```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


data = pd.read_csv('/content/world_population.csv')
data
```

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	Popul
0	36	AFG	Afghanistan	Kabul	Asia	41128771	38972230	337
1	138	ALB	Albania	Tirana	Europe	2842321	2866849	28
2	34	DZA	Algeria	Algiers	Africa	44903225	43451666	395
3	213	ASM	American Samoa	Pago Pago	Oceania	44273	46189	
4	203	AND	Andorra	Andorra la Vella	Europe	79824	77700	
...	...	...	...	...	...	...	...	
229	226	WLF	Wallis and Futuna	Mata-Utu	Oceania	11572	11655	
230	172	ESH	Western Sahara	El Aaiún	Africa	575986	556048	4
231	46	YEM	Yemen	Sanaa	Asia	33696614	32284046	285
232	63	ZMB	Zambia	Lusaka	Africa	20017675	18927715	162
233	74	ZWE	Zimbabwe	Harare	Africa	16320537	15669666	141

234 rows × 17 columns

```
data.head()
```

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2 Populat
0	36	AFG	Afghanistan	Kabul	Asia	41128771	38972230	33753
1	138	ALB	Albania	Tirana	Europe	2842321	2866849	2882
2	34	DZA	Algeria	Algiers	Africa	44903225	43451666	39543
3	213	ASM	American Samoa	Pago Pago	Oceania	44273	46189	51
4	203	AND	Andorra	Andorra la Vella	Europe	79824	77700	71

```
data.tail()
```

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	Popul
229	226	WLF	Wallis and Futuna	Mata-Utu	Oceania	11572	11655	
230	172	ESH	Western Sahara	El Aaiún	Africa	575986	556048	4
231	46	YEM	Yemen	Sanaa	Asia	33696614	32284046	285
232	63	ZMB	Zambia	Lusaka	Africa	20017675	18927715	162
233	74	ZWE	Zimbabwe	Harare	Africa	16320537	15669666	141

```
data.shape
```

```
(234, 17)
```

```
data.dtypes
```

```
Rank                            int64
CCA3                           object
Country/Territory              object
Capital                        object
Continent                      object
2022 Population                 int64
2020 Population                 int64
2015 Population                 int64
2010 Population                 int64
2000 Population                 int64
1990 Population                 int64
1980 Population                 int64
1970 Population                 int64
Area (km²)                      int64
Density (per km²)             float64
Growth Rate                   float64
World Population Percentage   float64
dtype: object
```

```
print(data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 234 entries, 0 to 233
Data columns (total 17 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   Rank                         234 non-null    int64
 1   CCA3                         234 non-null    object
 2   Country/Territory            234 non-null    object
 3   Capital                      234 non-null    object
 4   Continent                    234 non-null    object
 5   2022 Population              234 non-null    int64
 6   2020 Population              234 non-null    int64
 7   2015 Population              234 non-null    int64
 8   2010 Population              234 non-null    int64
 9   2000 Population              234 non-null    int64
 10  1990 Population              234 non-null    int64
 11  1980 Population              234 non-null    int64
 12  1970 Population              234 non-null    int64
 13  Area (km²)                   234 non-null    int64
 14  Density (per km²)            234 non-null    float64
 15  Growth Rate                  234 non-null    float64
 16  World Population Percentage  234 non-null    float64
dtypes: float64(3), int64(10), object(4)
memory usage: 31.2+ KB
None
```

```
print(data.describe())
```

```
              Rank  2022 Population  2020 Population  2015 Population  \
count  234.000000     2.340000e+02     2.340000e+02     2.340000e+02
mean   117.500000     3.407441e+07     3.350107e+07     3.172996e+07
std     67.694165     1.367664e+08     1.355899e+08     1.304050e+08
min      1.000000     5.100000e+02     5.200000e+02     5.640000e+02
25%     59.250000     4.197385e+05     4.152845e+05     4.046760e+05
50%    117.500000     5.559944e+06     5.493074e+06     5.307400e+06
75%    175.750000     2.247650e+07     2.144798e+07     1.973085e+07
max    234.000000     1.425887e+09     1.424930e+09     1.393715e+09

       2010 Population  2000 Population  1990 Population  1980 Population  \
count     2.340000e+02     2.340000e+02     2.340000e+02     2.340000e+02
mean      2.984524e+07     2.626947e+07     2.271022e+07     1.898462e+07
std       1.242185e+08     1.116982e+08     9.783217e+07     8.178519e+07
min       5.960000e+02     6.510000e+02     7.000000e+02     7.330000e+02
25%       3.931490e+05     3.272420e+05     2.641158e+05     2.296142e+05
50%       4.942770e+06     4.292907e+06     3.825410e+06     3.141146e+06
75%       1.915957e+07     1.576230e+07     1.186923e+07     9.826054e+06
max       1.348191e+09     1.264099e+09     1.153704e+09     9.823725e+08

       1970 Population    Area (km²)  Density (per km²)  Growth Rate  \
count     2.340000e+02  2.340000e+02         234.000000   234.000000
mean      1.578691e+07  5.814494e+05         452.127044     1.009577
std       6.779509e+07  1.761841e+06        2066.121904     0.013385
min       7.520000e+02  1.000000e+00           0.026100     0.912000
25%       1.559970e+05  2.650000e+03          38.417875     1.001775
50%       2.604830e+06  8.119950e+04          95.346750     1.007900
75%       8.817329e+06  4.304258e+05         238.933250     1.016950
max       8.225344e+08  1.709824e+07       23172.266700     1.069100

       World Population Percentage
count                   234.000000
mean                      0.427051
std                       1.714977
min                       0.000000
25%                       0.010000
50%                       0.070000
75%                       0.280000
max                      17.880000
```

```
col = data['Capital']
print(col)
```

```
0                Kabul
1               Tirana
2              Algiers
3            Pago Pago
4      Andorra la Vella
             ...
229           Mata-Utu
230           El Aaiún
231              Sanaa
232             Lusaka
233             Harare
Name: Capital, Length: 234, dtype: object
```

```
a1 = data[data['Rank']> 200]
print(a1)
```

```
78            56184           55599           50106           45434
83            59114           57727           52860           52656
97            75562           68865           64022           55298
116           33026           28765           25003           21089
126           54224           46047           31988           23969
134           32465           30329           27076           24270
137            5138           10805           11452           11402
142           10377            9598            7635            6663
150            2074            2533            3637            5185
153           80338           48002           17613           10143
157           19726           15293           12252           11366
173            7082            5168            2983            2417
174           45461           40636           43097           44968
176           29610           28127            7776            5802
177            6274            6324            6106            5537
180           26823           23132           21346           18169
188           30489           27845           12243            6260
209            1666            1669            1647            1714
215           18744           11709            7598            5665
216            9638            9182            7731            5814
226             651             700             733             752
229           14723           13454           11315            9377

     Area (km²)  Density (per km²)  Growth Rate  World Population Percentage
3           199           222.4774       0.9831                          0.0
4           468           170.5641       1.0100                          0.0
6            91           174.2527       1.0066                          0.0
7           442           212.1335       1.0058                          0.0
22           54          1188.5926       1.0000                          0.0
28          151           207.3179       1.0059                          0.0
37          264           260.2500       1.0084                          0.0
44          236            72.0805       1.0005                          0.0
53          751            96.8535       1.0045                          0.0
64        12173             0.3105       1.0043                          0.0
65         1393            38.1120       1.0038                          0.0
76            6          5441.5000       0.9994                          0.0
78      2166086             0.0261       1.0040                          0.0
83           78           811.5513       1.0037                          0.0
97          572           147.7605       1.0030                          0.0
116         160           245.7937       1.0074                          0.0
126         181           229.6630       0.9886                          0.0
134           2         18234.5000       0.9941                          0.0
137         102            43.0392       0.9939                          0.0
142          21           603.2381       1.0125                          0.0
150         260             7.4385       0.9985                          0.0
153         464           106.7909       1.0014                          0.0
157         459            39.3355       1.0017                          0.0
173          21           522.2381       1.0098                          0.0
174         261           182.5939       1.0011                          0.0
176          53           599.8302       0.9951                          0.0
177         242            24.2231       0.9964                          0.0
180          61           551.8033       0.9975                          0.0
188          34          1299.2647       1.0030                          0.0
209          12           155.9167       1.0119                          0.0
215         948            48.2099       1.0131                          0.0
216          26           435.0769       1.0096                          0.0
226           1           510.0000       0.9980                          0.0
229         142            81.4930       0.9953                          0.0
```

```
data['Population_Growth'] = data['2022 Population'] - data['2010 Population']
print(data['Population_Growth'])
```

```
0      12939099
1        -71078
2       9046881
3        -10576
4          8305
```

```
        ...
229      -1570
230     162690
231    8952668
232    6225589
233    3480766
Name: Population_Growth, Length: 234, dtype: int64
```

```
dup = data[data.duplicated()]
print(dup)
```

```
Empty DataFrame
Columns: [Rank, CCA3, Country/Territory, Capital, Continent, 2022 Population, 2020 Population, 2015 Population, 2010 Population, 20(
Index: []
```

```
data.drop_duplicates(inplace=True)
```

```
data.shape
```

```
(234, 18)
```

```
missing_val = data.isnull().sum()
print(missing_val)
```

```
Rank                           0
CCA3                           0
Country/Territory              0
Capital                        0
Continent                      0
2022 Population                0
2020 Population                0
2015 Population                0
2010 Population                0
2000 Population                0
1990 Population                0
1980 Population                0
1970 Population                0
Area (km²)                     0
Density (per km²)              0
Growth Rate                    0
World Population Percentage    0
Population_Growth              0
dtype: int64
```

```
count = data[data['Country/Territory'].str.contains('United')]
print(count)
```

```
     Rank CCA3             Country/Territory           Capital      Continent  \
219    97  ARE          United Arab Emirates         Abu Dhabi           Asia
220    21  GBR                United Kingdom            London         Europe
221     3  USA                 United States  Washington, D.C.  North America
222   200  VIR  United States Virgin Islands  Charlotte Amalie  North America

     2022 Population  2020 Population  2015 Population  2010 Population  \
219          9441129          9287289          8916899          8481771
220         67508936         67059474         65224364         62760039
221        338289857        335942003        324607776        311182845
222            99465           100442           102803           106142

     2000 Population  1990 Population  1980 Population  1970 Population  \
219          3275333          1900151          1014048           298084
220         58850043         57210442         56326328         55650166
221        282398554        248083732        223140018        200328340
222           108185           100685            96640            63446

     Area (km²)  Density (per km²)  Growth Rate  World Population Percentage  \
219       83600           112.9322       1.0081                         0.12
220      242900           277.9289       1.0034                         0.85
221     9372610            36.0935       1.0038                         4.24
222         347           286.6427       0.9937                         0.00

     Population_Growth
219             959358
220            4748897
221           27107012
222              -6677
```

```
cl = data.dropna()
print(cl)
```

```
     Rank CCA3  Country/Territory           Capital Continent  \
0      36  AFG        Afghanistan             Kabul      Asia
1     138  ALB            Albania            Tirana    Europe
2      34  DZA            Algeria           Algiers    Africa
3     213  ASM     American Samoa         Pago Pago   Oceania
4     203  AND            Andorra  Andorra la Vella    Europe
..    ...  ...                ...               ...       ...
229   226  WLF  Wallis and Futuna          Mata-Utu   Oceania
230   172  ESH     Western Sahara          El Aaiún    Africa
231    46  YEM              Yemen             Sanaa      Asia
232    63  ZMB             Zambia            Lusaka    Africa
233    74  ZWE           Zimbabwe            Harare    Africa

     2022 Population  2020 Population  2015 Population  2010 Population  \
0           41128771         38972230         33753499         28189672
1            2842321          2866849          2882481          2913399
2           44903225         43451666         39543154         35856344
3              44273            46189            51368            54849
4              79824            77700            71746            71519
..               ...              ...              ...              ...
229            11572            11655            12182            13142
230           575986           556048           491824           413296
231         33696614         32284046         28516545         24743946
232         20017675         18927715         16248230         13792086
233         16320537         15669666         14154937         12839771

     2000 Population  1990 Population  1980 Population  1970 Population  \
0           19542982         10694796         12486631         10752971
1            3182021          3295066          2941651          2324731
2           30774621         25518074         18739378         13795915
3              58230            47818            32886            27075
4              66097            53569            35611            19860
..               ...              ...              ...              ...
229            14723            13454            11315             9377
230           270375           178529           116775            76371
231         18628700         13375121          9204938          6843607
232          9891136          7686401          5720438          4281671
233         11834676         10113893          7049926          5202918

     Area (km²)  Density (per km²)  Growth Rate  World Population Percentage  \
0        652230            63.0587       1.0257                         0.52
1         28748            98.8702       0.9957                         0.04
2       2381741            18.8531       1.0164                         0.56
3           199           222.4774       0.9831                         0.00
4           468           170.5641       1.0100                         0.00
..          ...                ...          ...                          ...
229         142            81.4930       0.9953                         0.00
230      266000             2.1654       1.0184                         0.01
231      527968            63.8232       1.0217                         0.42
232      752612            26.5976       1.0280                         0.25
233      390757            41.7665       1.0204                         0.20

     Population_Growth
0             12939099
1               -71078
2              9046881
3               -10576
4                 8305
```

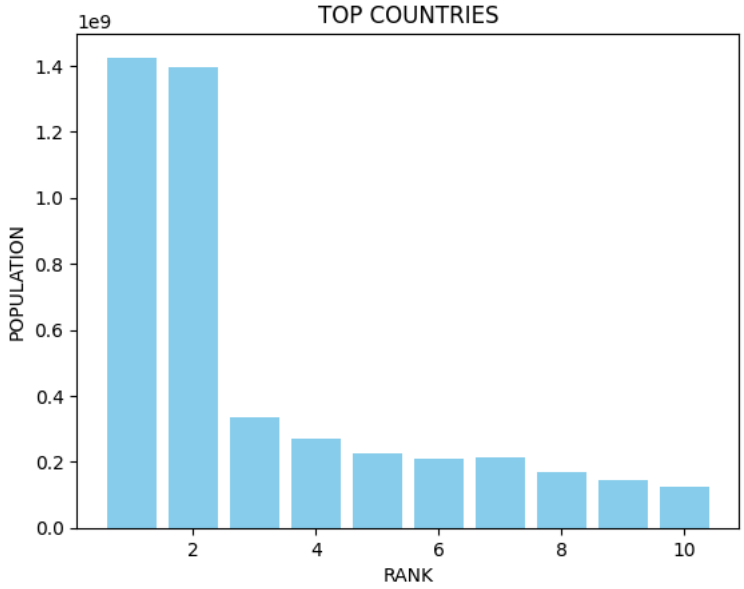
```
con = data['Continent'].value_counts()
print(con)
print()
dl = data['2022 Population'].value_counts()
print(dl)
```

```
Continent
Africa           57
Asia             50
Europe           50
North America    40
Oceania          23
South America    14
Name: count, dtype: int64

2022 Population
41128771     1
10142619     1
26207977     1
218541212    1
1934         1
            ..
17843908     1
63301        1
13859341     1
2105566      1
16320537     1
Name: count, Length: 234, dtype: int64
```

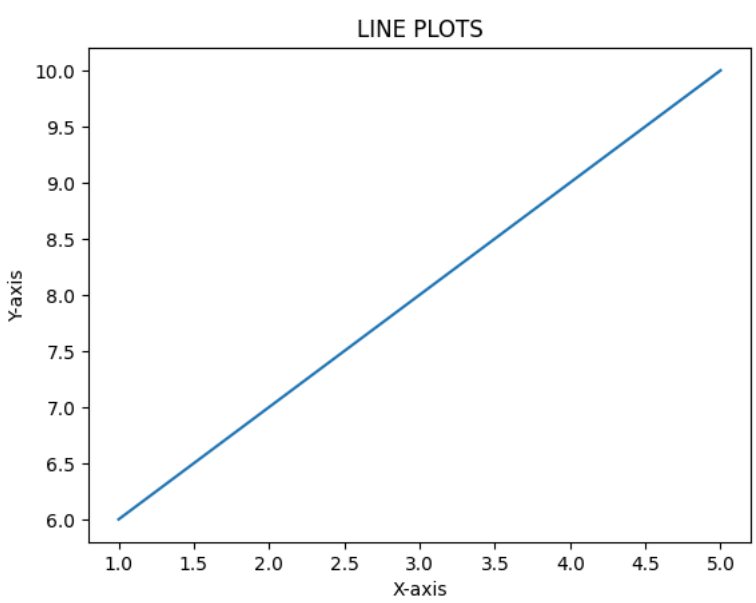
```
#sorting
sd = data.sort_values(by='2022 Population',ascending = False)
print(sd)
```

```
      Rank CCA3 Country/Territory           Capital      Continent  \
41       1  CHN             China           Beijing           Asia
92       2  IND             India         New Delhi           Asia
221      3  USA     United States  Washington, D.C.  North America
93       4  IDN         Indonesia           Jakarta           Asia
156      5  PAK          Pakistan         Islamabad           Asia
..     ...  ...               ...               ...            ...
137    230  MSR        Montserrat            Brades  North America
64     231  FLK  Falkland Islands           Stanley  South America
150    232  NIU              Niue             Alofi        Oceania
209    233  TKL           Tokelau          Nukunonu        Oceania
226    234  VAT      Vatican City      Vatican City         Europe

      2022 Population  2020 Population  2015 Population  2010 Population  \
41         1425887337       1424929781       1393715448       1348191368
92         1417173173       1396387127       1322866505       1240613620
221         338289857        335942003        324607776        311182845
93          275501339        271857970        259091970        244016173
156         235824862        227196741        210969298        194454498
..                ...              ...              ...              ...
137              4390             4500             5059             4938
64               3780             3747             3408             3187
150              1934             1942             1847             1812
209              1871             1827             1454             1367
226               510              520              564              596

      2000 Population  1990 Population  1980 Population  1970 Population  \
41         1264099069       1153704252        982372466        822534450
92         1059633675        870452165        696828385        557501301
221         282398554        248083732        223140018        200328340
93          214072421        182159874        148177096        115228394
156         154369924        115414069         80624057         59290872
..                ...              ...              ...              ...
137              5138            10805            11452            11402
64               3080             2332             2240             2274
150              2074             2533             3637             5185
209              1666             1669             1647             1714
226               651              700              733              752

      Area (km²)  Density (per km²)  Growth Rate  World Population Percentage  \
41       9706961           146.8933       1.0000                        17.88
92       3287590           431.0675       1.0068                        17.77
221      9372610            36.0935       1.0038                         4.24
93       1904569           144.6529       1.0064                         3.45
156       881912           267.4018       1.0191                         2.96
..           ...                ...          ...                          ...
137          102            43.0392       0.9939                         0.00
64         12173             0.3105       1.0043                         0.00
150          260             7.4385       0.9985                         0.00
209           12           155.9167       1.0119                         0.00
226            1           510.0000       0.9980                         0.00

      Population_Growth
41             77695969
92            176559553
221            27107012
93             31485166
156            41370364
```

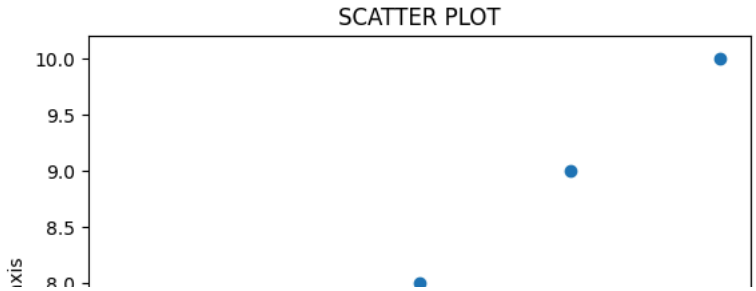
```
#visualization

tp = data.nlargest(10,'2020 Population')
plt.bar(tp['Rank'],tp['2020 Population'],color='skyblue')
plt.xlabel('RANK')
plt.ylabel('POPULATION')
plt.title('TOP COUNTRIES')
plt.show()
```



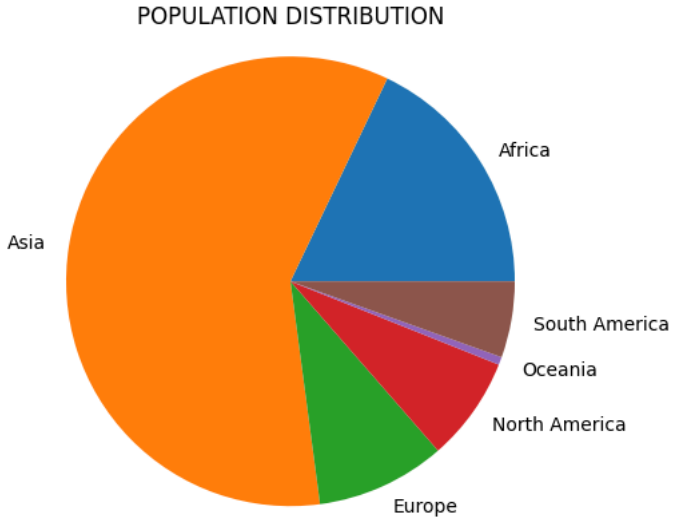

```
x = [1,2,3,4,5]
y = [6,7,8,9,10]

plt.plot(x,y)
plt.xlabel('X-axis')
plt.ylabel('Y-axis')
plt.title('LINE PLOTS')
plt.show()
```




```
x = [1,2,3,4,5]
y = [6,7,8,9,10]

plt.scatter(x,y)
plt.xlabel('X-axis')
plt.ylabel('Y-axis')
plt.title('SCATTER PLOT')
plt.show()
```




```
cp = data.groupby('Continent')['2022 Population'].sum()
plt.pie(cp,labels=cp.index)
plt.axis('equal')
plt.title('POPULATION DISTRIBUTION')
plt.show()
```




```
ye = data.nlargest(5, '2020 Population')
po = ['2020','2022','2032','2019','2017']
plt.plot(ye['2020 Population'], po, marker='o', ls='-', color='b', linewidth=2, ms=8)
plt.xlabel('Year')
plt.ylabel('Population')
plt.title('Population Over Time')
plt.grid(True)
plt.show()
```