BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

Course – Advanced Data Visualization (ADV)

Name	Tejal Subhash Komb
UID	2021600037
Batch	A
Lab no	2

Aim :- Create advanced charts using Python to be performed on the dataset - Socio economic data

Objectives :-

1. To Visualize Cost of Living Data
2. To Compare Indices Across Countries
3. To Identify Trends and Patterns
4. To Provide Insights and Interpretations

Dataset :-

Rank	Country	Cost of Liv	Rent Index	Cost of Liv	Groceries	Restauran	Local Purchasing Power Index
1	Switzerlan	101.1	46.5	74.9	109.1	97	158.7
2	Bahamas	85	36.7	61.8	81.6	83.3	54.6
3	Iceland	83	39.2	62	88.4	86.8	120.3
4	Singapore	76.7	67.2	72.1	74.6	50.4	111.1
5	Barbados	76.6	19	48.9	80.8	69.4	43.5
6	Norway	76	26.2	52.1	79	73.5	114.7
7	Denmark	72.3	26.4	50.2	64.8	81.3	127.2
8	Hong Kong	70.8	59.4	65.3	84.6	46.2	109.3
9	United Sta	70.4	41.7	56.6	75	67.2	142.3
10	Australia	70.2	33.4	52.5	77.3	62.5	127.4
11	Austria	65.1	22.5	44.7	66.4	59.3	102.4
12	Canada	64.8	33.2	49.6	71.2	61.7	103.7
13	New Zeala	64.6	25.9	46	71.9	57.1	121
14	Ireland	64.4	42.3	53.8	59.3	65.3	101.1
15	France	63.7	21	43.2	69.5	56.3	102.4
16	Puerto Ric	63.3	19.5	42.3	64.1	53.1	103.5
17	Finland	63.2	19.7	42.4	63	63.6	118
18	Netherlan	63.1	33.5	48.9	59.9	61.2	124.9
19	Israel	62.7	27.1	45.6	59.6	68	99.4

BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

Dataset description :-

Cost of Living Index by Country, 2024 Mid Year data An index of 100 reflects the same living cost as in New York City, United States.
Here's a breakdown of each index and its meaning:
Cost of Living Index (Excl. Rent): This index indicates the relative prices of consumer goods like groceries, restaurants, transportation, and utilities. It excludes accommodation expenses such as rent or mortgage. For instance, a city with a Cost of Living Index of 120 is estimated to be 20% more expensive than New York City (excluding rent).
Rent Index: This index estimates the prices of renting apartments in a city compared to New York City. If the Rent Index is 80, it suggests that the average rental prices in that city are approximately 20% lower than those in New York City.
Cost of Living Plus Rent Index: This index estimates consumer goods prices, including rent, in comparison to New York City.
Groceries Index: This index provides an estimation of grocery prices in a city relative to New York City. Numbeo uses item weights from the "Markets" section to calculate this index for each city.
Restaurants Index: This index compares the prices of meals and drinks in restaurants and bars to those in NYC.
Local Purchasing Power: This index indicates the relative purchasing power in a given city based on the average net salary. A domestic purchasing power of 40 means that residents with an average salary can afford, on average, 60% less goods and services compared to residents of New York City with an average salary.

Implementation :-

```
from wordcloud import WordCloud
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
import plotly.express as px
import plotly.graph_objects as go
import plotly.figure_factory as ff
import plotly.io as pio

#read the dataset
df = pd.read_csv('D:\Downloads new location\Cost_of_Living_Index_by_Country_2024.csv')
```

```
df.head()

df.columns

"""Box and Whisker Plot"""

plt.figure(figsize=(12, 6))
sns.boxplot(data=df[['Cost of Living Index', 'Rent Index', 'Groceries Index', 'Restaurant
Price Index', 'Local Purchasing Power Index']])
plt.title('Distribution of Different Indices')
plt.xticks(rotation=45)
plt.show()

"""**Violin Plot**"""

df_top_30 = df.sort_values(by='Cost of Living Index', ascending=False).head(30)

# Step 2: Create a Violin Plot for 'Cost of Living Index' for the top 30 countries
plt.figure(figsize=(12, 8))
sns.violinplot(x='Country', y='Cost of Living Index', data=df_top_30, inner='quartile')
plt.title('Violin Plot of Cost of Living Index for Top 30 Countries')
plt.xticks(rotation=90)  # Rotate x-axis labels for better visibility
plt.show()

"""**Regression Plot**"""

plt.figure(figsize=(10, 6))
sns.regplot(x='Cost of Living Index', y='Rent Index', data=df)
plt.title('Regression Plot: Cost of Living Index vs. Rent Index')
```

```
plt.show()

"""**3D Scatter Plot**"""

fig = px.scatter_3d(df, x='Cost of Living Index', y='Rent Index', z='Local Purchasing Power Index', color='Country')
fig.update_layout(title='3D Scatter Plot of Cost of Living, Rent, and Purchasing Power Indices')
fig.show()

"""**Tree Map**"""

fig = px.treemap(df, path=['Country'], values='Cost of Living Index', title='Treemap of Cost of Living Index by Country')
fig.show()

"""**Word Cloud**"""

text = " ".join(df['Country'].tolist())

# Generate word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)
# Plot word cloud
plt.figure(figsize=(10, 6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Word Cloud of Countries')
plt.show()
```

```
"""**Jitter Plot**"""
# Jitter Plot for Cost of Living Index
plt.figure(figsize=(12, 8))
sns.stripplot(x='Country', y='Cost of Living Index', data=df_top_30, jitter=True)
plt.title('Jitter Plot of Cost of Living Index for Top 30 Countries')
plt.xticks(rotation=90)
plt.show()

"""**Line Chart**"""
# Assuming 'Rank' is a proxy for time or ranking change
plt.figure(figsize=(12, 6))
plt.plot(df['Rank'], df['Cost of Living Index'], marker='o')
plt.title('Cost of Living Index by Rank')
plt.xlabel('Rank')
plt.ylabel('Cost of Living Index')
plt.grid(True)
plt.show()

"""**Area Chart**"""
# Area chart for Cost of Living Index by Rank
plt.figure(figsize=(12, 6))
plt.fill_between(df['Rank'], df['Cost of Living Index'], color="skyblue", alpha=0.4)
plt.plot(df['Rank'], df['Cost of Living Index'], color="Slateblue", alpha=0.6, linewidth=2)
plt.title('Area Chart of Cost of Living Index by Rank')
plt.xlabel('Rank')
plt.ylabel('Cost of Living Index')
plt.show()

"""**Waterfall Chart**"""
```

```
fig = go.Figure(go.Waterfall(
    name="20", orientation="v",
    measure=["relative", "relative", "total", "relative", "total"],
    x=["Start", "Increased Living Cost", "Taxes", "Expenses", "Net"],
    textposition="outside",
    y=[500, 150, -100, -60, 0],
    connector={"line": {"color": "rgb(63, 63, 63)"}},
))
fig.update_layout(title="Waterfall Chart Example")
fig.show()

"""**Donut Chart**"""
values = [450, 300, 150, 100]
names = ["Cost of Living Index", "Rent Index", "Groceries Index", "Restaurant Price Index"]

# Plot donut chart
fig = px.pie(values=values, names=names, hole=0.3, title='Index Distribution')
fig.show()

"""**Funnel Chart**"""
stages = ['Total Population', 'Employed', 'Educated', 'Wealthy']
values = [1000, 800, 600, 400]
# Plot funnel chart
fig = px.funnel(y=stages, x=values, title='Socio-Economic Funnel Chart')
fig.show()
```

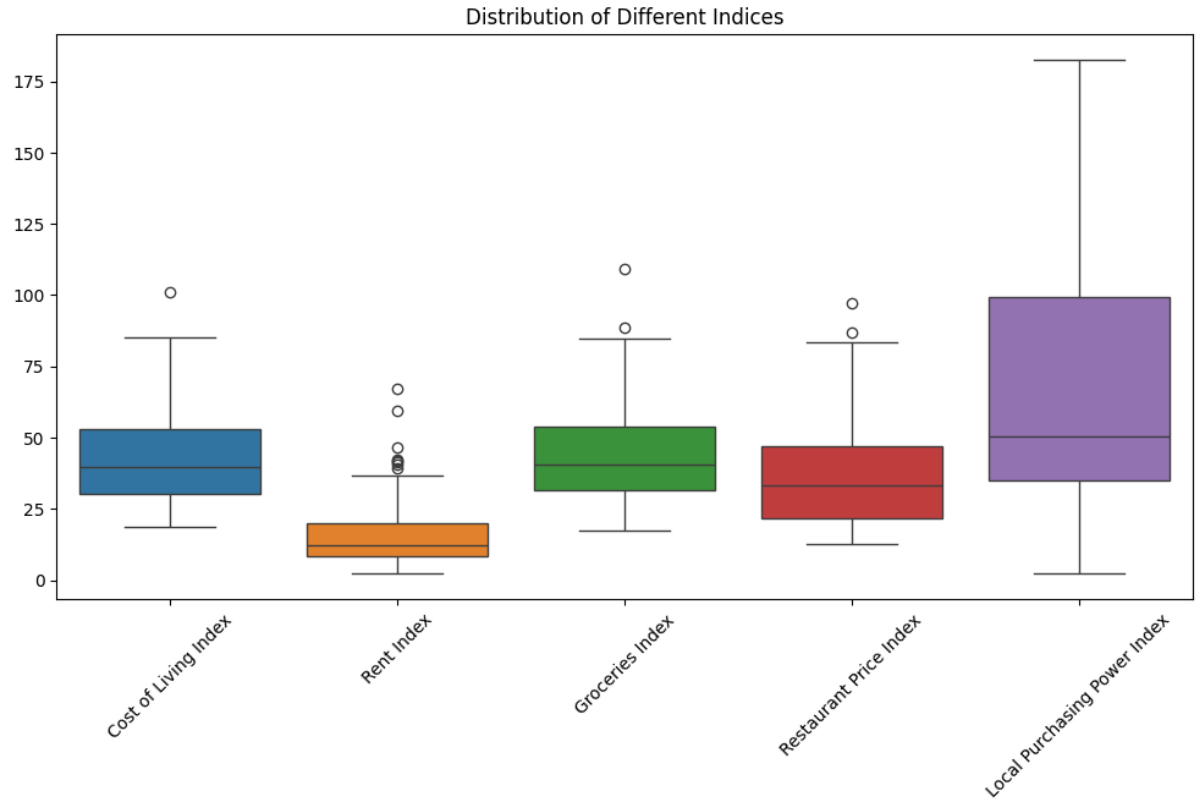
BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

Output :-

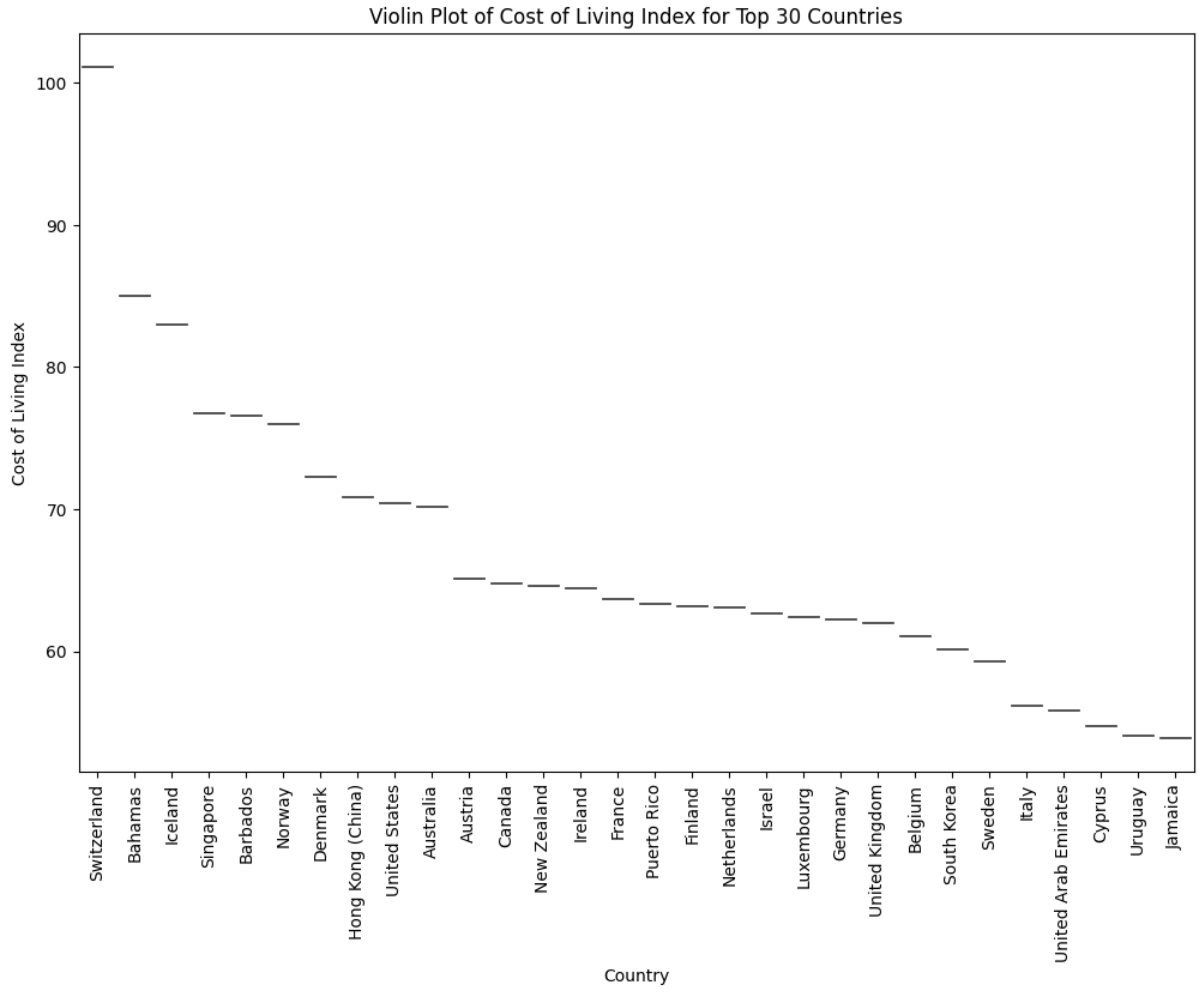
1] Box and Whisker Plot :-

A box plot (or box-and-whisker plot) visually summarizes data by displaying the median, quartiles, and outliers of the data set. It show the distribution of various indices (Cost of Living, Rent, Groceries, Restaurant, and Local Purchasing Power) across different countries. Each "box" represents the interquartile range (IQR), with the median as a line inside the box, and "whiskers" extending to show variability outside the upper and lower quartiles. Outliers, if any, are plotted as individual points.



2] Violin Plot :-

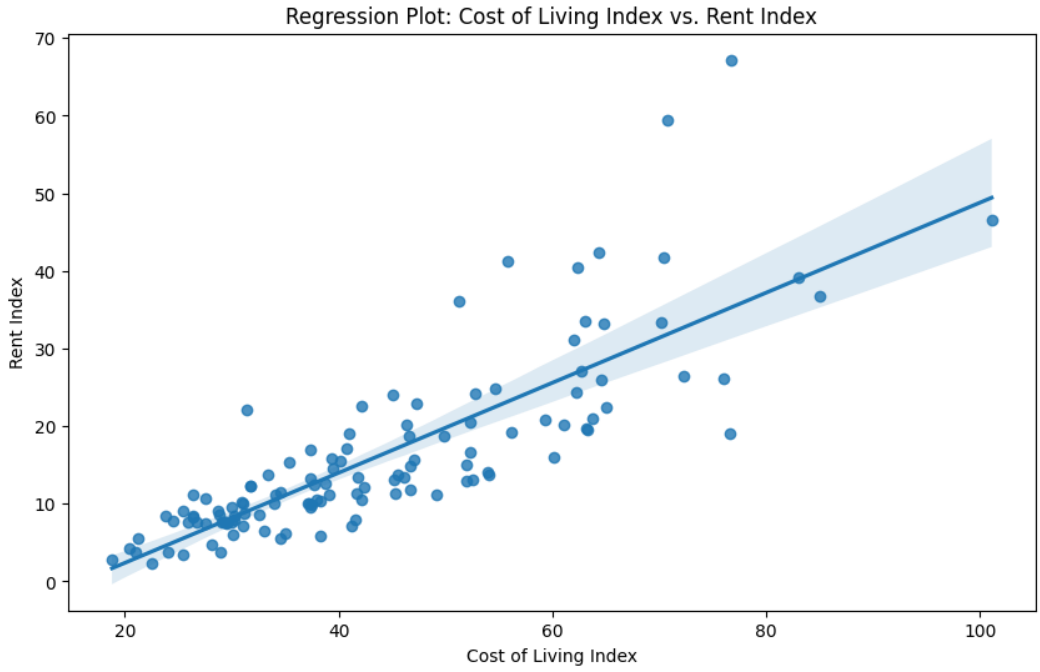
A violin plot combines elements of a box plot and a kernel density plot to show the distribution of data points.
What it Shows: Here, the violin plot shows the distribution of the "Cost of Living Index" for the top 30 countries, with the shape of the "violin" indicating the density of data points at different values. This allows for a better understanding of where most data points are concentrated within the range.



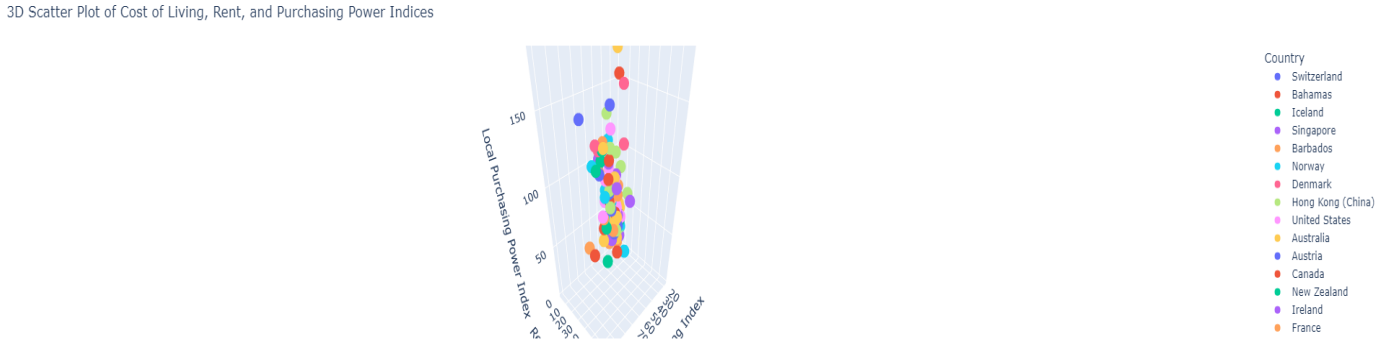
BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

3] Regression Plot :- A regression plot visualizes the relationship between two variables along with a regression line. This plot shows the correlation between the "Cost of Living Index" and "Rent Index." The regression line helps indicate the trend, showing whether higher living costs tend to be associated with higher rent.



4] 3D Scatter Plot :- • A 3D scatter plot plots data across three dimensions, with each axis representing a different variable. This plot displays the "Cost of Living Index," "Rent Index," and "Local Purchasing Power Index" for each country. It allows you to see relationships and patterns across three indices, with color indicating different countries for easy differentiation.

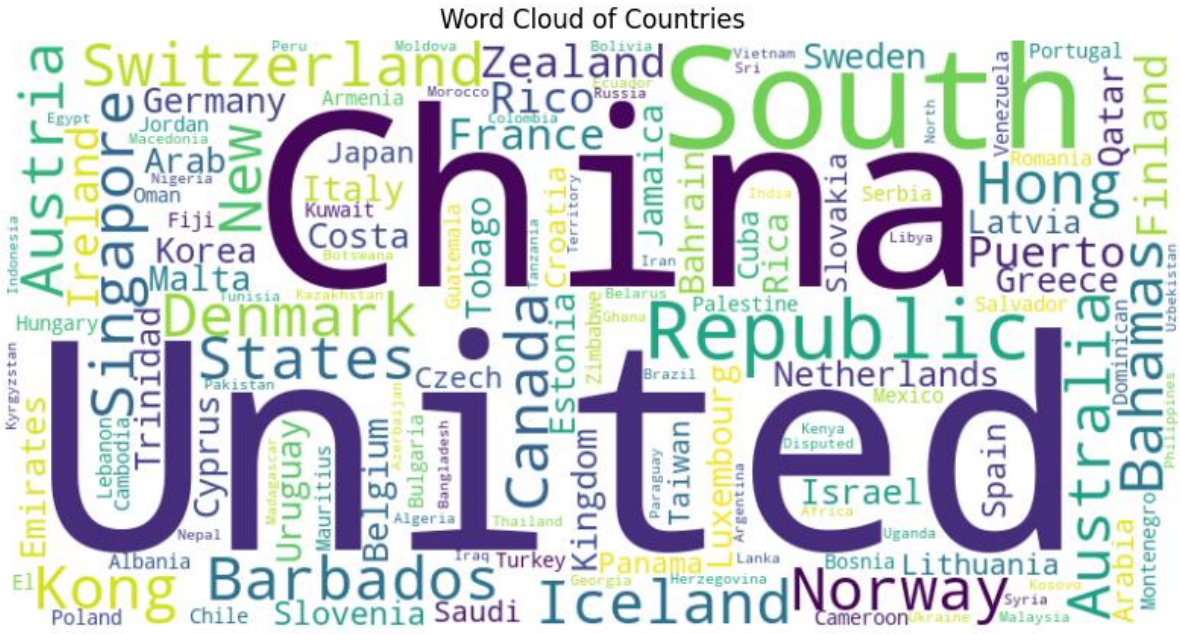


BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

5] Tree Map :- A tree map uses nested rectangles to represent hierarchical data, where the size of each rectangle reflects a numerical value. The tree map displays the "Cost of Living Index" by country, with the size of each rectangle representing the relative cost of living. It helps compare the cost of living visually across different countries.

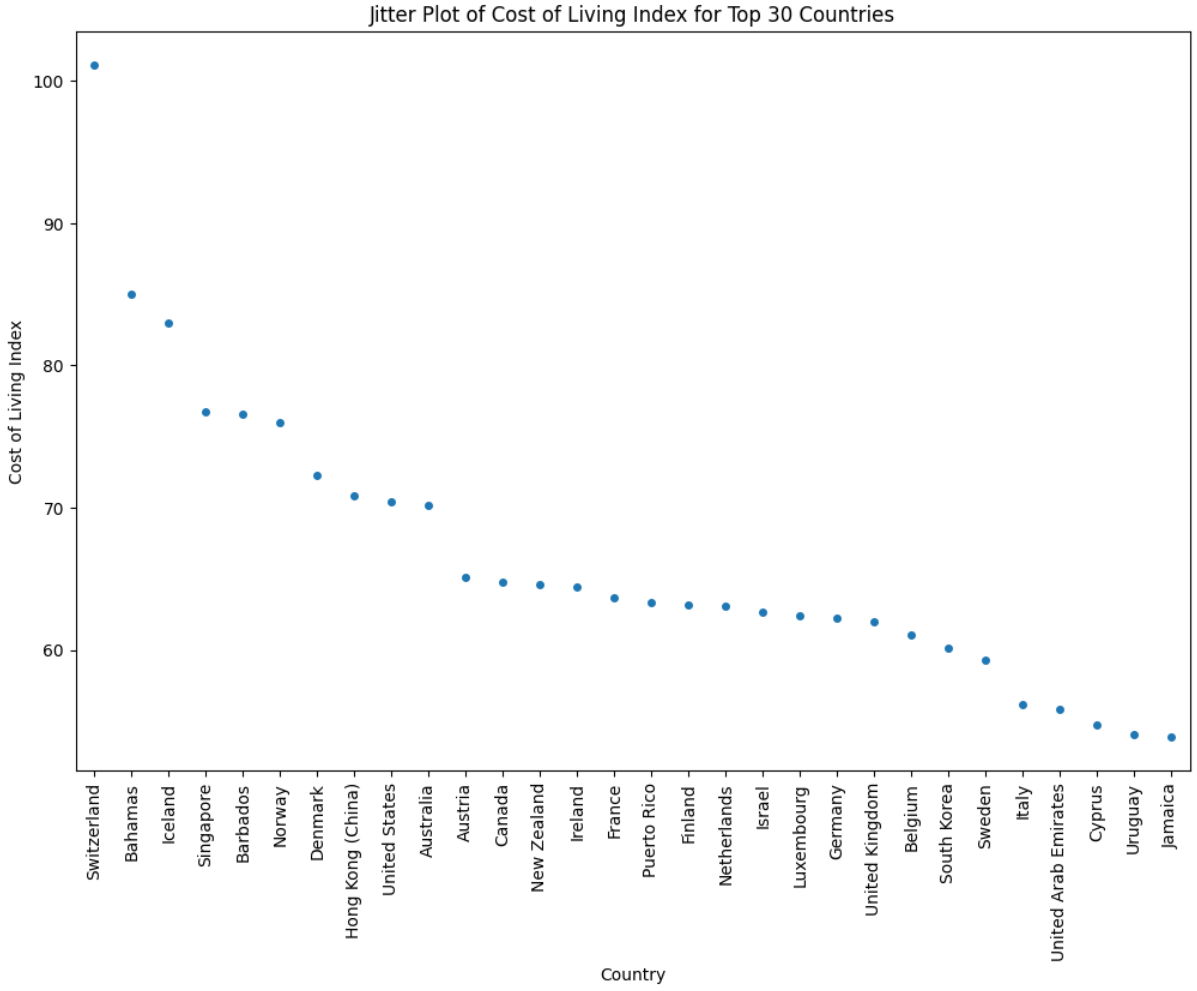
6] Word Cloud :- A word cloud is a visual representation of text data where the size of each word reflects its frequency or importance. Here, it shows the names of countries, with each country's name sized equally. This gives a visual impression of the dataset's geographic diversity.



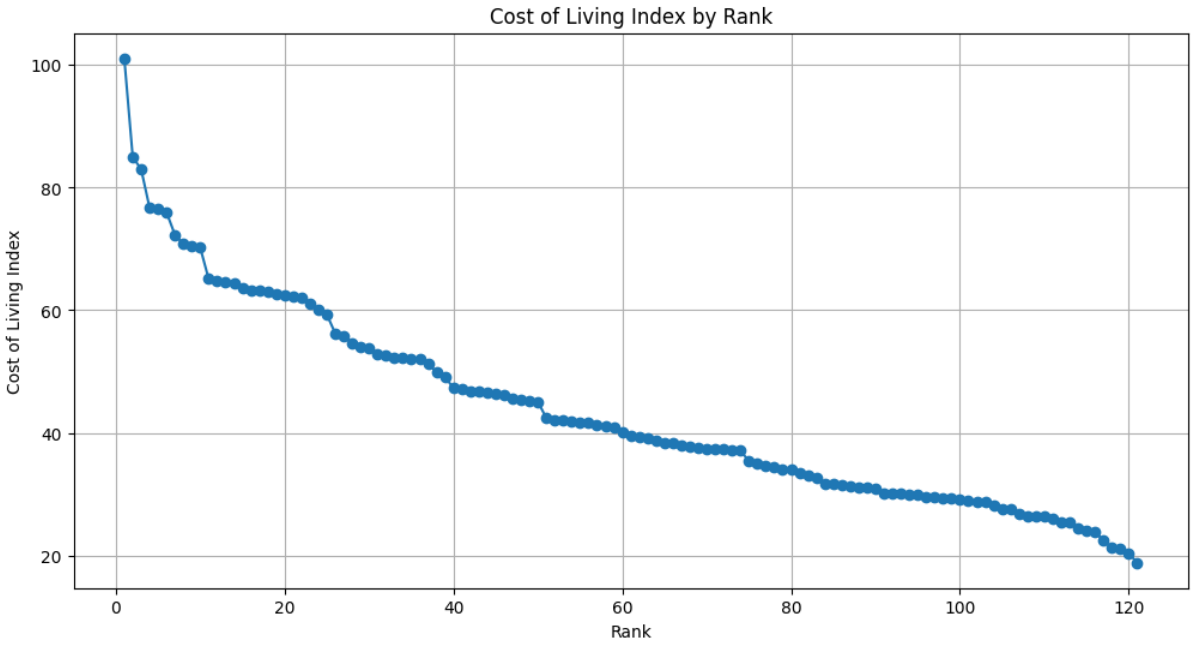
BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

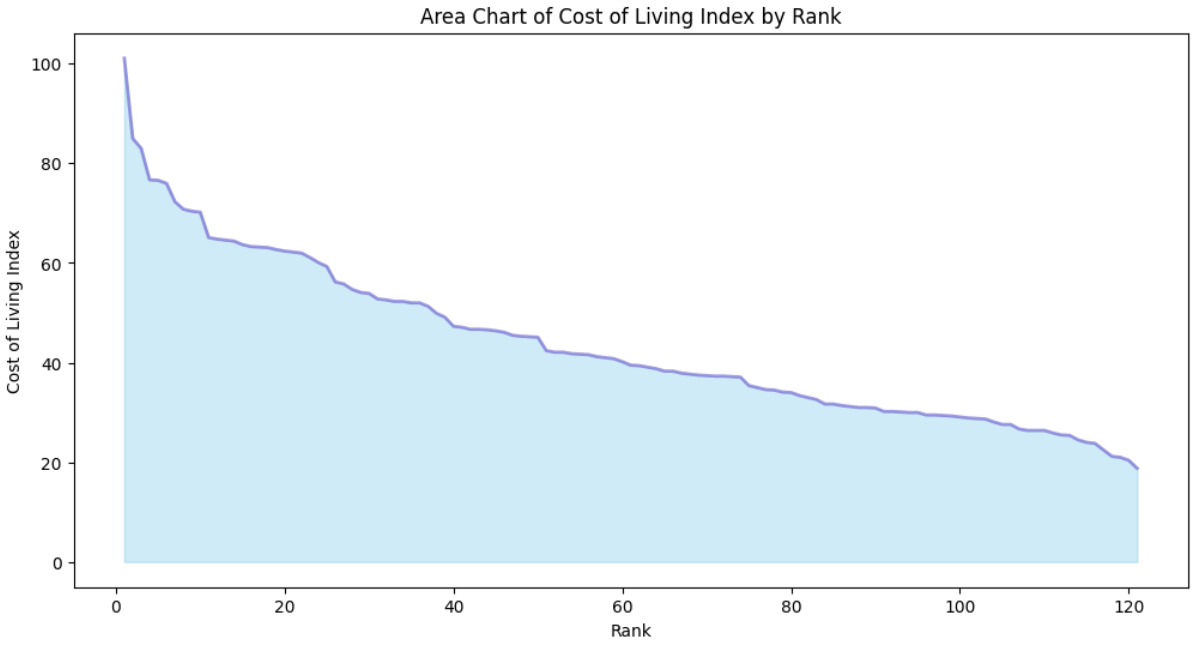
7] Jitter Plot :- A jitter plot is similar to a scatter plot but adds random noise to the position of points to prevent overlap and make individual points visible. This jitter plot displays the "Cost of Living Index" for the top 30 countries, allowing you to see individual data points without overlapping, especially if multiple countries have similar costs of living.



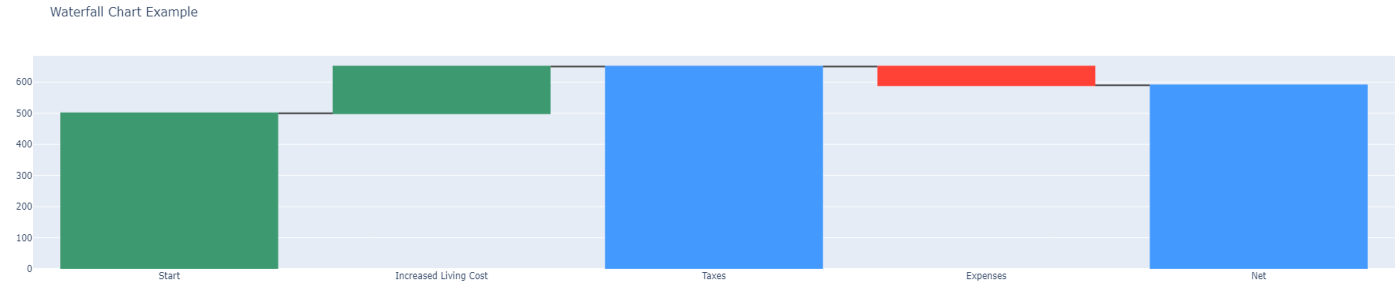
8] Line Chart :- A line chart shows data points connected by lines, often used to track changes over time or across ranks. Here, it shows the "Cost of Living Index" by rank, providing insight into how living costs vary as ranks change.



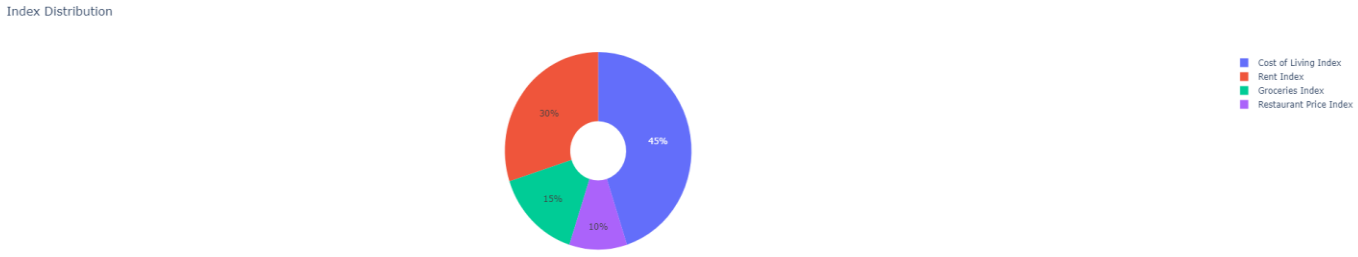
9] Area Chart :- An area chart is similar to a line chart, but it fills the area below the line, emphasizing the magnitude of values. This area chart highlights the "Cost of Living Index" across ranks, offering a clearer view of cost magnitude.



10] Waterfall Chart :- A waterfall chart shows the cumulative effect of sequentially introduced values, helping to visualize how individual contributions affect a total. This example waterfall chart illustrates hypothetical changes in living costs, taxes, and expenses to visualize net impact.



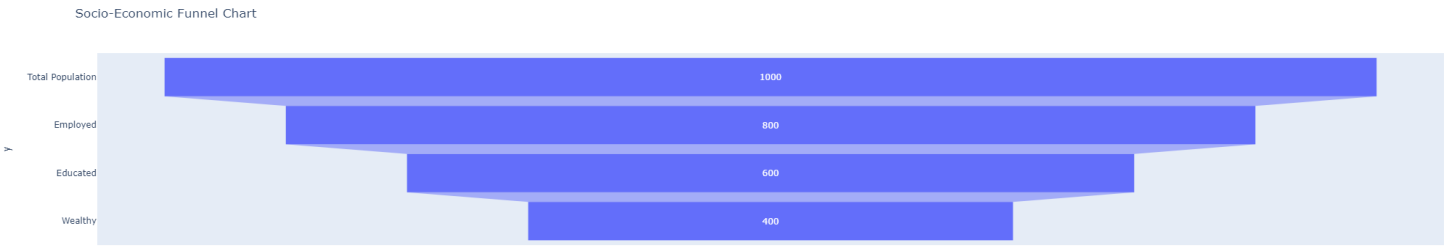
11] Donut Chart :- A donut chart is a variant of the pie chart with a hole in the middle, used to display part-to-whole relationships. This donut chart visualizes the distribution of different indices, providing an overall perspective of how each index contributes to the total.



BHARATIYA VIDYA BHAVAN'S
SARDAR PATEL INSTITUTE OF TECHNOLOGY
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai – 400093-India
(Autonomous College Affiliated to University of Mumbai)

Department of Computer Science and Engineering

12] Funnel Chart :- A funnel chart shows values at different stages in a process, narrowing to represent drop-offs or changes at each stage. This chart, though hypothetical here, could be used to represent stages in socio-economic metrics, like population, employment, and education, to show drop-off rates.



Conclusion :-

The experiment effectively visualized socio-economic data, highlighting key trends and patterns in cost of living, rent, groceries, and purchasing power indices across countries. Advanced charts provided insights into relationships between these indices, with notable findings including a positive correlation between cost of living and rent, and higher purchasing power in developed regions despite high costs. These visualizations offer valuable comparative insights for economic analysis and decision-making on a global scale.