

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import plotly.express as ppx
from scipy import stats
```

```
df=pd.read_csv(r"C:\Users\HP\Downloads\PandasYouTubeSeries-331621f89f679892d18af059ea9061f91321a934\PandasYouTubeSeries-331621f89f679892d18af059ea9061f91321a934\world_population.csv")
```

```
df.head()
```

|              | Rank | CCA3 | Country        | Capital          | Continent | 2022       |
|--------------|------|------|----------------|------------------|-----------|------------|
| Population \ |      |      |                |                  |           |            |
| 0            | 36   | AFG  | Afghanistan    | Kabul            | Asia      | 41128771.0 |
| 1            | 138  | ALB  | Albania        | Tirana           | Europe    | 2842321.0  |
| 2            | 34   | DZA  | Algeria        | Algiers          | Africa    | 44903225.0 |
| 3            | 213  | ASM  | American Samoa | Pago Pago        | Oceania   | 44273.0    |
| 4            | 203  | AND  | Andorra        | Andorra la Vella | Europe    | 79824.0    |

|   | 2020 | Population | 2015 | Population | 2010 | Population | 2000 | Population |
|---|------|------------|------|------------|------|------------|------|------------|
| \ |      |            |      |            |      |            |      |            |
| 0 |      | 38972230.0 |      | 33753499.0 |      | 28189672.0 |      | 19542982.0 |
| 1 |      | 2866849.0  |      | 2882481.0  |      | 2913399.0  |      | 3182021.0  |
| 2 |      | 43451666.0 |      | 39543154.0 |      | 35856344.0 |      | 30774621.0 |
| 3 |      | 46189.0    |      | 51368.0    |      | 54849.0    |      | 58230.0    |
| 4 |      | 77700.0    |      | 71746.0    |      | 71519.0    |      | 66097.0    |

|   | 1990 | Population | 1980 | Population | 1970 | Population | Area (km <sup>2</sup> ) | \ |
|---|------|------------|------|------------|------|------------|-------------------------|---|
| 0 |      | 10694796.0 |      | 12486631.0 |      | 10752971.0 | 652230.0                |   |
| 1 |      | 3295066.0  |      | 2941651.0  |      | 2324731.0  | 28748.0                 |   |
| 2 |      | 25518074.0 |      | 18739378.0 |      | 13795915.0 | 2381741.0               |   |
| 3 |      | 47818.0    |      | 32886.0    |      | 27075.0    | 199.0                   |   |
| 4 |      | 53569.0    |      | 35611.0    |      | 19860.0    | 468.0                   |   |

|   | Density (per km <sup>2</sup> ) | Growth Rate | World Population Percentage |
|---|--------------------------------|-------------|-----------------------------|
| 0 | 63.0587                        | 1.0257      | 0.52                        |
| 1 | 98.8702                        | 0.9957      | 0.04                        |
| 2 | 18.8531                        | 1.0164      | 0.56                        |

```
3          222.4774          0.9831          0.00
4          170.5641          1.0100          0.00
```

```
# lets format the decimal value i will set upto two decimal place
pd.set_option('display.float_format',lambda x:'%.2f' % x)
```

```
df
```

|     | Rank | CCA3 | Country           | 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            | 41128771.00     | 38972230.00     | 33753499.00     |      |
| 28189672.00  |                 |                 |                 |      |
| 1            | 2842321.00      | 2866849.00      | 2882481.00      |      |
| 2913399.00   |                 |                 |                 |      |
| 2            | 44903225.00     | 43451666.00     | 39543154.00     |      |
| 35856344.00  |                 |                 |                 |      |
| 3            | 44273.00        | 46189.00        | 51368.00        |      |
| 54849.00     |                 |                 |                 |      |
| 4            | 79824.00        | 77700.00        | 71746.00        |      |
| 71519.00     |                 |                 |                 |      |
| ...          | ...             | ...             | ...             | ..   |
| .            |                 |                 |                 |      |
| 229          | 11572.00        | 11655.00        | 12182.00        |      |
| 13142.00     |                 |                 |                 |      |
| 230          | 575986.00       | 556048.00       | 491824.00       |      |
| 413296.00    |                 |                 |                 |      |
| 231          | 33696614.00     | 32284046.00     | 28516545.00     |      |
| 24743946.00  |                 |                 |                 |      |
| 232          | 20017675.00     | 18927715.00     | NaN             |      |
| 13792086.00  |                 |                 |                 |      |
| 233          | 16320537.00     | 15669666.00     | 14154937.00     |      |
| 12839771.00  |                 |                 |                 |      |

|              | 2000 Population | 1990 Population | 1980 Population | 1970 |
|--------------|-----------------|-----------------|-----------------|------|
| Population \ |                 |                 |                 |      |
| 0            | 19542982.00     | 10694796.00     | 12486631.00     |      |
| 10752971.00  |                 |                 |                 |      |

```

1      3182021.00      3295066.00      2941651.00
2324731.00
2      30774621.00      25518074.00      18739378.00
13795915.00
3      58230.00      47818.00      32886.00
27075.00
4      66097.00      53569.00      35611.00
19860.00
..      ...      ...      ...
.
229      14723.00      13454.00      11315.00
9377.00
230      270375.00      178529.00      116775.00
76371.00
231      18628700.00      13375121.00      9204938.00
6843607.00
232      9891136.00      7686401.00      5720438.00
4281671.00
233      11834676.00      10113893.00      7049926.00
5202918.00

```

```

      Area (km²) Density (per km²) Growth Rate World Population
Percentage
0      652230.00      63.06      1.03
0.52
1      28748.00      98.87      1.00
0.04
2      2381741.00      18.85      1.02
0.56
3      199.00      222.48      0.98
0.00
4      468.00      170.56      1.01
0.00
..      ...      ...      ...
...
229      142.00      81.49      1.00
0.00
230      266000.00      2.17      1.02
0.01
231      527968.00      63.82      1.02
0.42
232      752612.00      26.60      1.03
0.25
233      390757.00      41.77      1.02
0.20

```

```
[234 rows x 17 columns]
```

```
df.info() # for getting a view of dataset the value types and null
```

```
<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                        | 234 non-null   | object  |
| 3  | Capital                        | 234 non-null   | object  |
| 4  | Continent                      | 234 non-null   | object  |
| 5  | 2022 Population                | 230 non-null   | float64 |
| 6  | 2020 Population                | 233 non-null   | float64 |
| 7  | 2015 Population                | 230 non-null   | float64 |
| 8  | 2010 Population                | 227 non-null   | float64 |
| 9  | 2000 Population                | 227 non-null   | float64 |
| 10 | 1990 Population                | 229 non-null   | float64 |
| 11 | 1980 Population                | 229 non-null   | float64 |
| 12 | 1970 Population                | 230 non-null   | float64 |
| 13 | Area (km <sup>2</sup> )        | 232 non-null   | float64 |
| 14 | Density (per km <sup>2</sup> ) | 230 non-null   | float64 |
| 15 | Growth Rate                    | 232 non-null   | float64 |
| 16 | World Population Percentage    | 234 non-null   | float64 |

```
dtypes: float64(12), int64(1), object(4)
```

```
memory usage: 31.2+ KB
```

```
df.isnull().sum() # function to check how many null values are in each column
```

|                                |   |
|--------------------------------|---|
| Rank                           | 0 |
| CCA3                           | 0 |
| Country                        | 0 |
| Capital                        | 0 |
| Continent                      | 0 |
| 2022 Population                | 4 |
| 2020 Population                | 1 |
| 2015 Population                | 4 |
| 2010 Population                | 7 |
| 2000 Population                | 7 |
| 1990 Population                | 5 |
| 1980 Population                | 5 |
| 1970 Population                | 4 |
| Area (km <sup>2</sup> )        | 2 |
| Density (per km <sup>2</sup> ) | 4 |
| Growth Rate                    | 2 |
| World Population Percentage    | 0 |

```
dtype: int64
```

```
df.describe() # to display the overall relation between the numeric columns
```

|              |                             |                         |                                |                 |   |
|--------------|-----------------------------|-------------------------|--------------------------------|-----------------|---|
|              | Rank                        | 2022 Population         | 2020 Population                | 2015 Population | \ |
| count        | 234.00                      | 230.00                  | 233.00                         | 230.00          |   |
| mean         | 117.50                      | 34632250.88             | 33600710.95                    | 32066004.16     |   |
| std          | 67.69                       | 137889172.44            | 135873196.61                   | 131507146.34    |   |
| min          | 1.00                        | 510.00                  | 520.00                         | 564.00          |   |
| 25%          | 59.25                       | 419738.50               | 406471.00                      | 394295.00       |   |
| 50%          | 117.50                      | 5762857.00              | 5456681.00                     | 5244415.00      |   |
| 75%          | 175.75                      | 22653719.00             | 21522626.00                    | 19730853.75     |   |
| max          | 234.00                      | 1425887337.00           | 1424929781.00                  | 1393715448.00   |   |
|              |                             |                         |                                |                 |   |
|              | 2010 Population             | 2000 Population         | 1990 Population                | 1980            |   |
| Population   | \                           |                         |                                |                 |   |
| count        | 227.00                      | 227.00                  | 229.00                         |                 |   |
| 229.00       |                             |                         |                                |                 |   |
| mean         | 30270164.48                 | 26840495.26             | 19330463.93                    |                 |   |
| 16282884.78  |                             |                         |                                |                 |   |
| std          | 126074183.54                | 113352454.57            | 81309624.96                    |                 |   |
| 69345465.54  |                             |                         |                                |                 |   |
| min          | 596.00                      | 651.00                  | 700.00                         |                 |   |
| 733.00       |                             |                         |                                |                 |   |
| 25%          | 382726.50                   | 329470.00               | 261928.00                      |                 |   |
| 223752.00    |                             |                         |                                |                 |   |
| 50%          | 4889741.00                  | 4491202.00              | 3785847.00                     |                 |   |
| 3135123.00   |                             |                         |                                |                 |   |
| 75%          | 16825852.50                 | 15625467.00             | 11882762.00                    |                 |   |
| 9817257.00   |                             |                         |                                |                 |   |
| max          | 1348191368.00               | 1264099069.00           | 1153704252.00                  |                 |   |
| 982372466.00 |                             |                         |                                |                 |   |
|              |                             |                         |                                |                 |   |
|              | 1970 Population             | Area (km <sup>2</sup> ) | Density (per km <sup>2</sup> ) | Growth Rate     | \ |
| count        | 230.00                      | 232.00                  | 230.00                         | 232.00          |   |
| mean         | 15866499.13                 | 581663.75               | 456.81                         | 1.01            |   |
| std          | 68355859.75                 | 1769133.06              | 2083.74                        | 0.01            |   |
| min          | 752.00                      | 1.00                    | 0.03                           | 0.91            |   |
| 25%          | 145880.50                   | 2567.25                 | 36.60                          | 1.00            |   |
| 50%          | 2511718.00                  | 77141.00                | 95.35                          | 1.01            |   |
| 75%          | 8817329.00                  | 414643.25               | 236.88                         | 1.02            |   |
| max          | 822534450.00                | 17098242.00             | 23172.27                       | 1.07            |   |
|              |                             |                         |                                |                 |   |
|              | World Population Percentage |                         |                                |                 |   |
| count        | 234.00                      |                         |                                |                 |   |
| mean         | 0.43                        |                         |                                |                 |   |
| std          | 1.71                        |                         |                                |                 |   |
| min          | 0.00                        |                         |                                |                 |   |
| 25%          | 0.01                        |                         |                                |                 |   |
| 50%          | 0.07                        |                         |                                |                 |   |
| 75%          | 0.28                        |                         |                                |                 |   |
| max          | 17.88                       |                         |                                |                 |   |

```
#lets check how many unique value our data ste have
df.nunique()
```

```
Rank                234
CCA3                234
Country            234
Capital            234
Continent           6
2022 Population    230
2020 Population    233
2015 Population    230
2010 Population    227
2000 Population    227
1990 Population    229
1980 Population    229
1970 Population    230
Area (km²)         231
Density (per km²)  230
Growth Rate        178
World Population Percentage  70
dtype: int64
```

```
# lets see that top 5 highest population values in year 2022
df.sort_values("2022 Population",ascending=False).head(5)
```

|     | Rank | CCA3 | Country       | 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          |   |

|     | 2022 Population | 2020 Population | 2015 Population | 2010 Population | \ |
|-----|-----------------|-----------------|-----------------|-----------------|---|
| 41  | 1425887337.00   | 1424929781.00   | 1393715448.00   | 1348191368.00   |   |
| 92  | 1417173173.00   | 1396387127.00   | 1322866505.00   | 1240613620.00   |   |
| 221 | 338289857.00    | 335942003.00    | 324607776.00    | 311182845.00    |   |
| 93  | 275501339.00    | 271857970.00    | 259091970.00    | 244016173.00    |   |
| 156 | 235824862.00    | 227196741.00    | 210969298.00    | 194454498.00    |   |

|    | 2000 Population | 1990 Population | 1980 Population | 1970 Population | \ |
|----|-----------------|-----------------|-----------------|-----------------|---|
| 41 | 1264099069.00   | 1153704252.00   | 982372466.00    |                 |   |

```

822534450.00
92      1059633675.00      NaN      NaN
557501301.00
221      282398554.00      248083732.00      223140018.00
200328340.00
93      214072421.00      182159874.00      148177096.00
115228394.00
156      154369924.00      115414069.00      80624057.00
59290872.00

```

|            | Area (km <sup>2</sup> ) | Density (per km <sup>2</sup> ) | Growth Rate | World Population |
|------------|-------------------------|--------------------------------|-------------|------------------|
| Percentage |                         |                                |             |                  |
| 41         | 9706961.00              | 146.89                         | 1.00        |                  |
| 17.88      |                         |                                |             |                  |
| 92         | 3287590.00              | 431.07                         | 1.01        |                  |
| 17.77      |                         |                                |             |                  |
| 221        | 9372610.00              | 36.09                          | 1.00        |                  |
| 4.24       |                         |                                |             |                  |
| 93         | 1904569.00              | 144.65                         | 1.01        |                  |
| 3.45       |                         |                                |             |                  |
| 156        | 881912.00               | 267.40                         | 1.02        |                  |
| 2.96       |                         |                                |             |                  |

```

# lets see the correlation for each and very numeric column its an
important part in eda
# first we will filter out the rows from original dataset i mean
columns thta are numeric
numeric=df.select_dtypes(include='number')
numeric

```

|     | Rank | 2022 Population | 2020 Population | 2015 Population | 2010 Population |
|-----|------|-----------------|-----------------|-----------------|-----------------|
| 0   | 36   | 41128771.00     | 38972230.00     | 33753499.00     | 28189672.00     |
| 1   | 138  | 2842321.00      | 2866849.00      | 2882481.00      | 2913399.00      |
| 2   | 34   | 44903225.00     | 43451666.00     | 39543154.00     | 35856344.00     |
| 3   | 213  | 44273.00        | 46189.00        | 51368.00        | 54849.00        |
| 4   | 203  | 79824.00        | 77700.00        | 71746.00        | 71519.00        |
| ..  | ...  | ...             | ...             | ...             | ...             |
| ... |      |                 |                 |                 |                 |
| 229 | 226  | 11572.00        | 11655.00        | 12182.00        | 13142.00        |
| 230 | 172  | 575986.00       | 556048.00       | 491824.00       | 413296.00       |
| 231 | 46   | 33696614.00     | 32284046.00     | 28516545.00     | 24743946.00     |

|             |    |             |             |             |
|-------------|----|-------------|-------------|-------------|
| 232         | 63 | 20017675.00 | 18927715.00 | NaN         |
| 13792086.00 |    |             |             |             |
| 233         | 74 | 16320537.00 | 15669666.00 | 14154937.00 |
| 12839771.00 |    |             |             |             |

|   | 2000 Population | 1990 Population | 1980 Population | 1970 Population \ |
|---|-----------------|-----------------|-----------------|-------------------|
| 0 | 19542982.00     | 10694796.00     | 12486631.00     | 10752971.00       |
| 1 | 3182021.00      | 3295066.00      | 2941651.00      | 2324731.00        |
| 2 | 30774621.00     | 25518074.00     | 18739378.00     | 13795915.00       |
| 3 | 58230.00        | 47818.00        | 32886.00        | 27075.00          |
| 4 | 66097.00        | 53569.00        | 35611.00        | 19860.00          |

|     |             |             |            |            |
|-----|-------------|-------------|------------|------------|
| ..  | ...         | ...         | ...        | ..         |
| .   |             |             |            |            |
| 229 | 14723.00    | 13454.00    | 11315.00   | 9377.00    |
| 230 | 270375.00   | 178529.00   | 116775.00  | 76371.00   |
| 231 | 18628700.00 | 13375121.00 | 9204938.00 | 6843607.00 |
| 232 | 9891136.00  | 7686401.00  | 5720438.00 | 4281671.00 |
| 233 | 11834676.00 | 10113893.00 | 7049926.00 | 5202918.00 |

|     | Area (km <sup>2</sup> ) | Density (per km <sup>2</sup> ) | Growth Rate | World Population Percentage |
|-----|-------------------------|--------------------------------|-------------|-----------------------------|
| 0   | 652230.00               | 63.06                          | 1.03        | 0.52                        |
| 1   | 28748.00                | 98.87                          | 1.00        | 0.04                        |
| 2   | 2381741.00              | 18.85                          | 1.02        | 0.56                        |
| 3   | 199.00                  | 222.48                         | 0.98        | 0.00                        |
| 4   | 468.00                  | 170.56                         | 1.01        | 0.00                        |
| ..  | ...                     | ...                            | ...         | ...                         |
| ... |                         |                                |             |                             |
| 229 | 142.00                  | 81.49                          | 1.00        | 0.00                        |
| 230 | 266000.00               | 2.17                           | 1.02        | 0.01                        |
| 231 | 527968.00               | 63.82                          | 1.02        |                             |



```

0.42
232 752612.00          26.60          1.03
0.25
233 390757.00          41.77          1.02
0.20

```

```
[234 rows x 13 columns]
```

```
numeric.corr()
```

|                                | Rank  | 2022 Population | 2020 Population \ |
|--------------------------------|-------|-----------------|-------------------|
| Rank                           | 1.00  | -0.36           | -0.36             |
| 2022 Population                | -0.36 | 1.00            | 1.00              |
| 2020 Population                | -0.36 | 1.00            | 1.00              |
| 2015 Population                | -0.35 | 1.00            | 1.00              |
| 2010 Population                | -0.35 | 1.00            | 1.00              |
| 2000 Population                | -0.34 | 0.99            | 1.00              |
| 1990 Population                | -0.33 | 0.99            | 0.99              |
| 1980 Population                | -0.33 | 0.99            | 0.99              |
| 1970 Population                | -0.34 | 0.97            | 0.98              |
| Area (km <sup>2</sup> )        | -0.38 | 0.45            | 0.45              |
| Density (per km <sup>2</sup> ) | 0.13  | -0.03           | -0.03             |
| Growth Rate                    | -0.22 | -0.02           | -0.03             |
| World Population Percentage    | -0.36 | 1.00            | 1.00              |

|                                | 2015 Population | 2010 Population \ |
|--------------------------------|-----------------|-------------------|
| Rank                           | -0.35           | -0.35             |
| 2022 Population                | 1.00            | 1.00              |
| 2020 Population                | 1.00            | 1.00              |
| 2015 Population                | 1.00            | 1.00              |
| 2010 Population                | 1.00            | 1.00              |
| 2000 Population                | 1.00            | 1.00              |
| 1990 Population                | 0.99            | 1.00              |
| 1980 Population                | 0.99            | 0.99              |
| 1970 Population                | 0.98            | 0.98              |
| Area (km <sup>2</sup> )        | 0.46            | 0.46              |
| Density (per km <sup>2</sup> ) | -0.03           | -0.03             |
| Growth Rate                    | -0.03           | -0.04             |
| World Population Percentage    | 1.00            | 1.00              |

|                 | 2000 Population | 1990 Population \ |
|-----------------|-----------------|-------------------|
| Rank            | -0.34           | -0.33             |
| 2022 Population | 0.99            | 0.99              |
| 2020 Population | 1.00            | 0.99              |
| 2015 Population | 1.00            | 0.99              |
| 2010 Population | 1.00            | 1.00              |
| 2000 Population | 1.00            | 1.00              |
| 1990 Population | 1.00            | 1.00              |
| 1980 Population | 1.00            | 1.00              |
| 1970 Population | 0.99            | 1.00              |

|                                |                                |                 |      |
|--------------------------------|--------------------------------|-----------------|------|
| Area (km <sup>2</sup> )        | 0.47                           | 0.52            |      |
| Density (per km <sup>2</sup> ) | -0.03                          | -0.03           |      |
| Growth Rate                    | -0.05                          | -0.07           |      |
| World Population Percentage    | 0.99                           | 0.99            |      |
|                                | 1980 Population                | 1970 Population | Area |
| (km <sup>2</sup> ) \           |                                |                 |      |
| Rank                           | -0.33                          | -0.34           | -    |
| 0.38                           |                                |                 |      |
| 2022 Population                | 0.99                           | 0.97            |      |
| 0.45                           |                                |                 |      |
| 2020 Population                | 0.99                           | 0.98            |      |
| 0.45                           |                                |                 |      |
| 2015 Population                | 0.99                           | 0.98            |      |
| 0.46                           |                                |                 |      |
| 2010 Population                | 0.99                           | 0.98            |      |
| 0.46                           |                                |                 |      |
| 2000 Population                | 1.00                           | 0.99            |      |
| 0.47                           |                                |                 |      |
| 1990 Population                | 1.00                           | 1.00            |      |
| 0.52                           |                                |                 |      |
| 1980 Population                | 1.00                           | 1.00            |      |
| 0.53                           |                                |                 |      |
| 1970 Population                | 1.00                           | 1.00            |      |
| 0.51                           |                                |                 |      |
| Area (km <sup>2</sup> )        | 0.53                           | 0.51            |      |
| 1.00                           |                                |                 |      |
| Density (per km <sup>2</sup> ) | -0.03                          | -0.03           | -    |
| 0.06                           |                                |                 |      |
| Growth Rate                    | -0.08                          | -0.08           | -    |
| 0.01                           |                                |                 |      |
| World Population Percentage    | 0.99                           | 0.97            |      |
| 0.45                           |                                |                 |      |
|                                | Density (per km <sup>2</sup> ) | Growth Rate \   |      |
| Rank                           | 0.13                           | -0.22           |      |
| 2022 Population                | -0.03                          | -0.02           |      |
| 2020 Population                | -0.03                          | -0.03           |      |
| 2015 Population                | -0.03                          | -0.03           |      |
| 2010 Population                | -0.03                          | -0.04           |      |
| 2000 Population                | -0.03                          | -0.05           |      |
| 1990 Population                | -0.03                          | -0.07           |      |
| 1980 Population                | -0.03                          | -0.08           |      |
| 1970 Population                | -0.03                          | -0.08           |      |
| Area (km <sup>2</sup> )        | -0.06                          | -0.01           |      |
| Density (per km <sup>2</sup> ) | 1.00                           | -0.07           |      |
| Growth Rate                    | -0.07                          | 1.00            |      |
| World Population Percentage    | -0.03                          | -0.02           |      |
|                                | World Population Percentage    |                 |      |

```

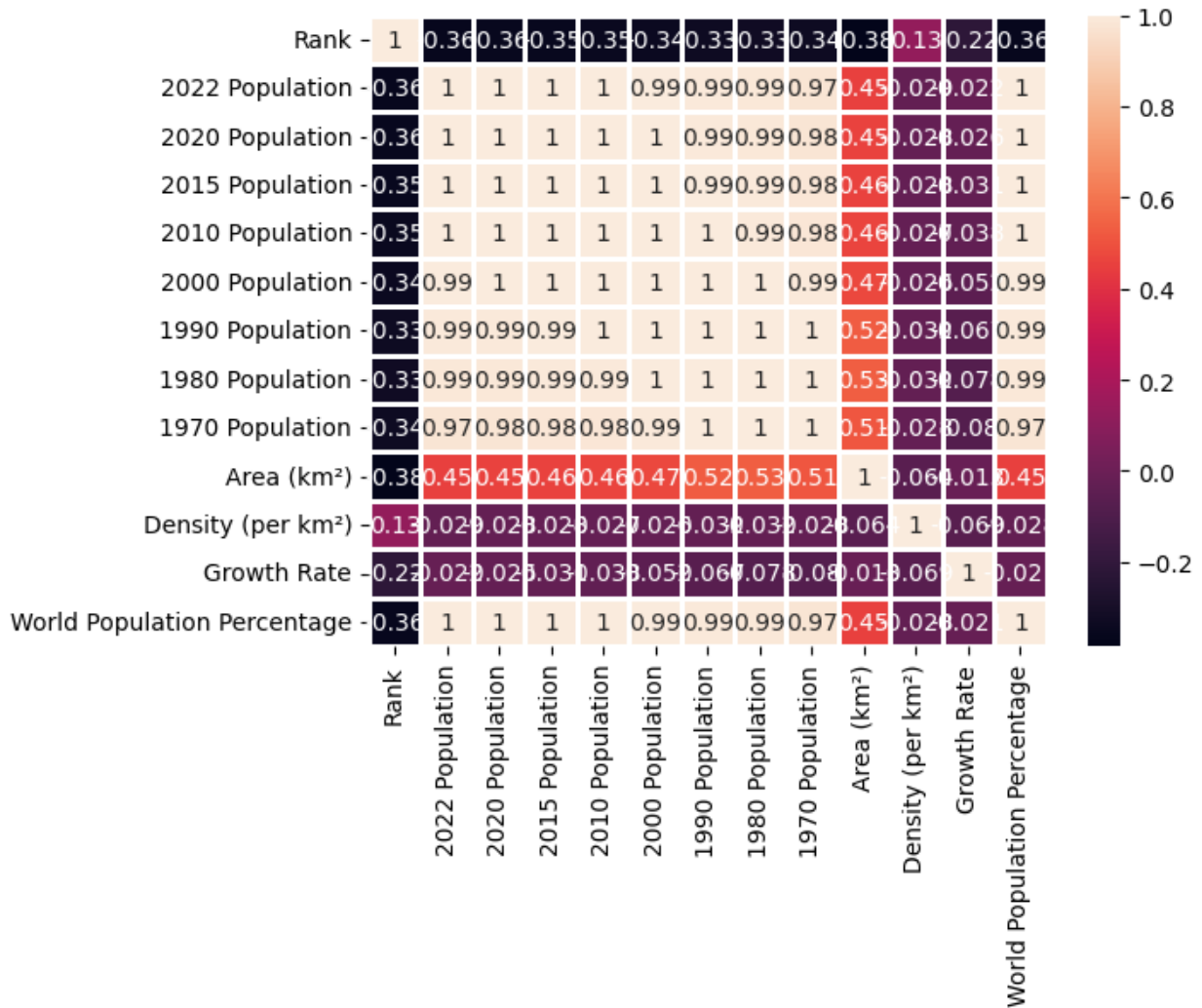
Rank -0.36
2022 Population 1.00
2020 Population 1.00
2015 Population 1.00
2010 Population 1.00
2000 Population 0.99
1990 Population 0.99
1980 Population 0.99
1970 Population 0.97
Area (km²) 0.45
Density (per km²) -0.03
Growth Rate -0.02
World Population Percentage 1.00

```

```

sns.heatmap(numeric.corr(),annot=True,linewidth=1)
plt.figure(figsize = (30,30))
plt.show()

```



<Figure size 3000x3000 with 0 Axes>

*# lets check out which continent has largest population and accordingly we will look top 5 continenet with largest population*  
numeric.columns

```
Index(['Rank', '2022 Population', '2020 Population', '2015 Population',  
      '2010 Population', '2000 Population', '1990 Population',  
      '1980 Population', '1970 Population', 'Area (km²)', 'Density (per km²)',  
      'Growth Rate', 'World Population Percentage'],  
      dtype='object')
```

*# for this i will group by continent and their population in those consecutive years*

```
continent=df.groupby("Continent")[['1970 Population', '1980 Population', '1990 Population', '2000 Population', '2010 Population', '2015 Population', '2020 Population', '2022 Population']].mean()  
continent
```

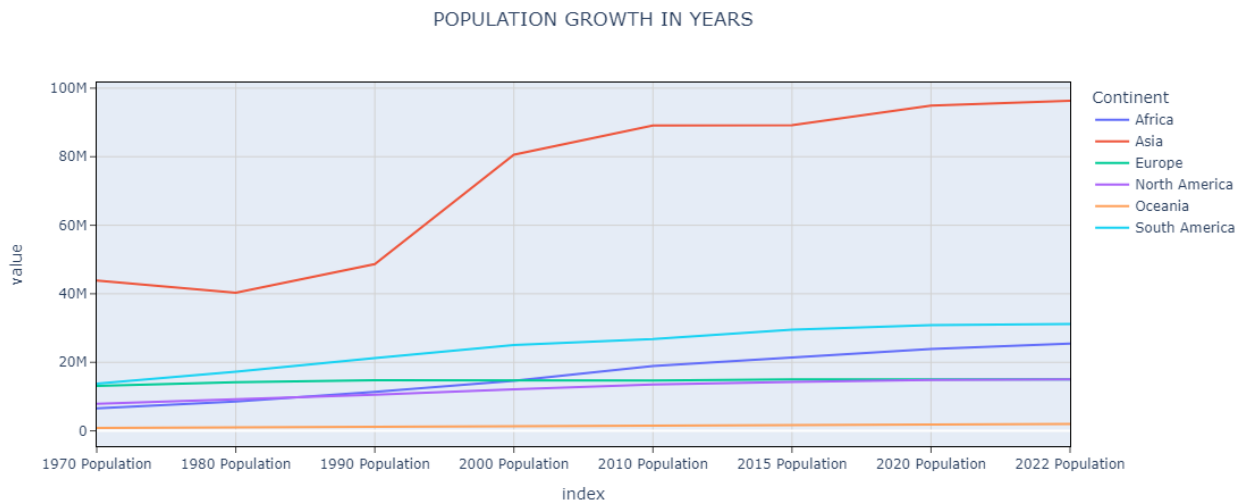
|               | 1970 Population | 1980 Population | 1990 Population | \ |
|---------------|-----------------|-----------------|-----------------|---|
| Continent     |                 |                 |                 |   |
| Africa        | 6567175.27      | 8586031.98      | 11376964.52     |   |
| Asia          | 43839877.83     | 40278333.33     | 48639995.33     |   |
| Europe        | 13118479.82     | 14200004.52     | 14785203.94     |   |
| North America | 7885865.15      | 9207334.03      | 10531660.62     |   |
| Oceania       | 846968.26       | 996532.17       | 1162774.87      |   |
| South America | 13781939.71     | 17270643.29     | 21224743.93     |   |

|               | 2000 Population | 2010 Population | 2015 Population | \ |
|---------------|-----------------|-----------------|-----------------|---|
| Continent     |                 |                 |                 |   |
| Africa        | 14598365.95     | 18898197.31     | 21419703.57     |   |
| Asia          | 80580835.11     | 89087770.00     | 89165003.64     |   |
| Europe        | 14817685.71     | 14712278.68     | 15027454.12     |   |
| North America | 12151739.60     | 13568016.28     | 14259596.25     |   |
| Oceania       | 1357512.09      | 1613163.65      | 1756664.48      |   |
| South America | 25015888.69     | 26789395.54     | 29509599.71     |   |

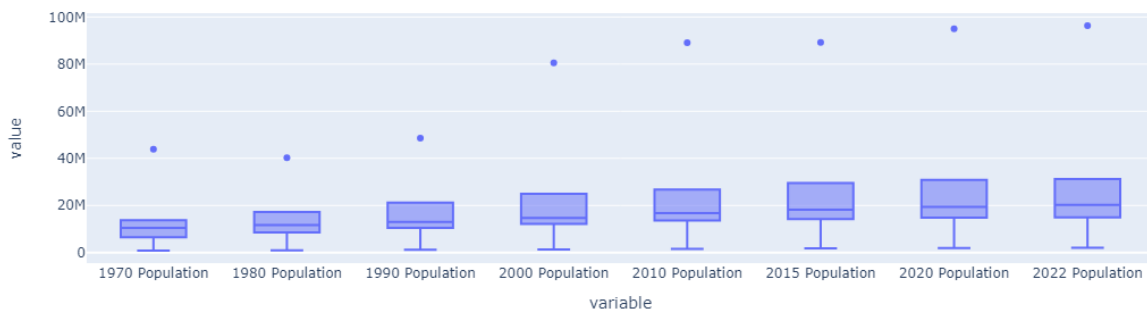
|               | 2020 Population | 2022 Population |
|---------------|-----------------|-----------------|
| Continent     |                 |                 |
| Africa        | 23871435.26     | 25455879.68     |
| Asia          | 94955134.37     | 96327387.31     |
| Europe        | 14915843.92     | 15055371.82     |
| North America | 14855914.82     | 15007403.40     |
| Oceania       | 1910148.96      | 2046386.32      |
| South America | 30823574.50     | 31201186.29     |

lets plot a graph to visualize the growth of population over years

```
fig = ppx.line(continent, title='POPULATION GROWTH IN YEARS')
fig.update_layout(title_x=0.47,height=500)
fig.update_xaxes(
    mirror=True,
    ticks='outside',
    showline=True,
    linecolor='black',
    gridcolor='lightgrey'
)
fig.update_yaxes(
    mirror=True,
    ticks='outside',
    showline=True,
    linecolor='black',
    gridcolor='lightgrey'
)
fig.show()
```



```
# lets take a boxplot to detect the outliers in the dataset
fig = ppx.box(continent)
fig.show()
# from this we can detect outliers in the columns and as we can see
the population has exceeded above the maximum range years
```



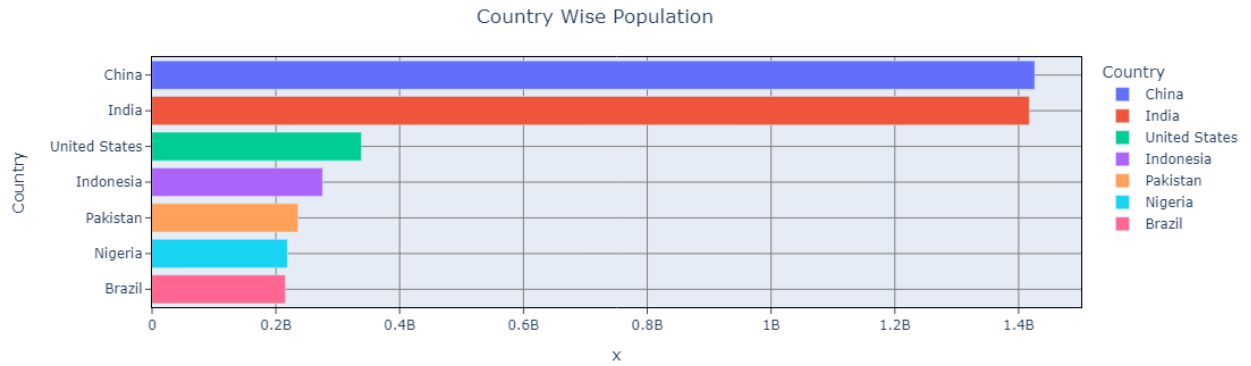
*# lets take a look at countries having highest population we will see the top 7 countries via population of 2022*

df

```
count=df.groupby("Country")["2022
Population"].mean().sort_values(ascending=False).head(7)
count
```

```
Country
China      1425887337.00
India      1417173173.00
United States  338289857.00
Indonesia  275501339.00
Pakistan   235824862.00
Nigeria    218541212.00
Brazil     215313498.00
Name: 2022 Population, dtype: float64
```

```
fig=ppx.bar(count,x=count.values,y=count.index,color=count.index)
fig.update_layout(title="Country Wise Population",title_x=0.47)
fig.update_xaxes(
    mirror=True,
    ticks='outside',
    showline=True,
    linecolor='black',
    gridcolor='grey'
)
fig.update_yaxes(
    mirror=True,
    ticks='outside',
    showline=True,
    linecolor='black',
    gridcolor='grey'
)
fig.show()
```



# from the observation we can see that china stands at the top while india stands at the second position  
# with this i would like to end this Eda on world population