Introduction:

The world population dataset is essential for government organizations in determining policies for its population in the region of a country, continent, or world. Each region needs to know the information extracted from this dataset which can be in the form of population numbers, growth rate, and population density from year to year. Extracting and analyzing this dataset will help plan how each sector will be managed, in terms of health, economy, education, and others.

Data Description:

In this Dataset, we have Historical Population data for every Country/Territory in the world by different parameters like Area Size of the Country/Territory, Name of the Continent, Name of the Capital, Density, Population Growth Rate, Ranking based on Population, World Population Percentage, etc.

- Rank: Ranked particular country by population
- CCA3: 3-digit country code.
- Country/Territories: Country name
- Capital: Capital of Country
- Continent: Name of the Continent where the Country Belongs
- 2022 population: The Population of the Country in the year 2022.
- 2020 population: The Population of the Country in the year 2020.
- 2015 population: The Population of the Country in the year 2015.
- 2010 population: The Population of the Country in the year 2010.
- 2000 population: The Population of the Country in the year 2000.
- 1990 population: The Population of the Country in the year 1990.
- 1980 population: The Population of the Country in the year 1980.
- 1970 population: The Population of the Country in the year 1970.
- Area: The land area of the Country (measured in km^2).
- Growth Rate: The Population Growth Rate of the Country.
- World Population Percentage: The percentage of the World Population residing in that Country.

We will import the libraries :

```
import numpy as np
import pandas as pd
import seaborn as sns
from plotly.subplots import make_subplots
import matplotlib.pyplot as plt
import plotly.express as px
import plotly.graph_objects as go
import missingno as msno
import plotly.offline as py
from plotly.offline import iplot
import plotly.io as pio
from plotly.offline import init_notebook_mode
py.init_notebook_mode()
import os
from wordcloud import WordCloud
from sklearn.preprocessing import LabelEncoder
```

Reading the dataset:

```
df= pd.read_csv("../input/world-population-dataset/world_population.csv")
df.head()
```

Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	2010 Population	2000 Population	1990 Population	1980 Population	Po
36	AFG	Afghanistan	Kabul	Asia	41128771	38972230	33753499	28189672	19542982	10694796	12486631	
138	ALB	Albania	Tirana	Europe	2842321	2866849	2882481	2913399	3182021	3295066	2941651	
34	DZA	Algeria	Algiers	Africa	44903225	43451666	39543154	35856344	30774621	25518074	18739378	
213	ASM	American Samoa	Pago Pago	Oceania	44273	46189	51368	54849	58230	47818	32886	
203	AND	Andorra	Andorra la Vella	Europe	79824	77700	71746	71519	66097	53569	35611	

Data exploration:

```
print(df.shape)
```

```
(234, 17)
```

The dataset contains 234 rows and 17 columns.

```
df.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
```

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

```
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
dtype: int64
```

There is no missing value.

```
df.duplicated().sum()
```

```
0
```

There is no duplicate value.

The data set is clean.

Summary statistic:

```
df.describe().T.sort_values(ascending = 0,by = "mean").style.background_gradient(cmap = "viridis")
```

	count	mean	std	min	25%	50%	75%	max
2022 Population	234.000000	34074414.709402	136766424.804763	510.000000	419738.500000	5559944.500000	22476504.750000	1425887337.000000
2020 Population	234.000000	33501070.952991	135589876.924439	520.000000	415284.500000	5493074.500000	21447979.500000	1424929781.000000
2015 Population	234.000000	31729956.243590	130404992.751760	564.000000	404676.000000	5307400.000000	19730853.750000	1393715448.000000
2010 Population	234.000000	29845235.034188	124218487.632998	596.000000	393149.000000	4942770.500000	19159567.500000	1348191368.000000
2000 Population	234.000000	26269468.816239	111698206.719070	651.000000	327242.000000	4292907.000000	15762301.000000	1264099069.000000
1990 Population	234.000000	22710220.790598	97832173.346751	700.000000	264115.750000	3825409.500000	11869231.000000	1153704252.000000
1980 Population	234.000000	18984616.970085	81785186.084201	733.000000	229614.250000	3141145.500000	9826053.750000	982372466.000000
1970 Population	234.000000	15786908.807692	67795091.643236	752.000000	155997.000000	2604830.000000	8817329.000000	822534450.000000
Area (km²)	234.000000	581449.384615	1761840.864063	1.000000	2650.000000	81199.500000	430425.750000	17098242.000000
Density (per km²)	234.000000	452.127044	2066.121904	0.026100	38.417875	95.346750	238.933250	23172.266700
Rank	234.000000	117.500000	67.694165	1.000000	59.250000	117.500000	175.750000	234.000000
Growth Rate	234.000000	1.009577	0.013385	0.912000	1.001775	1.007900	1.016950	1.069100
World Population Percentage	234.000000	0.427051	1.714977	0.000000	0.010000	0.070000	0.280000	17.880000

Data Preparation :

- The data set is clean, but we will rename the column to a shorter name.

```
# Renaming year columns from "Country/Territory" to "Country" for shorter name

df.rename(columns = {'Country/Territory':'Country'}, inplace = True)

df.head()
```

	Rank	Country	Continent	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
0	36	Afghanistan	Asia	41128771	38972230	33753499	28189672	19542982	10694796	12486631	10752971	652230	63.0587	1.0257	0.52
1	138	Albania	Europe	2842321	2866849	2882481	2913399	3182021	3295066	2941651	2324731	28748	98.8702	0.9957	0.04
2	34	Algeria	Africa	44903225	43451666	39543154	35856344	30774621	25518074	18739378	13795915	2381741	18.8531	1.0164	0.56
3	213	American Samoa	Oceania	44273	46189	51368	54849	58230	47818	32886	27075	199	222.4774	0.9831	0.00
4	203	Andorra	Europe	79824	77700	71746	71519	66097	53569	35611	19860	468	170.5641	1.0100	0.00

```
# Renaming year columns from "Year Population" to just "Year" for shorter name

for col in df.columns:
    if 'Population' and '0' in col:
        df = df.rename(columns={col: col.split(' ')[0]})

df.head()
```

	Rank	Country	Continent	2022	2020	2015	2010	2000	1990	1980	1970	Area (km^2)	Density (per km^2)	Growth Rate	World Population Percentage
0	36	Afghanistan	Asia	41128771	38972230	33753499	28189672	19542982	10694796	12486631	10752971	652230	63.0587	1.0257	0.52
1	138	Albania	Europe	2842321	2866849	2882481	2913399	3182021	3295066	2941651	2324731	28748	98.8702	0.9957	0.04
2	34	Algeria	Africa	44903225	43451666	39543154	35856344	30774621	25518074	18739378	13795915	2381741	18.8531	1.0164	0.56
3	213	American Samoa	Oceania	44273	46189	51368	54849	58230	47818	32886	27075	199	222.4774	0.9831	0.00
4	203	Andorra	Europe	79824	77700	71746	71519	66097	53569	35611	19860	468	170.5641	1.0100	0.00

Correlation:

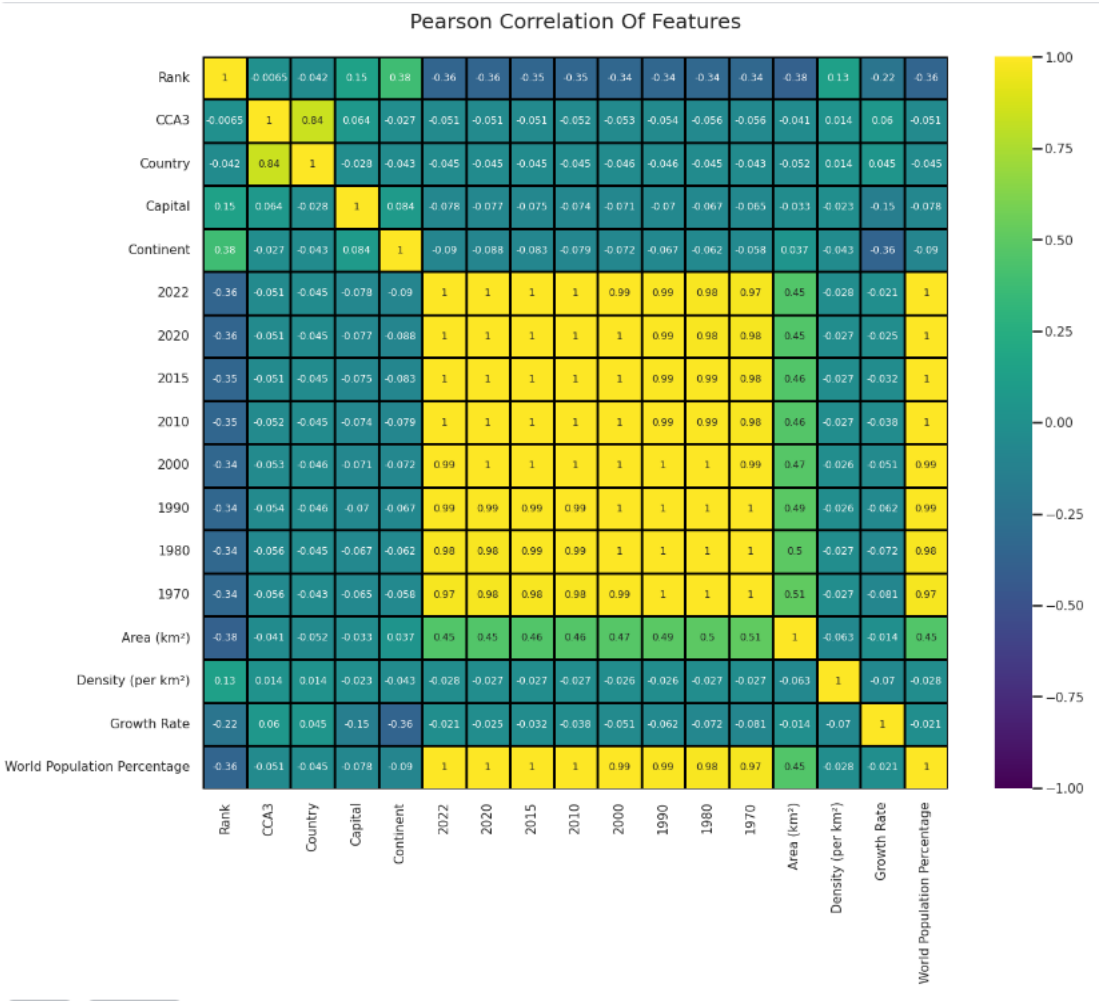
```
sns.set_style("white")
sns.set(rc={"axes.facecolor":"white","figure.facecolor":"white"})
sns.set_context("poster",font_scale = .7)

palette = ["#006837","#1A9850","#66BD63","#A6D96A","#D9EF8B","#FFFFBF","#FEE08B","#FDAE61","#F46D43","#D73027"
palette_cmap =['viridis']
```

```
catcol = ["CCA3","Country","Capital","Continent"]
le = LabelEncoder()
for col in catcol:
        df[col] = le.fit_transform(df[col])
```

```
plt.subplots(figsize =(20, 20))

sns.heatmap(df.corr(), cmap = 'viridis', square=True, cbar_kws=dict(shrink =.82),
            annot=True, vmin=-1, vmax=1, linewidths=3,linecolor='black',annot_kws=dict(fontsize =12))
plt.title("Pearson Correlation Of Features\n", fontsize=25)
plt.xticks(rotation=90)
plt.show()
```



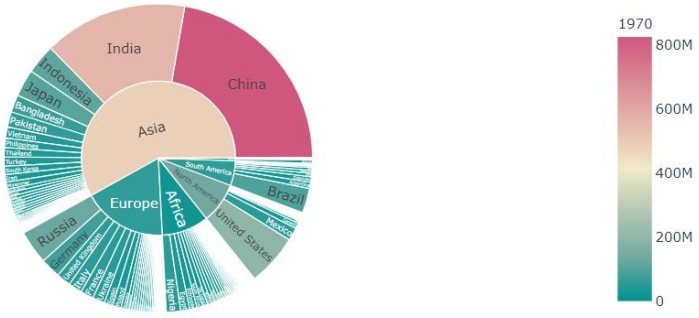
- A high correlation between population and world population percentage. Also, a high correlation between country and CCA3 code as expected.
- Medium correlation between area and population, area and world population percentage.
- Medium inverse correlation between growth rate and continent, rank and world population percentage, rank and area, rank and population.

```
years=['1970','1980','1990','2000','2010','2015','2020','2022']
for x in years:
    fig = px.sunburst(df, path=['Continent', 'Country'], values=x,color_continuous_scale='Tealrose',
                template='plotly_white', color=x,title=x)
    fig.show()
    fig = px.choropleth(df,
                    locations='Country',
                    locationmode='country names',
                    color=x,
                    color_continuous_scale='Tealrose',
                    template='plotly_white',
                    title = x)

    fig.show()

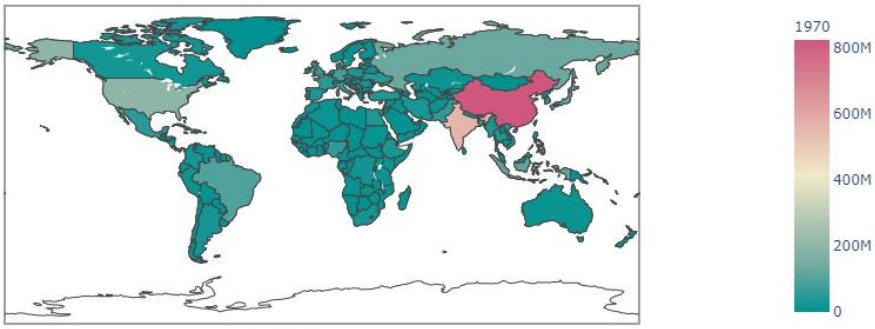
    fig = px.treemap(df, path=[px.Constant("world"), 'Continent', 'Country'], values=x,
                 color=x,template='plotly_white',title = x,
                 color_continuous_scale='Tealrose')
    fig.show()
```

1970

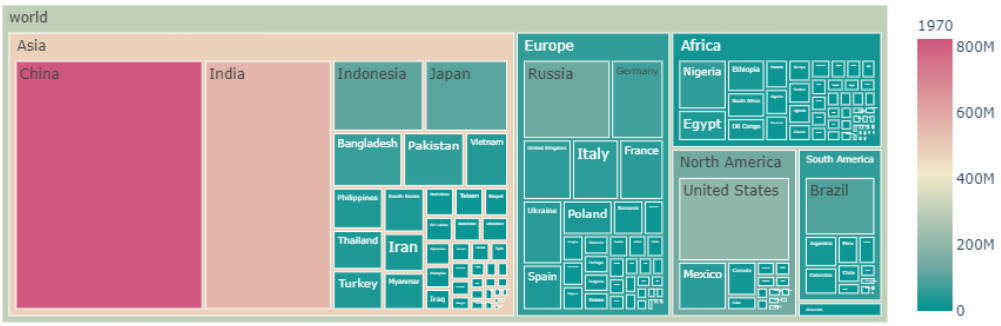


-

1970

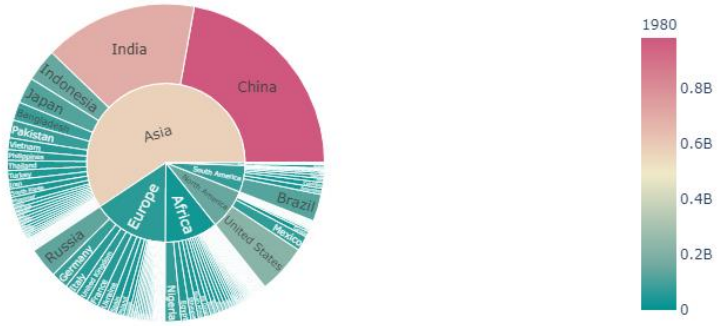


1970

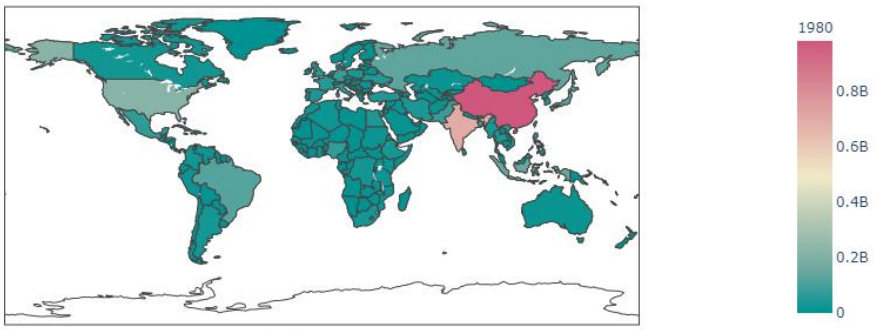


In 1970 China was the country with the highest population in the world with a population of 822 M and India was the second country with a population of 557M.

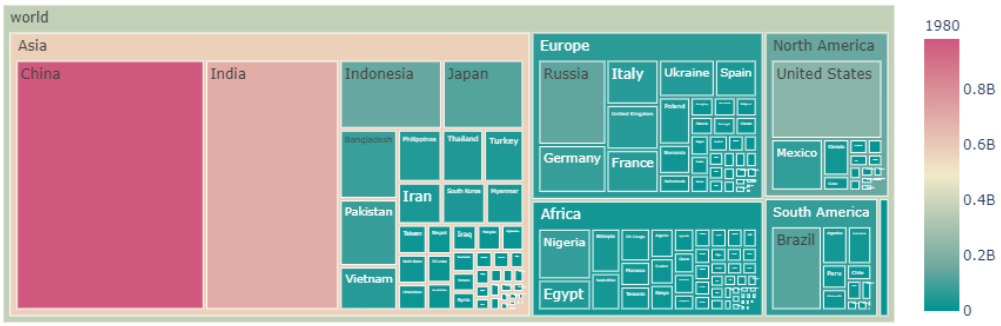
1980



1980

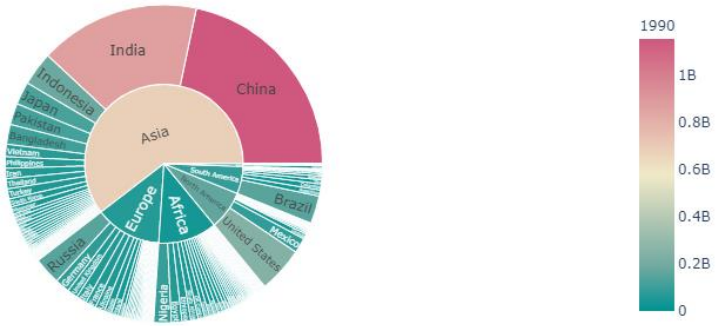


1980

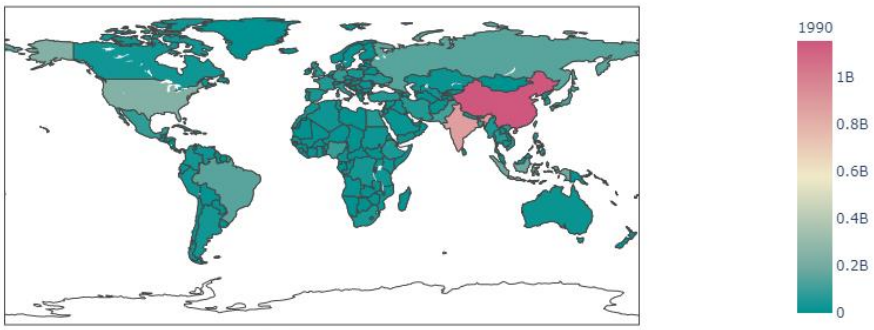


- In 1980 China was the country with the highest population in the world with a population of 982M and India was the second country with a population of 696M.

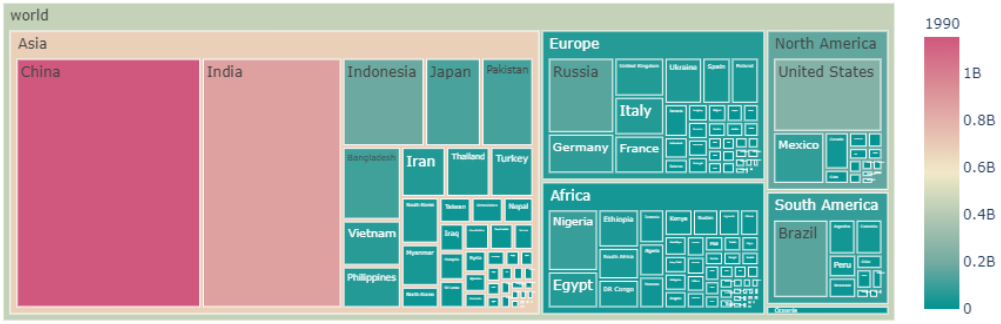
1990



1990



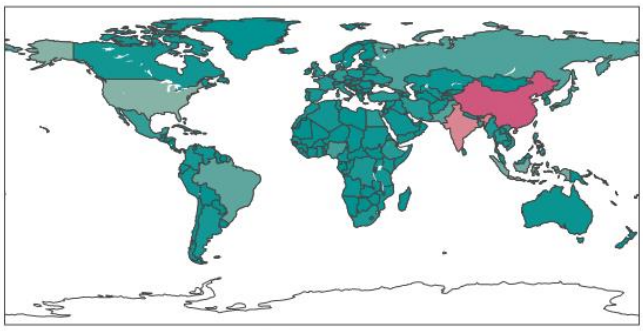
1990



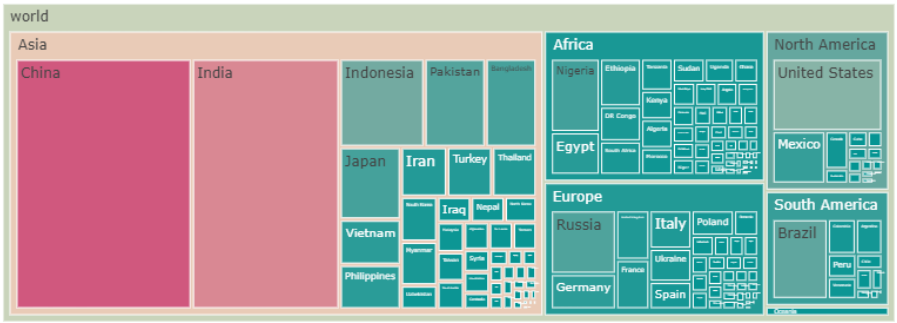
- In 1990 China was the country with the highest population in the world with a population of 1.15B and India was the second country with a population of 870M.

2000

2000

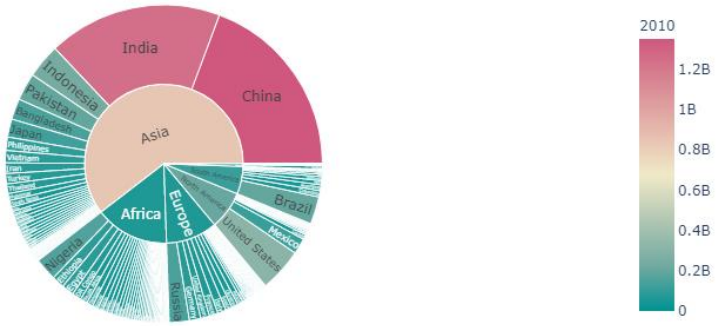


2000

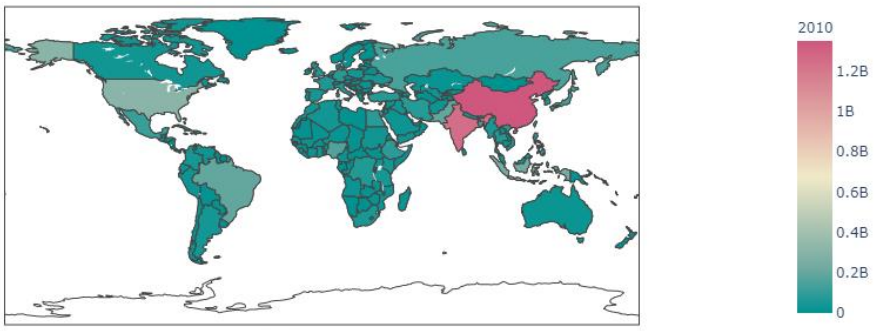


- In 2000 China was the country with the highest population in the world with a population of 1.26B and India was the second country with a population of 1B.

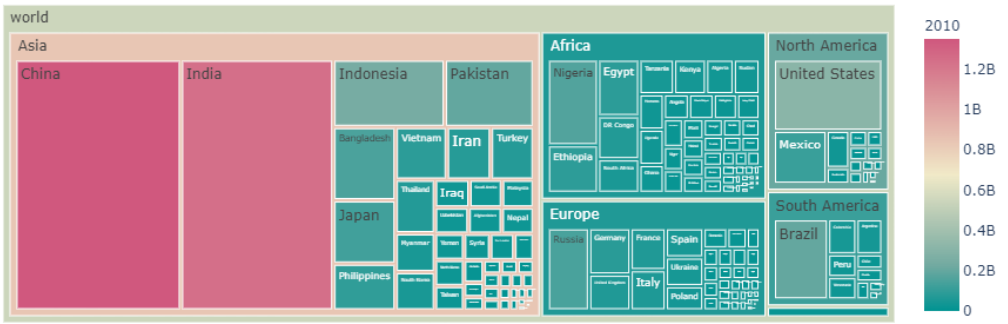
2010



2010

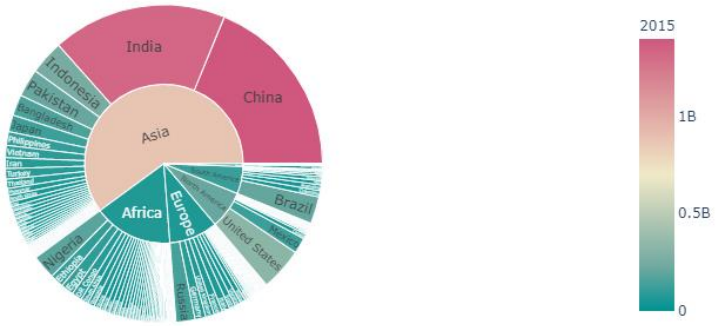


2010

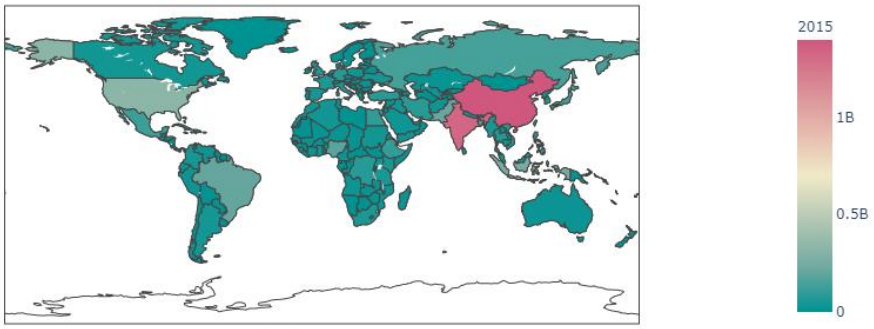


- In 2010 China was the country with the highest population in the world with a population of 1.34B and India was the second country with a population of 1.24B.

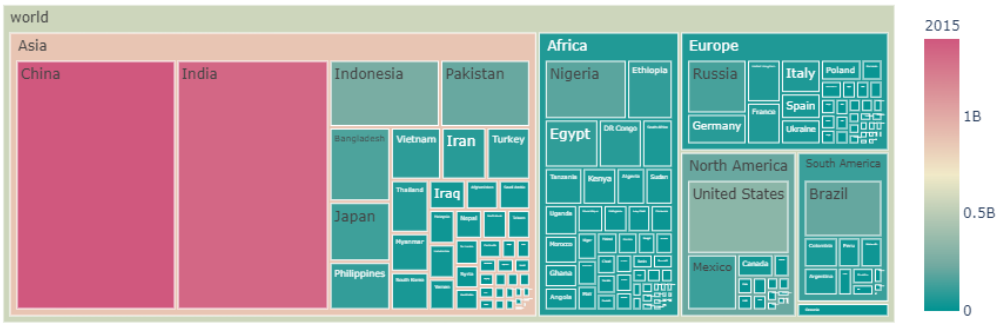
2015



2015

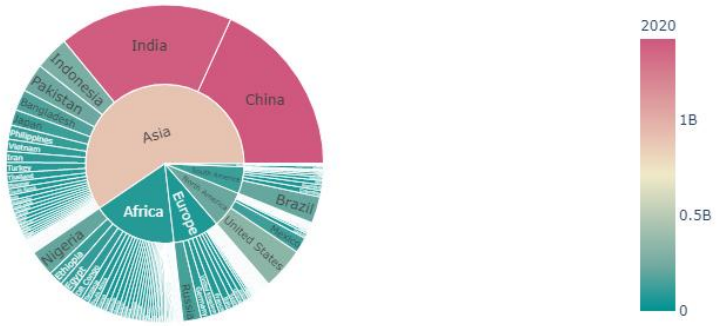


2015



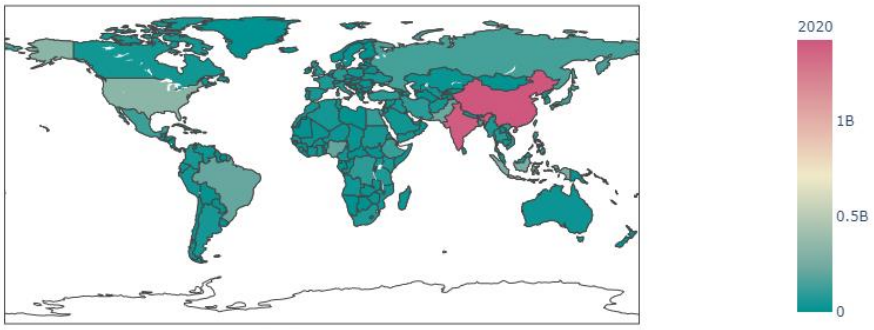
- In 2015 China was the country with the highest population in the world with a population of 1.39B and India was the second country with a population of 1.32B.

2020

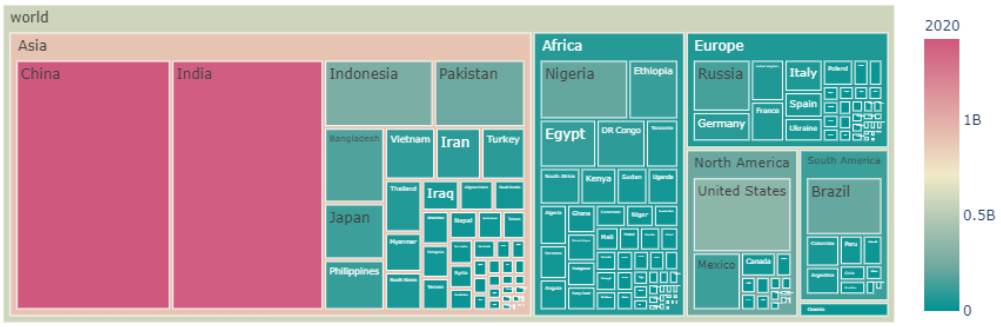


-

2020

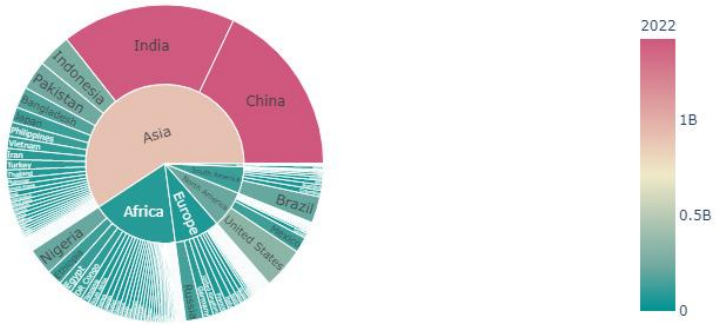


2020



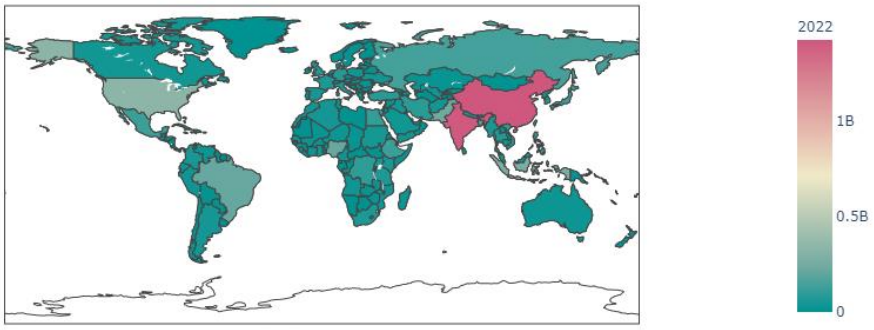
In 2020 China was the country with the highest population in the world with a population of 1.42B and India was the second country with a population of 1.39B.

2022

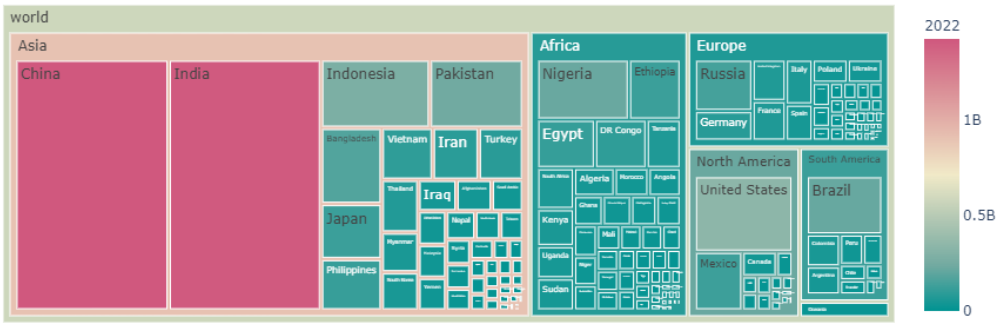


-

2022



2022



In 2022 China was the country with the highest population in the world with a population of 1.42B and India is the second country with a population of 1.41B.

- ```
  by_continent = df.groupby('Continent').sum().sort_values(by = '2022',ascending = True)
  ```

```
x = by_continent.index
fig = go.Figure()
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,1].values, name="2022",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,2].values, name="2020",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,3].values, name="2015",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,4].values, name="2010",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,5].values, name="2000",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,6].values, name="1990",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,7].values, name="1980",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))
fig.add_trace(go.Scatter(x=x, y=by_continent.iloc[:,8].values, name="1970",marker = dict(size = 8,line_width=1.5),
                    mode='lines'))

fig.update_traces(mode='lines+markers')
fig.update_layout(
    title = {
        'text': "Continent Poulation Growth (1970 to 2022)",
        'x':0.5,
        'xanchor': 'center'},

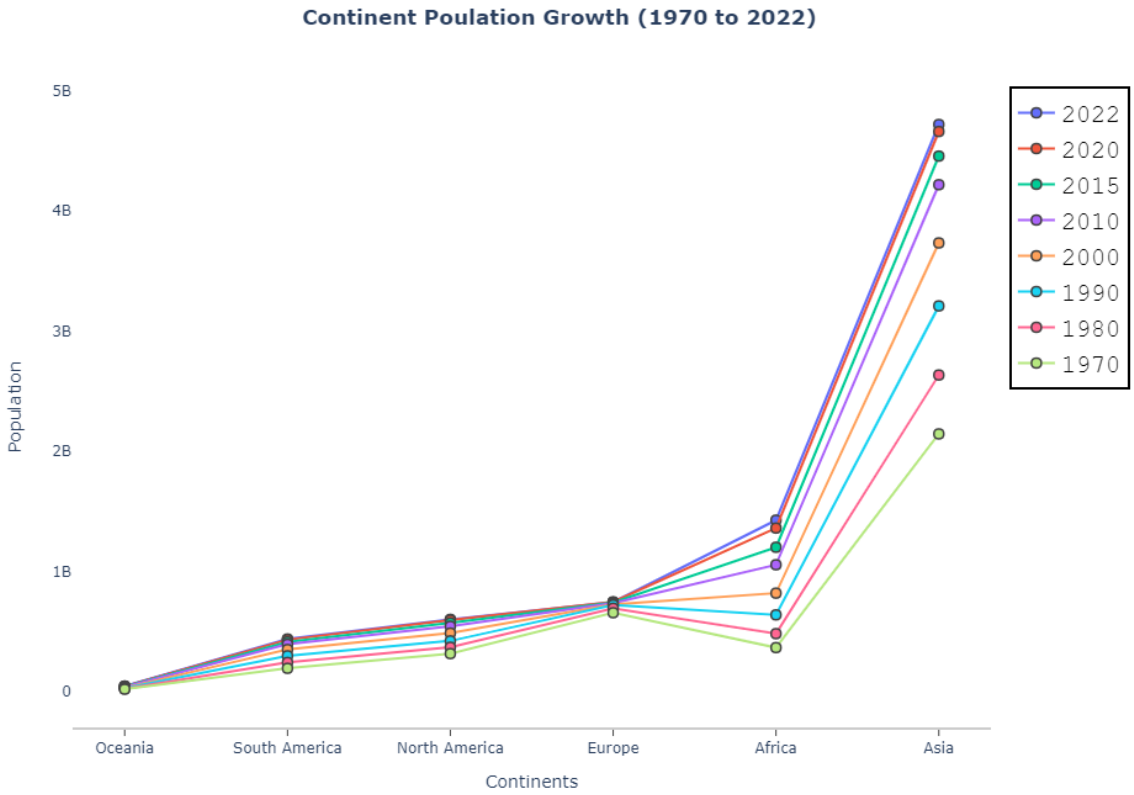
     xaxis=dict(
        title = 'Continents',
        showline=True,
```

```
      xaxis=dict(
         title = 'Continents',
         showline=True,
         showgrid = False,
         linecolor='rgb(204, 204, 204)',
         linewidth=2,
         ticks='outside'),

     yaxis = dict(
         title = 'Population',
         showgrid = False),

         legend=dict(
         title_font_family="Times New Roman",
         font=dict(
             family="Courier",
             size=20,
             color="black"),
         bgcolor="white",
         bordercolor="Black",
         borderwidth=2),

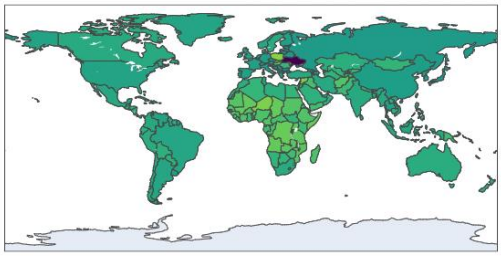
     width = 950,
     height = 700,
     plot_bgcolor='white'
)
iplot(fig)
```
```

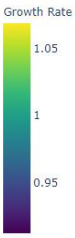


- The Population growth of Asia (1970 to 2022) is from 2.144B to 4.721B with a difference of 2.577B.
- The Population growth of Africa (1970 to 2022) is from 365.44M to 1.426B with a difference of 1.06B.
- The Population growth of Europe (1970 to 2022) is from 655.94M to 745.79M with a difference of 89.85M.
- The Population growth of North America (1970 to 2022) is from 315.43M to 600.269M with a difference of 284.839M.
- The Population growth of South America (1970 to 2022) is from 192.94M to 436.81M with a difference of 243.87M
- The Population growth of Oceania (1970 to 2022) is from 19.480M to 45.038M with a difference of 25.558M.

```
#Growth Rate
fig = px.choropleth(df,
                    locations='Country',
                    locationmode='country names',
                    color='Growth Rate',
                    hover_name='Country',
                    title = 'Growth Rate',
                    color_continuous_scale='Viridis'
                    )
fig.show()
```

Growth Rate



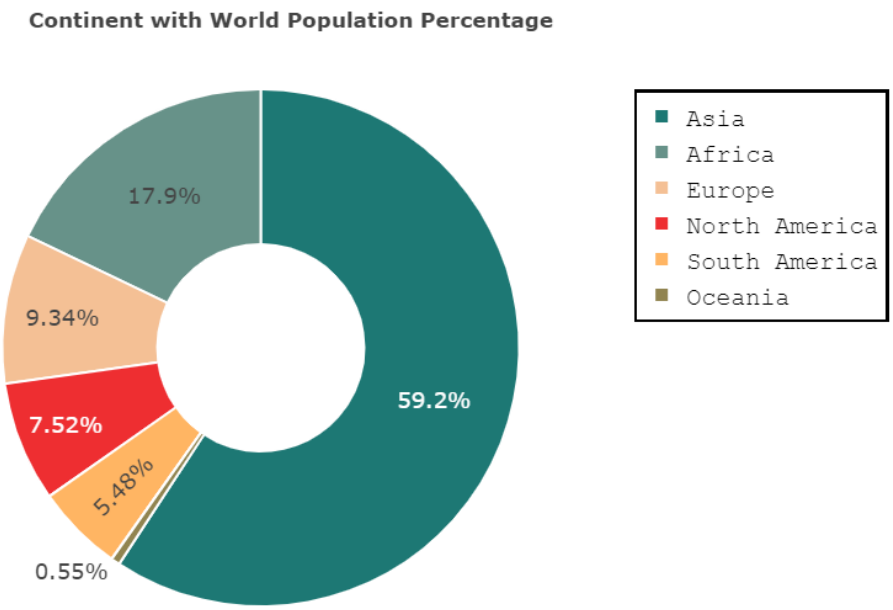


- The Growth Rate Of Ukraine is lowest at 0.912 and that of Moldova is highest at 1.0691.

```
cont_pop = df.groupby('Continent',)[['World Population Percentage']].sum().sort_values(by = 'World Population Percentage', ascending = False)
cont_pop
```

Continent	World Population Percentage
Asia	59.19
Africa	17.87
Europe	9.33
North America	7.51
South America	5.48
Oceania	0.55

```
fig = go.Figure(data = go.Pie(labels = cont_pop.index, values = cont_pop['World Population Percentage'].values))
fig.update_traces(hoverinfo='label',
                  hole = 0.4,
                  textfont_size = 18,
                  textposition ='auto',
                  marker=dict(colors = colors,
                              line = dict(color = 'white',
                                          width = 2)))
fig.update_layout(title ={'text' : '<b>Continent with World Population Percentage</b>',
                          'x' : 0.21},
                          template = 'xgridoff',
                          width = 900, height = 600,
                     legend=dict(
                        title_font_family="Times New Roman",
                        font=dict(
                        family="Courier",
                        size=20,
                        color="black"
                        ),
                        bgcolor="white",
                        bordercolor="Black",
                        borderwidth=2.5)
                  )
iplot(fig)
```



Asia is the continent with the highest population in the world with a 59.2% percentage

And Oceania is the continent with the lowest population in the world with a 0.55% percentage

Data visualization:

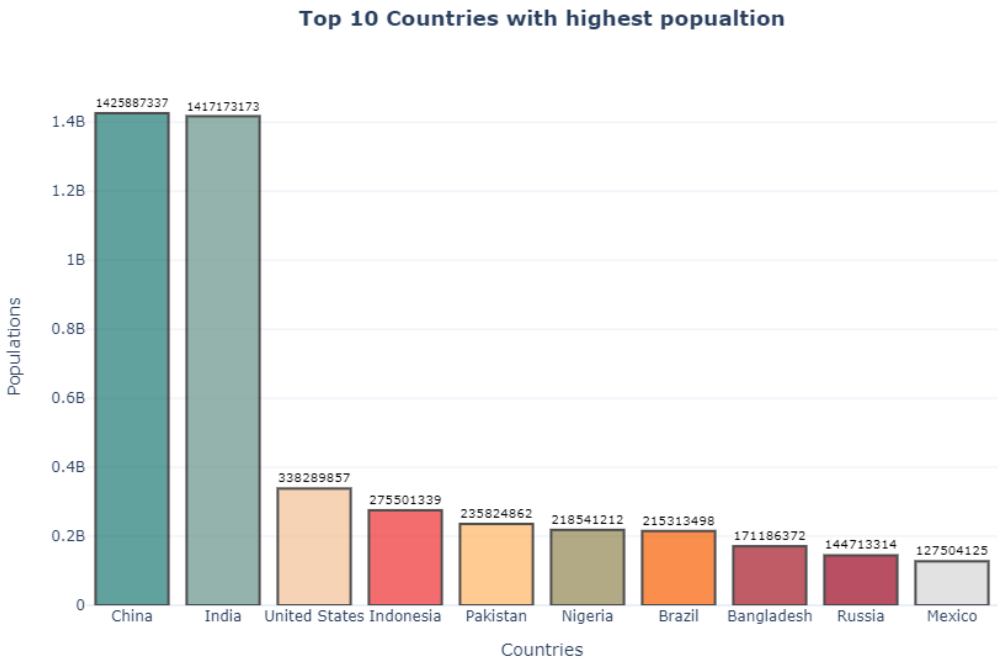
What are the top 10 Countries with the highest population?

```
top_pop = df.sort_values(by = '2022', ascending = False).head(10)
top_pop[['Country', '2022']]
```

	Country	2022
41	China	1425887337
92	India	1417173173
221	United States	338289857
93	Indonesia	275501339
156	Pakistan	235824862
149	Nigeria	218541212
27	Brazil	215313498
16	Bangladesh	171186372
171	Russia	144713314
131	Mexico	127504125

```
data = go.Bar(x = top_pop['Country'], y = top_pop['2022'], text = top_pop['2022'],textposition ='outside',
              textfont = dict(size = 30,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2))
layout = go.Layout(title = {'text': "<b>Top 10 Countries with highest popualtion</b>",
                            'x':0.5,
                            'xanchor': 'center'},
                   xaxis = dict(title='Countries' ),
                   yaxis =dict(title='Populations'),
                   width = 900,
                   height = 600,
                   template = 'plotly_white')
fig=go.Figure(data = data, layout = layout)

iplot(fig)
```



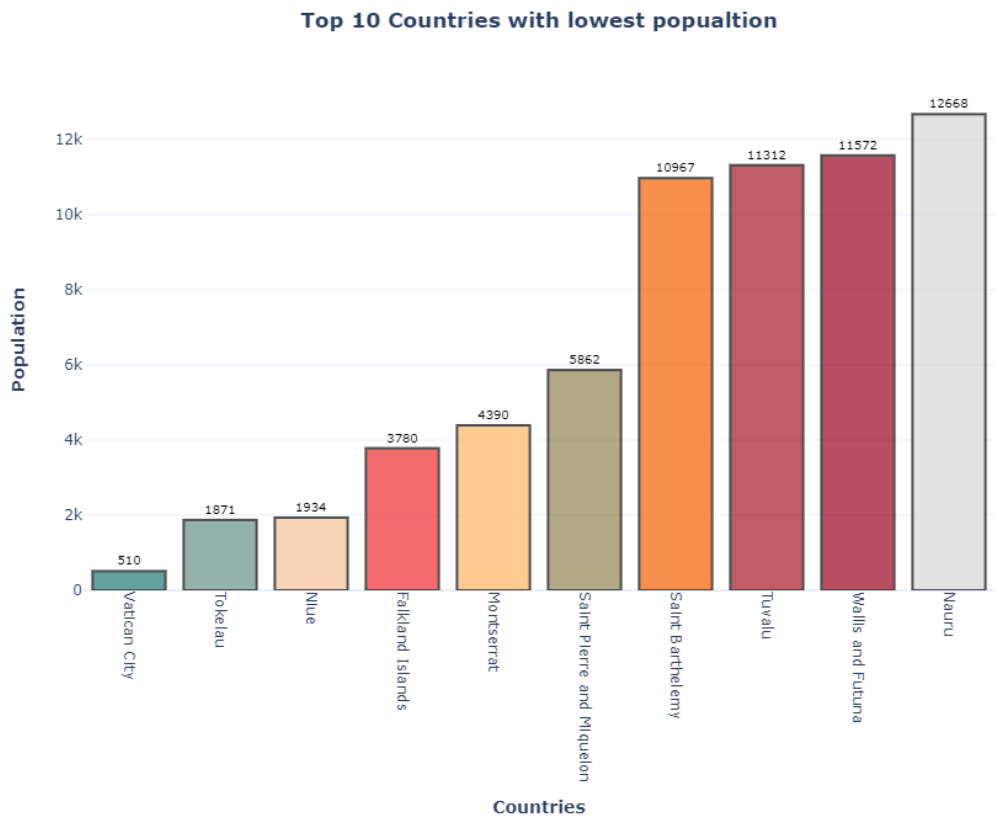
From the above graph, we can see that China is the country with the highest population of 1.42B in the world & India is the second country with a population of 1.41B.

What are the top 10 Countries with the lowest population?

```
less_pop = df.sort_values(by = '2022', ascending = True).head(10)
less_pop[['Country', '2022']]
```

	Country	2022
226	Vatican City	510
209	Tokelau	1871
150	Niue	1934
64	Falkland Islands	3780
137	Montserrat	4390
177	Saint Pierre and Miquelon	5862
173	Saint Barthelemy	10967
216	Tuvalu	11312
229	Wallis and Futuna	11572
142	Nauru	12668

```
data = go.Bar(x = less_pop['Country'], y = less_pop['2022'],text = less_pop['2022'],textposition ='outside',
              textfont = dict(size = 10,
                              color = 'black'),
             marker = dict(color = colors, opacity = 0.7, line_color = 'black', line_width = 2))
layout = go.Layout(title = {'text' : '<b>Top 10 Countries with lowest popualtion</b>', 'x' : 0.5},
                   xaxis = dict(title = '<b>Countries</b>'),
                   yaxis = dict(title = '<b>Population</b>'),
                   width = 900,
                   height = 700,
                   template = 'plotly_white')
fig = go.Figure(data = data, layout = layout)
fig.update_xaxes(tickangle=90,tickfont_size = 12)
iplot(fig)
```



Vatican City is the country with the lowest population of 510

What are the top 5 Countries with the highest population of each Continent?

```
# Making function which gives the top5 countries with highest population of continent
def get_cont_top5_country(continent):
    return df[df['Continent'] == continent].groupby('Country')[['2022']].sum().sort_values(by = '2022', ascending=False).head(5)
```

```
# Storing the contitnet top5 countries data with same continent names
Asia = get_cont_top5_country('Asia')
Europe = get_cont_top5_country('Europe')
Africa = get_cont_top5_country('Africa')
North_America = get_cont_top5_country('North America')
South_America = get_cont_top5_country('South America')
Oceania = get_cont_top5_country('Oceania')
```

```
# Making subplots with 3 Rows & Columns
fig = make_subplots(rows=3, cols=2, subplot_titles=("Asia", 'Europe', "Africa", "Oceania", 'Nort America', 'South America'))

# Asia
fig.add_trace(go.Bar(x = Asia.index, y = Asia['2022'], text = Asia['2022'],textposition ='auto',

              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
             row = 1, col =1)
# Europe
fig.add_trace(go.Bar(x = Europe.index, y = Europe['2022'], text = Europe['2022'],textposition ='auto',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 1, col =2)
# Africa
fig.add_trace(go.Bar(x = Africa.index, y = Africa['2022'], text = Africa['2022'],textposition ='auto',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 2, col =1)
```

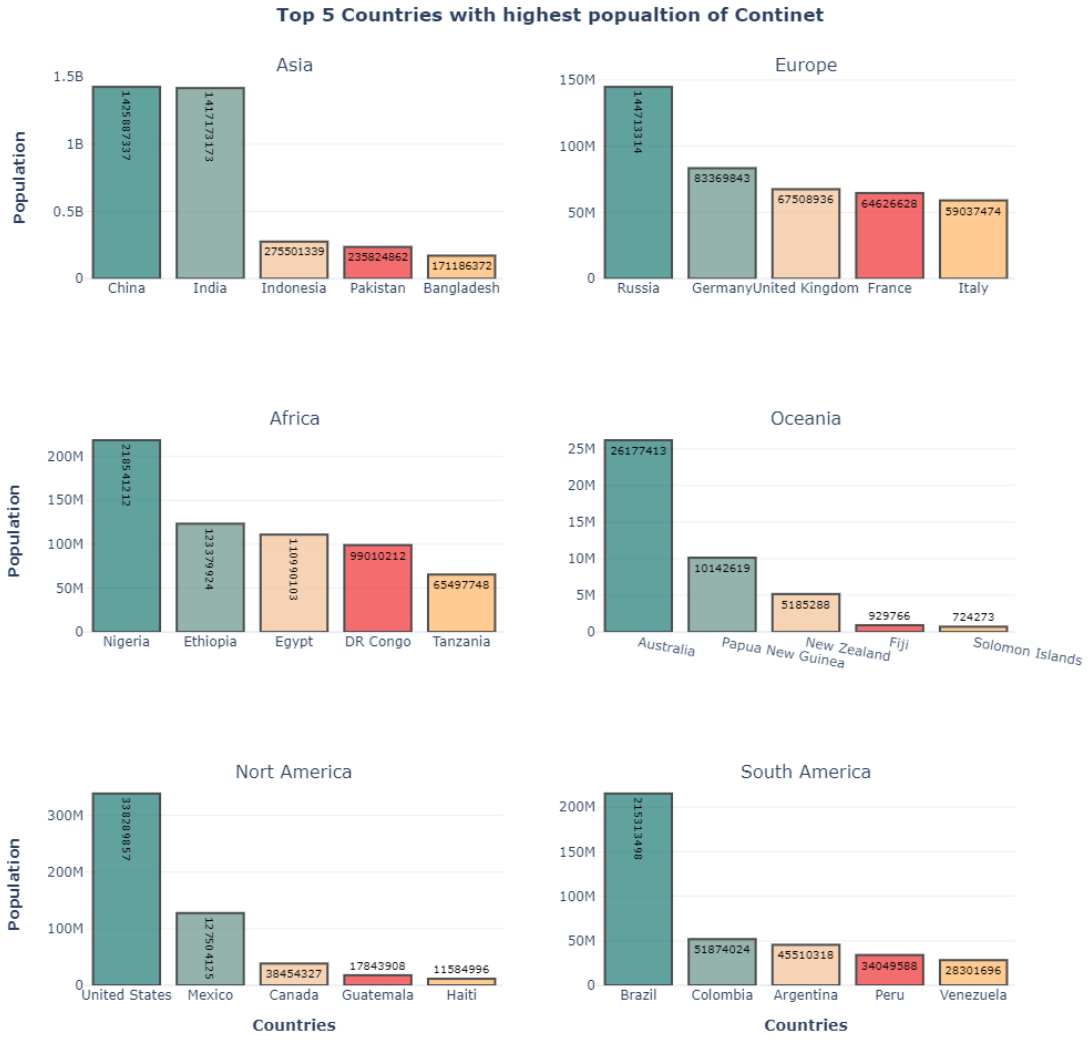
```
 # Oceania
 fig.add_trace(go.Bar(x = Oceania.index, y = Oceania['2022'], text = Oceania['2022'],textposition ='auto',
               textfont = dict(size = 10,
                               color = 'black'),
               marker = dict(color = colors,
                             opacity = 0.7,
                             line_color = 'black',
                             line_width = 2)),
                                   row = 2, col =2)

 # Nort America
 fig.add_trace(go.Bar(x = North_America.index, y = North_America['2022'], text = North_America['2022'],textposition ='auto',
               textfont = dict(size = 10,
                               color = 'black'),
               marker = dict(color = colors,
                             opacity = 0.7,
                             line_color = 'black',
                             line_width = 2)),
                                   row = 3, col =1)
 # South America
 fig.add_trace(go.Bar(x = South_America.index, y = South_America['2022'], text = South_America['2022'],
                      textposition ='auto',
               textfont = dict(size = 10,
                               color = 'black'),
               marker = dict(color = colors,
                             opacity = 0.7,
                             line_color = 'black',
                             line_width = 2)),
                                   row = 3, col =2)
```

```
 # updating Layout
 fig.update_layout(title = {'text': ' <b>Top 5 Countries with highest popualtion of Continet</b> ',
                            'x':0.5,
                            'xanchor': 'center'},
                   width = 1000,
                   height = 1000,
                   showlegend = False,
                   template = 'plotly_white')
 #  xaxis title
 fig.update_xaxes(tickangle = 0)
 fig.update_xaxes(tickangle = 10, row=2, col=2)
 fig.update_xaxes(title_text="<b>Countries</b>", row=3, col=1)
 fig.update_xaxes(title_text="<b>Countries</b>", row=3, col=2)

 # yaxis title
 fig.update_yaxes(title_text="<b>Population</b>", row = 1, col=1)
 fig.update_yaxes(title_text="<b>Population</b>", row = 2, col=1)
 fig.update_yaxes(title_text="<b>Population</b>", row = 3, col=1)

 iplot(fig)
```



- In Asia Continent the Country with Highest Population is China with a population of 1.42 B.
- In Europe Continent the Country with Highest Population is Russia with a population of 144.71 M.
- In Africa Continent the Country with Highest Population is Nigeria with a population of 218.54 M.
- In Oceania Continent the Country with Highest Population is Australia with a population of 26.17 M.
- In North America Continent the Country with Highest Population is the United State with a population of 338.28 M.
- In South America Continent the Country with Highest Population is Brazil with a population of 215.31 M.

What are the top 5 Countries with the lowest population of Each Continent?

```
# Making function which gives the bottom5 countries with lowest population of continent
def get_cont_bott5_country(continent):
    return df[df['Continent'] == continent].groupby('Country')[['2022']].sum().sort_values(by = '2022', ascending=True).head(5)
```

```
# Storing the contitnet bott5 countries data
bott5_Asia = get_cont_bott5_country('Asia')
bott5_Europe = get_cont_bott5_country('Europe')
bott5_Africa = get_cont_bott5_country('Africa')
bott5_North_America = get_cont_bott5_country('North America')
bott5_South_America = get_cont_bott5_country('South America')
bott5_Oceania = get_cont_bott5_country('Oceania')
```

```
# Making subplots with 3 Rows & Columns
fig = make_subplots(rows=3, cols=2, subplot_titles=("Asia", 'Europe', "Africa", "Oceania", 'Nort America', 'South America'))

# Asia
fig.add_trace(go.Bar(x = bott5_Asia.index, y = bott5_Asia['2022'], text = bott5_Asia['2022'],textposition ='auto',

              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
             row = 1, col =1)
# Europe
fig.add_trace(go.Bar(x = bott5_Europe.index, y = bott5_Europe['2022'], text = bott5_Europe['2022'],textposition ='auto',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 1, col =2)

# Africa
fig.add_trace(go.Bar(x = bott5_Africa.index, y = bott5_Africa['2022'], text = bott5_Africa['2022'],textposition ='auto',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 2, col =1)
```

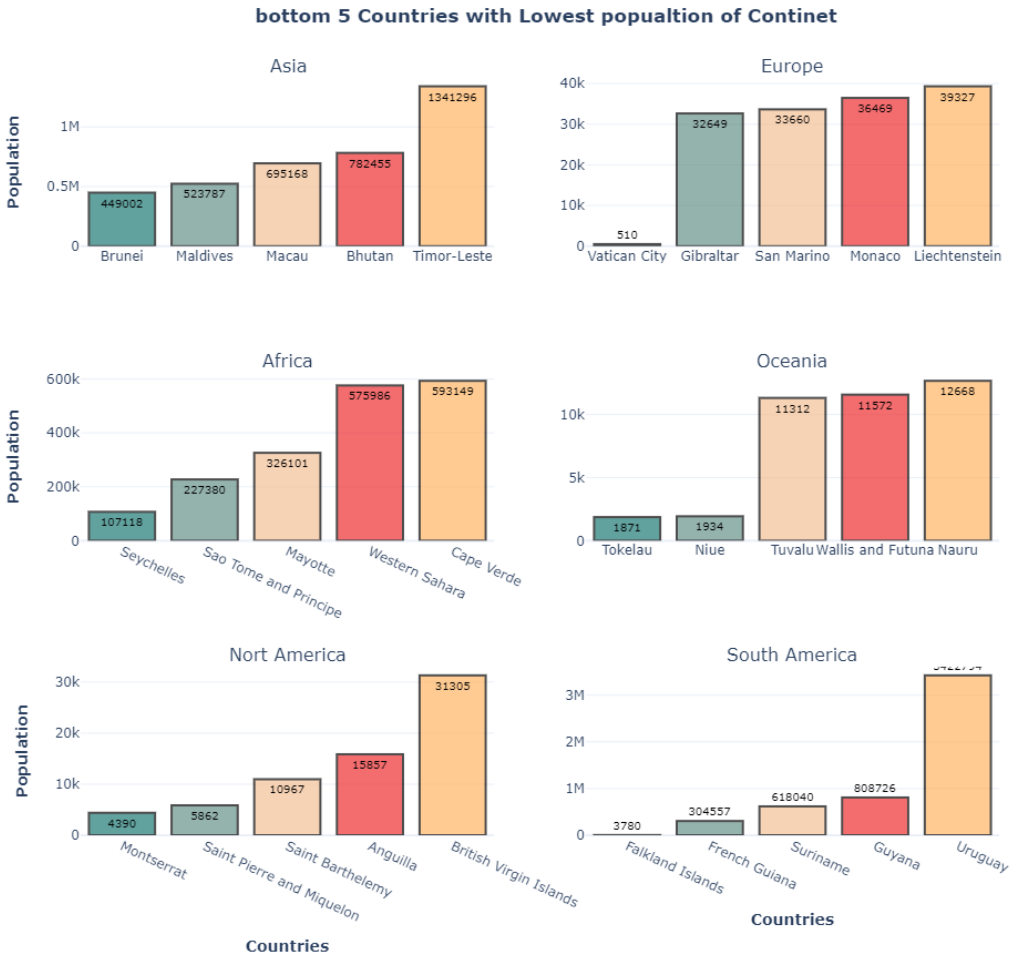
```
# Oceania
fig.add_trace(go.Bar(x = bott5_Oceania.index, y = bott5_Oceania['2022'], text = bott5_Oceania['2022'],textposition ='auto',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 2, col =2)

# Nort America
fig.add_trace(go.Bar(x = bott5_North_America.index, y = bott5_North_America['2022'], text = bott5_North_America['2022'],textposition ='auto',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 3, col =1)
# South America
fig.add_trace(go.Bar(x = bott5_South_America.index, y = bott5_South_America['2022'], text = bott5_South_America['2022'],
                     textposition ='outside',
              textfont = dict(size = 10,
                              color = 'black'),
              marker = dict(color = colors,
                            opacity = 0.7,
                            line_color = 'black',
                            line_width = 2)),
                                  row = 3, col =2)
```

```
# updating Layout
fig.update_layout(title = {'text': ' <b>bottom 5 Countries with Lowest popualtion of Continet</b> ',
                           'x':0.5,
                           'xanchor': 'center'},
                  width = 1000,
                  height = 900,
                  showlegend = False,
                  template = 'plotly_white')
#  xaxis title
fig.update_xaxes(tickangle = 0)
fig.update_xaxes(tickangle = 25, row=2, col=1)
fig.update_xaxes(tickangle = 25, row=3, col=1)
fig.update_xaxes(tickangle = 25, row=3, col=2)
fig.update_xaxes(title_text="<b>Countries</b>", row=3, col=1)
fig.update_xaxes(title_text="<b>Countries</b>", row=3, col=2)

# yaxis title
fig.update_yaxes(title_text="<b>Population</b>", row = 1, col=1)
fig.update_yaxes(title_text="<b>Population</b>", row = 2, col=1)
fig.update_yaxes(title_text="<b>Population</b>", row = 3, col=1)

iplot(fig)
```



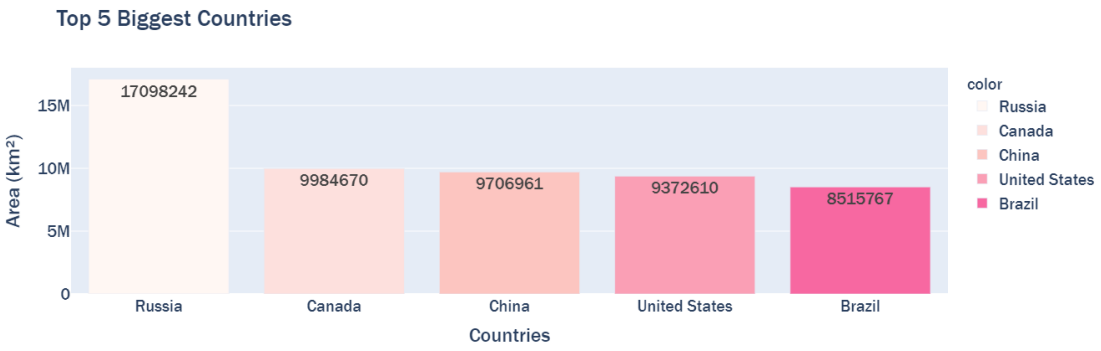
In Europe Continent the Country with Lowest Population is Vatican City with a population of 510 only.

- In Africa Continent the Country with Lowest Population is Seychelles with a population of 107.11 thousand.
- In Oceania Continent the Country with the Lowest Population is Tokelau with a population of 1.87 Thousand.

- In North America Continent the Country with the Lowest Population is Montserrat with a population of 4.39 Thousand.
- In South America Continent the Country with the Lowest Population is Falak land Island with a population of 3.78 Thousand.

What are the top 5 biggest Countries by area in the world?

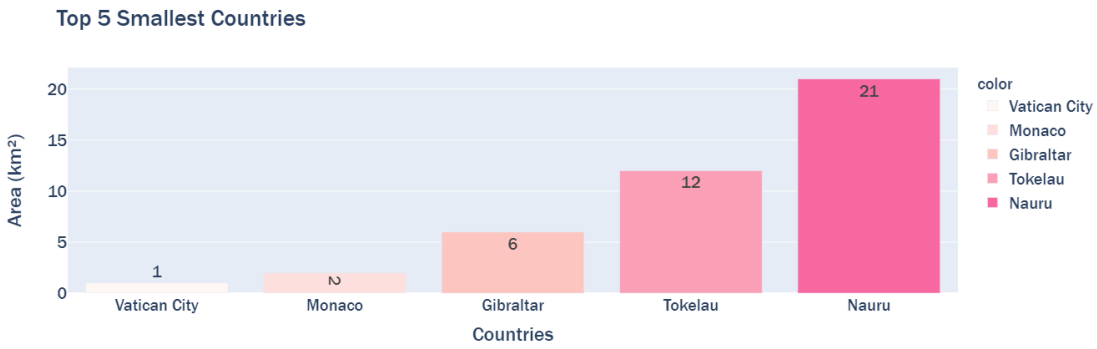
```
area_country = df.groupby(by = 'Country')['Area (km²)'].sum().sort_values(ascending=False).head()
fig = px.bar(y=area_country.values,
             x=area_country.index,
             color = area_country.index,
             color_discrete_sequence=px.colors.sequential.RdPu,
             text=area_country.values,
             title= 'Top 5 Biggest Countries',
             )
fig.update_layout(
    xaxis_title="Countries",
    yaxis_title="Area (km²)",
    font = dict(size=20,family="Franklin Gothic"))
fig.show()
```



Russia is the biggest country in the world with 17M km^2

What are the top 5 smallest Countries by area in the world?

```
area_country_2 = df.groupby(by = 'Country')['Area (km²)'].sum().sort_values(ascending=True).head()
fig = px.bar(y=area_country_2.values,
             x=area_country_2.index,
             color = area_country_2.index,
             color_discrete_sequence=px.colors.sequential.RdPu,
             text=area_country_2.values,
             title= 'Top 5 Smallest Countries',
             )
fig.update_layout(
    xaxis_title="Countries",
    yaxis_title="Area (km²)",
    font = dict(size=20,family="Franklin Gothic"))
fig.show()
```



Vatican City is the smallest country in the world with only 1 km^2 .

Counclusion :

- Africa is the Continent with the highest growth rate, Asia continent has the highest population and the largest World population percentage of 59.2
- The country with the highest population is China and the second is India, and the population growth difference of China from (1970 to 2022) is 603.35M and population growth difference of India from (1970 to 2022) is 859.67M and probably after 10 years India will surpass China and will become Top1 country with the highest Population.
- In Europe the population growth seems to be very low and Vatican City is the country with the lowest population from all over the world, Moldova has the highest population growth rate of 1.069 and Ukraine is the country with the lowest growth rate of 0.921.