

indian-economy

February 24, 2024

```
[2]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[3]: df = pd.read_csv("C:\\Users\\mathi\\Downloads\\indianEco.csv")
```

```
[9]: print(df.head())
```

	Year	Country Name	GDP (current US\$)	GDP per capita (current US\$)	\
0	1960	India	3.702988e+10		82
1	1961	India	3.923244e+10		85
2	1962	India	4.216148e+10		90
3	1963	India	4.842192e+10		101
4	1964	India	5.648029e+10		116

	GDP growth (annual %)	Imports of goods and services (% of GDP)	\
0	0.00		6.83
1	3.72		5.96
2	2.93		6.03
3	5.99		5.91
4	7.45		5.69

	Exports of goods and services (% of GDP)	\
0	4.46	
1	4.30	
2	4.17	
3	4.28	
4	3.73	

	Total reserves (includes gold, current US\$)	\
0	6.745366e+08	
1	6.663571e+08	
2	5.127918e+08	
3	6.078625e+08	
4	4.991451e+08	

	Inflation, consumer prices (annual %)	Population, total	\
0	1.78	445954579	

1	1.70	456351876
2	3.63	467024193
3	2.95	477933619
4	13.36	489059309

	Population growth (annual %)	Life expectancy at birth, total (years)
0	2.31	41.13
1	2.33	41.74
2	2.34	42.34
3	2.34	42.94
4	2.33	43.57

```
[10]: print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 61 entries, 0 to 60
Data columns (total 12 columns):
#   Column                                          Non-Null Count  Dtype
---  -
0   Year                                           61 non-null     int64
1   Country Name                                  61 non-null     object
2   GDP (current US$)                             61 non-null     float64
3   GDP per capita (current US$)                  61 non-null     int64
4   GDP growth (annual %)                        61 non-null     float64
5   Imports of goods and services (% of GDP)      61 non-null     float64
6   Exports of goods and services (% of GDP)      61 non-null     float64
7   Total reserves (includes gold, current US$)  61 non-null     float64
8   Inflation, consumer prices (annual %)         61 non-null     float64
9   Population, total                             61 non-null     int64
10  Population growth (annual %)                  61 non-null     float64
11  Life expectancy at birth, total (years)       61 non-null     float64
dtypes: float64(8), int64(3), object(1)
memory usage: 5.8+ KB
None
```

```
[11]: # Descriptive statistics
descriptive_stats = df.describe()
print(descriptive_stats)
```

	Year	GDP (current US\$)	GDP per capita (current US\$)	\
count	61.000000	6.100000e+01	61.000000	
mean	1990.000000	6.584725e+11	575.557377	
std	17.752934	8.129606e+11	584.079062	
min	1960.000000	3.702988e+10	82.000000	
25%	1975.000000	9.952590e+10	161.000000	
50%	1990.000000	2.882084e+11	340.000000	
75%	2005.000000	8.203816e+11	715.000000	
max	2020.000000	2.831552e+12	2101.000000	

	GDP growth (annual %)	Imports of goods and services (% of GDP) \
count	61.000000	61.000000
mean	4.938197	12.746393
std	3.344891	8.155110
min	-7.250000	3.710000
25%	3.720000	6.590000
50%	5.530000	8.570000
75%	7.450000	19.640000
max	9.630000	31.260000

	Exports of goods and services (% of GDP) \
count	61.000000
mean	10.885574
std	7.060458
min	3.310000
25%	5.200000
50%	7.050000
75%	18.690000
max	25.430000

	Total reserves (includes gold, current US\$) \
count	6.100000e+01
mean	9.802227e+10
std	1.497102e+11
min	4.991451e+08
25%	2.324650e+09
50%	1.151174e+10
75%	1.378248e+11
max	5.902274e+11

	Inflation, consumer prices (annual %)	Population, total \
count	61.000000	6.100000e+01
mean	7.413279	8.913946e+08
std	4.940153	2.974496e+08
min	-7.630000	4.459546e+08
25%	4.010000	6.235242e+08
50%	6.670000	8.704522e+08
75%	10.020000	1.154639e+09
max	28.600000	1.396387e+09

	Population growth (annual %)	Life expectancy at birth, total (years)
count	61.000000	61.000000
mean	1.927705	57.146230
std	0.419024	8.459559
min	0.960000	41.130000
25%	1.620000	50.630000
50%	2.150000	57.660000

75%	2.260000	64.310000
max	2.340000	69.730000

```
[12]: # Remove repeated values in the "Country Name" column
df['Country Name'] = df['Country Name'].unique()[0]
```

```
[13]: # Remove leading/trailing whitespaces from column names
df.columns = df.columns.str.strip()
```

```
[14]: # Convert relevant columns to numeric, handling errors by coercing them to NaN
numeric_columns = ['GDP (current US$)', 'GDP per capita (current US$)', 'GDP_
↳ growth (annual %)',
                    'Imports of goods and services (% of GDP)', 'Exports of_
↳ goods and services (% of GDP)',
                    'Total reserves (includes gold, current US$)', 'Inflation,_
↳ consumer prices (annual %)',
                    'Population, total', 'Population growth (annual %)', 'Life_
↳ expectancy at birth, total (years)']

df[numeric_columns] = df[numeric_columns].apply(pd.to_numeric, errors='coerce')
```

```
[24]: from statistics import mode

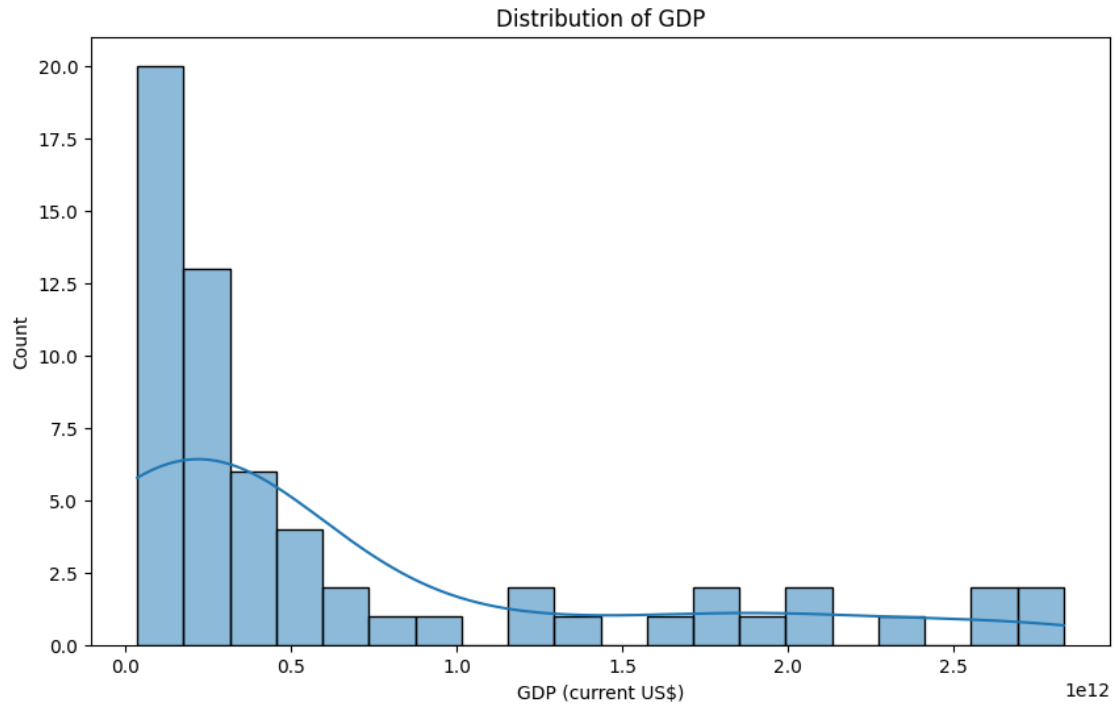
mode_values = df.apply(mode)
print("\nMode Values:")
print(mode_values)
```

```
Mode Values:
Year                                1960
Country Name                       India
GDP (current US$)                 37029883876.0
GDP per capita (current US$)              90
GDP growth (annual %)                7.86
Imports of goods and services (% of GDP)  6.83
Exports of goods and services (% of GDP)  4.03
Total reserves (includes gold, current US$) 674536630.9
Inflation, consumer prices (annual %)      1.78
Population, total                   445954579
Population growth (annual %)          2.23
Life expectancy at birth, total (years)    41.13
dtype: object
```

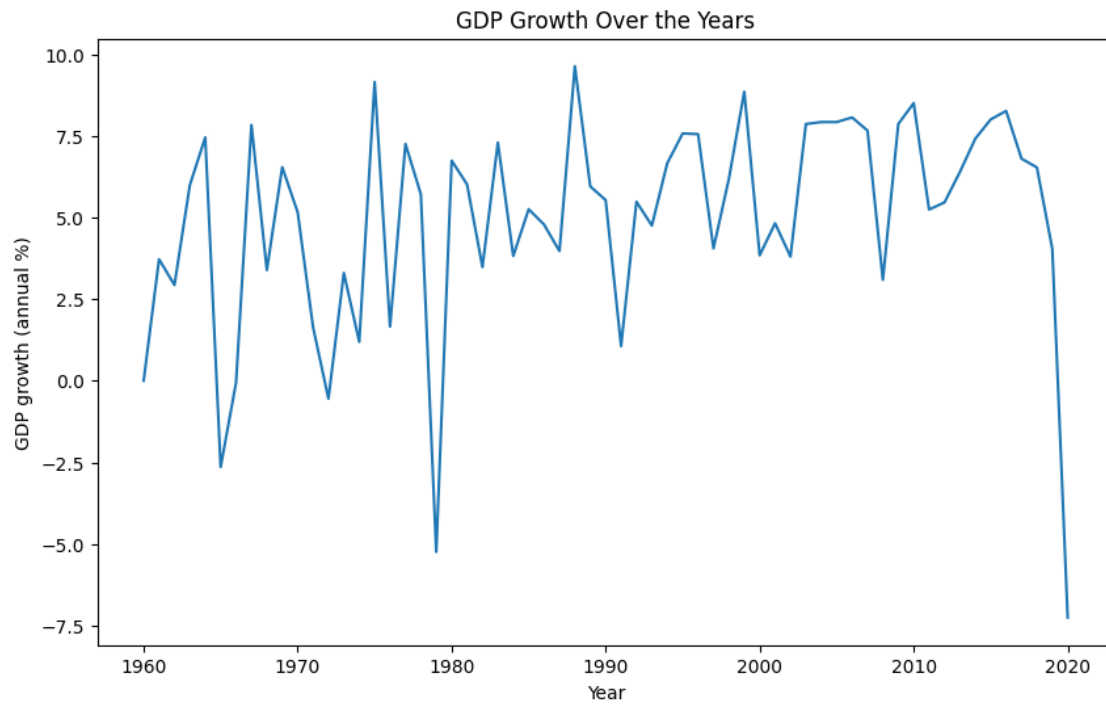
```
[25]: # Exploratory Data Analysis (EDA)

# Distribution of GDP
plt.figure(figsize=(10, 6))
```

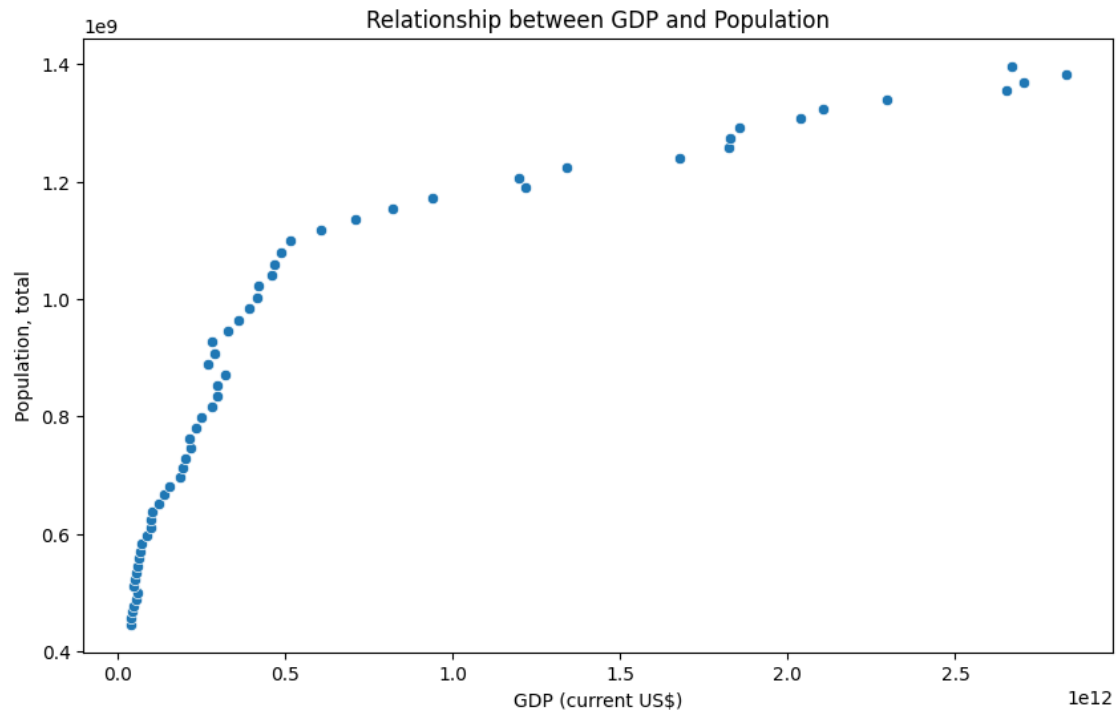
```
sns.histplot(df['GDP (current US$)'], bins=20, kde=True)
plt.title('Distribution of GDP')
plt.xlabel('GDP (current US$)')
plt.show()
```



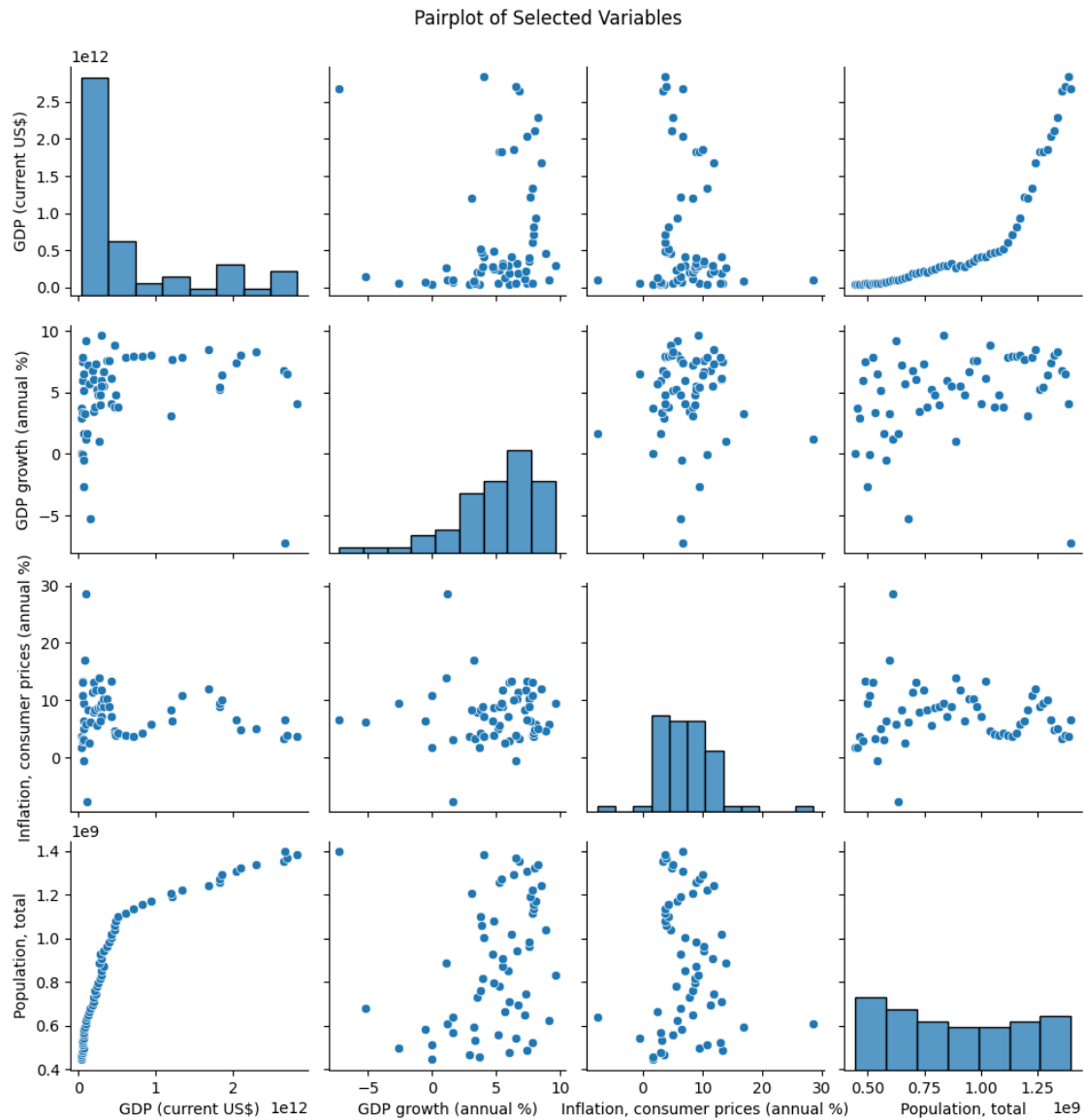
```
[26]: # GDP Growth over the years
plt.figure(figsize=(10, 6))
sns.lineplot(x='Year', y='GDP growth (annual %)', data=df)
plt.title('GDP Growth Over the Years')
plt.xlabel('Year')
plt.ylabel('GDP growth (annual %)')
plt.show()
```



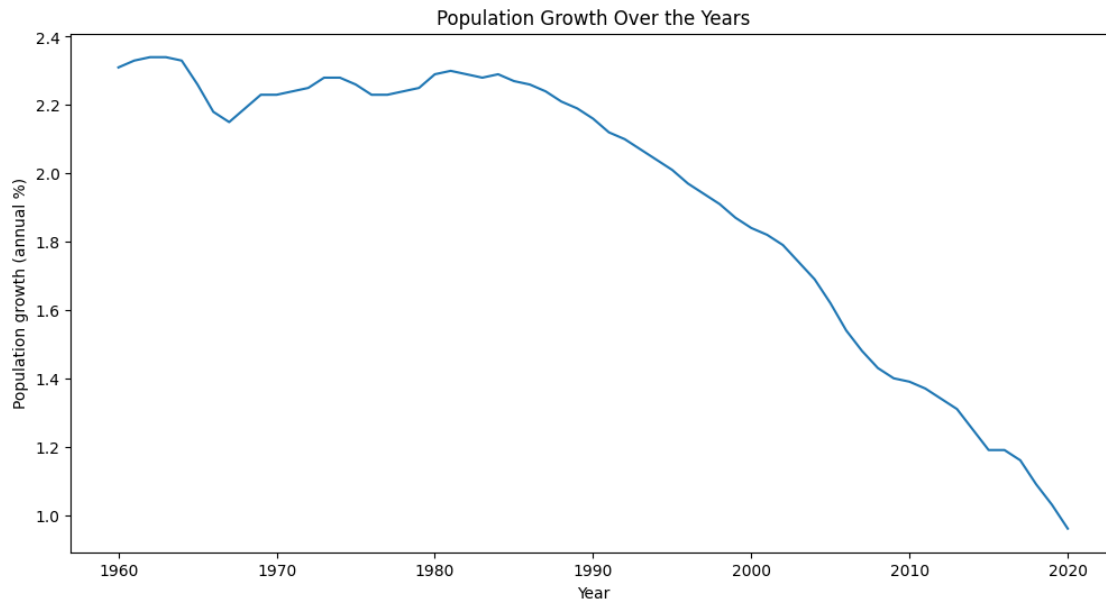
```
[27]: # Relationship between GDP and Population
plt.figure(figsize=(10, 6))
sns.scatterplot(x='GDP (current US$)', y='Population, total', data=df)
plt.title('Relationship between GDP and Population')
plt.xlabel('GDP (current US$)')
plt.ylabel('Population, total')
plt.show()
```



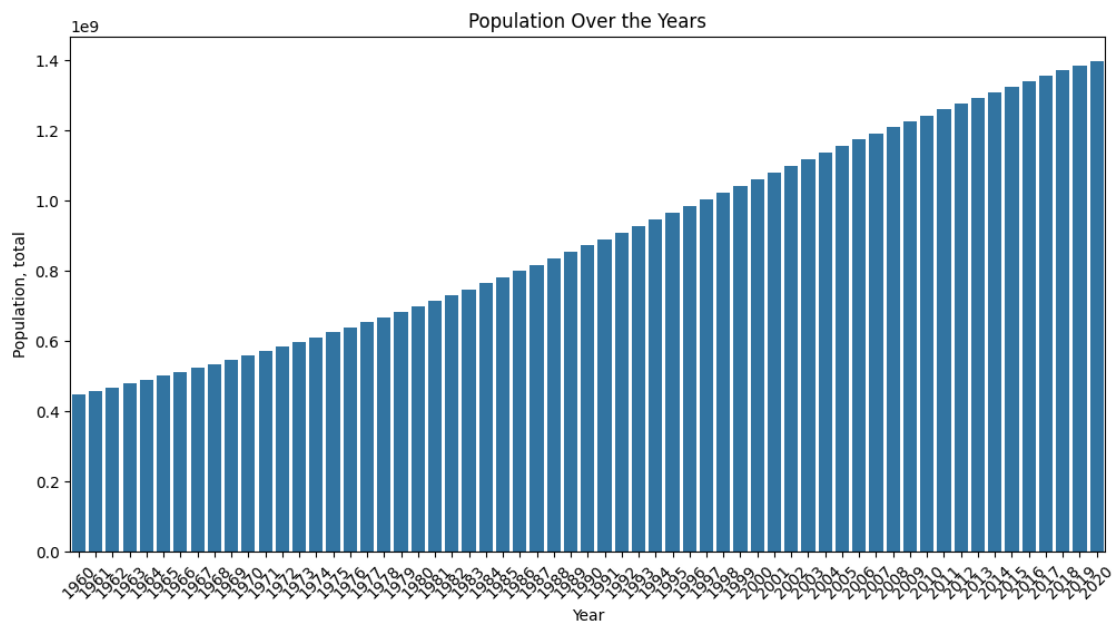
```
[28]: # Pairplot for selected variables
selected_columns = ['GDP (current US$)', 'GDP growth (annual %)', 'Inflation,
↳ consumer prices (annual %)', 'Population, total']
sns.pairplot(df[selected_columns])
plt.suptitle('Pairplot of Selected Variables', y=1.02)
plt.show()
```



```
[30]: # Visualize Population growth over the years
plt.figure(figsize=(12, 6))
sns.lineplot(x='Year', y='Population growth (annual %)', data=df)
plt.title('Population Growth Over the Years')
plt.show()
```

```
[32]: # Bar chart for Population over the years
plt.figure(figsize=(12, 6))
sns.barpplot(x='Year', y='Population, total', data=df)
plt.title('Population Over the Years')
plt.xlabel('Year')
plt.ylabel('Population, total')
plt.xticks(rotation=45)
plt.show()
```



[]: