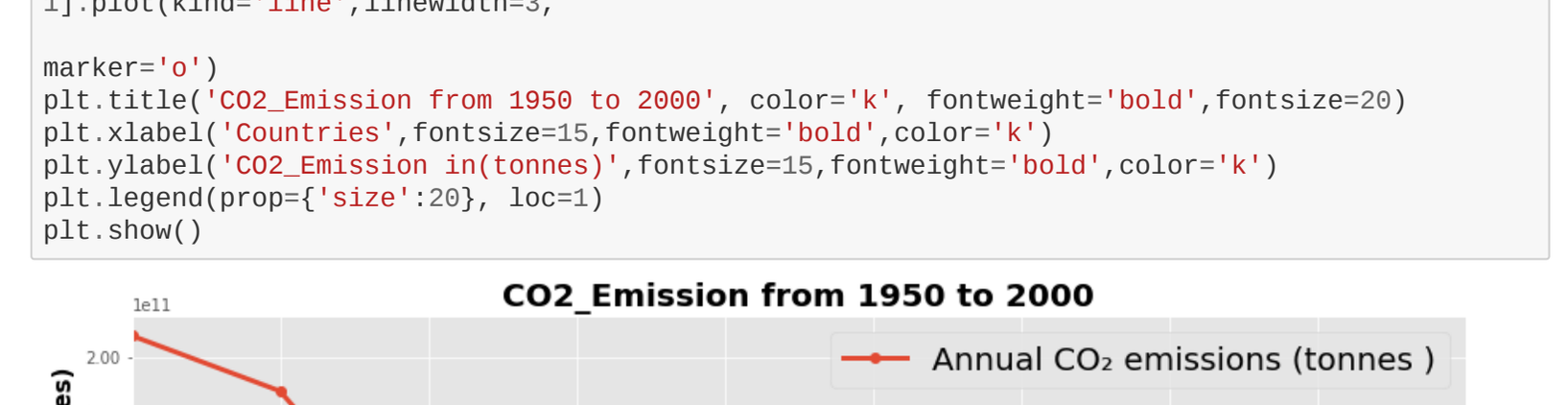
5. From the year 1950 to 2000

In [42]:
```
p5.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[42]:
```
Entity
United States               2.101561e+11
EU-28                       1.847867e+11
Europe (other)              1.068213e+11
Asia and Pacific (other)    9.525536e+10
Russia                      7.278933e+10
China                       7.093433e+10
Americas (other)            5.206207e+10
Germany                     4.704089e+10
Japan                       3.687116e+10
United Kingdom              2.931237e+10
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [43]:
```
plt.style.use('ggplot')
plt.figure(figsize=(15,5))
p5.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1].plot(kind='line',linewidth=3,

marker='o')
plt.title('CO2_Emission from 1950 to 2000', color='k', fontweight='bold',fontsize=20)
plt.xlabel('Countries',fontsize=15,fontweight='bold',color='k')
plt.ylabel('CO2_Emission in(tonnes)',fontsize=15,fontweight='bold',color='k')
plt.legend(prop={'size':20}, loc=1)
plt.show()
```



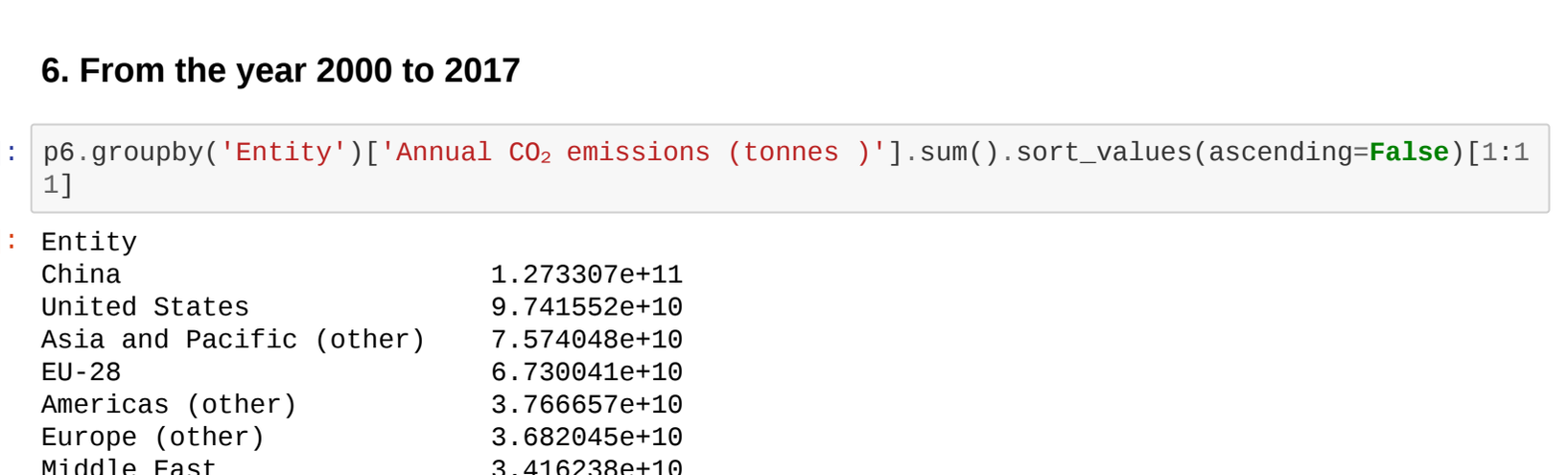
6. From the year 2000 to 2017

In [44]:
```
p6.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1]
```

Out[44]:
```
Entity
China                       1.273307e+11
United States               9.741552e+10
Asia and Pacific (other)    7.574048e+10
EU-28                       6.730041e+10
Americas (other)            3.766657e+10
Europe (other)              3.682045e+10
Middle East                 3.416238e+10
India                       2.830331e+10
Russia                      2.779977e+10
Japan                       2.131609e+10
Name: Annual CO₂ emissions (tonnes ), dtype: float64
```

In [45]:
```
plt.style.use('ggplot')
plt.figure(figsize=(15,5))
p6.groupby('Entity')['Annual CO₂ emissions (tonnes )'].sum().sort_values(ascending=False)[1:1
1].plot(kind='line',linewidth=3,

marker='o')
plt.title('CO2_Emission from 2000 to 2017', color='k', fontweight='bold',fontsize=20)
plt.xlabel('Countries',fontsize=15,fontweight='bold',color='k')
plt.ylabel('CO2_Emission in(tonnes)',fontsize=15,fontweight='bold',color='k')
plt.legend(prop={'size':20}, loc=1)
plt.show()
```



From this dataset CHINA and UNITED STATES have the highest CO2_Emission in last few year. In the Beginning of the year it has no singificant CO2-Emission. In early stage EU-28 & united have most contributed country among all.

In [ ]: