Bharatiya Vidya Bhavan's

SARDAR PATEL INSTITUTE OF TECHNOLOGY

Advanced Data Visualization

Experiment no. 2

Submitted To

Prof. Pranav Nerurkar

Submitted By

Name : Dhanashree Otari

Class : Btech Computer Engineering

Batch : A

UID : 2021300086

1. Dataset

This dataset contains information about the cost of living, rent, groceries, restaurant prices, and local purchasing power indices for 28 countries. The data ranks the countries based on their overall cost of living index and provides a comprehensive comparison of various living expenses across different countries.

For details,click on
https://www.kaggle.com/datasets/myrios/cost-of-living-index-by-country-by-number-2024

2. Description

Cost of Living Index by Country, 2024 Mid Year data
Data scraped from Numbeo: www.numbeo.com/cost-of-living/rankings_by_country.jsp
All credits to Numbeo: www.numbeo.com/cost-of-living/
An index of 100 reflects the same living cost as in New York City, United States.
As of 2024 Mid Year data, in NYC,
A family of four estimated monthly costs are $6,074.40 without rent.
A single person's estimated monthly costs are $1,640.90 without rent.

3. Metadata

Rank:

- *Description*: The position of the country in the ranking based on the cost of living index.
- *Data Type*: Integer

Country:

- *Description*: The name of the country.
- *Data Type*: String

Cost of Living Index:

- *Description*: A numerical index representing the overall cost of living in the country. The higher the value, the more expensive it is to live in that country.
- *Data Type*: Float

Rent Index:

- *Description*: A numerical index representing the cost of renting accommodation in the country.
- *Data Type*: Float

Cost of Living Plus Rent Index:

- *Description*: A combined index of the cost of living and rent, providing a comprehensive measure of living expenses including accommodation.
- *Data Type*: Float

Groceries Index:

- *Description*: A numerical index representing the cost of groceries in the country.
- *Data Type*: Float

Restaurant Price Index:

- *Description*: A numerical index representing the cost of dining out in restaurants in the country.

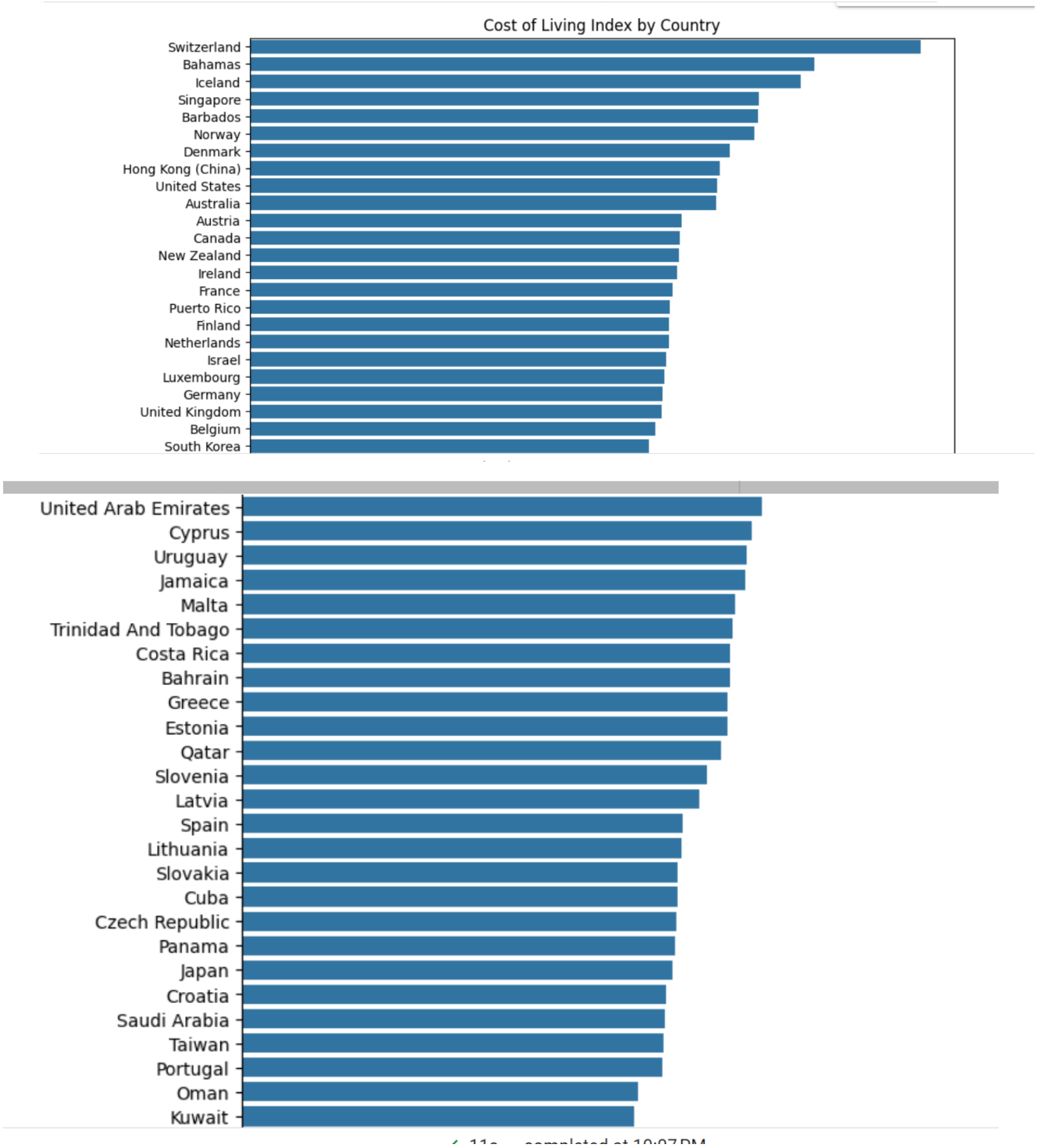
- *Data Type*: Float

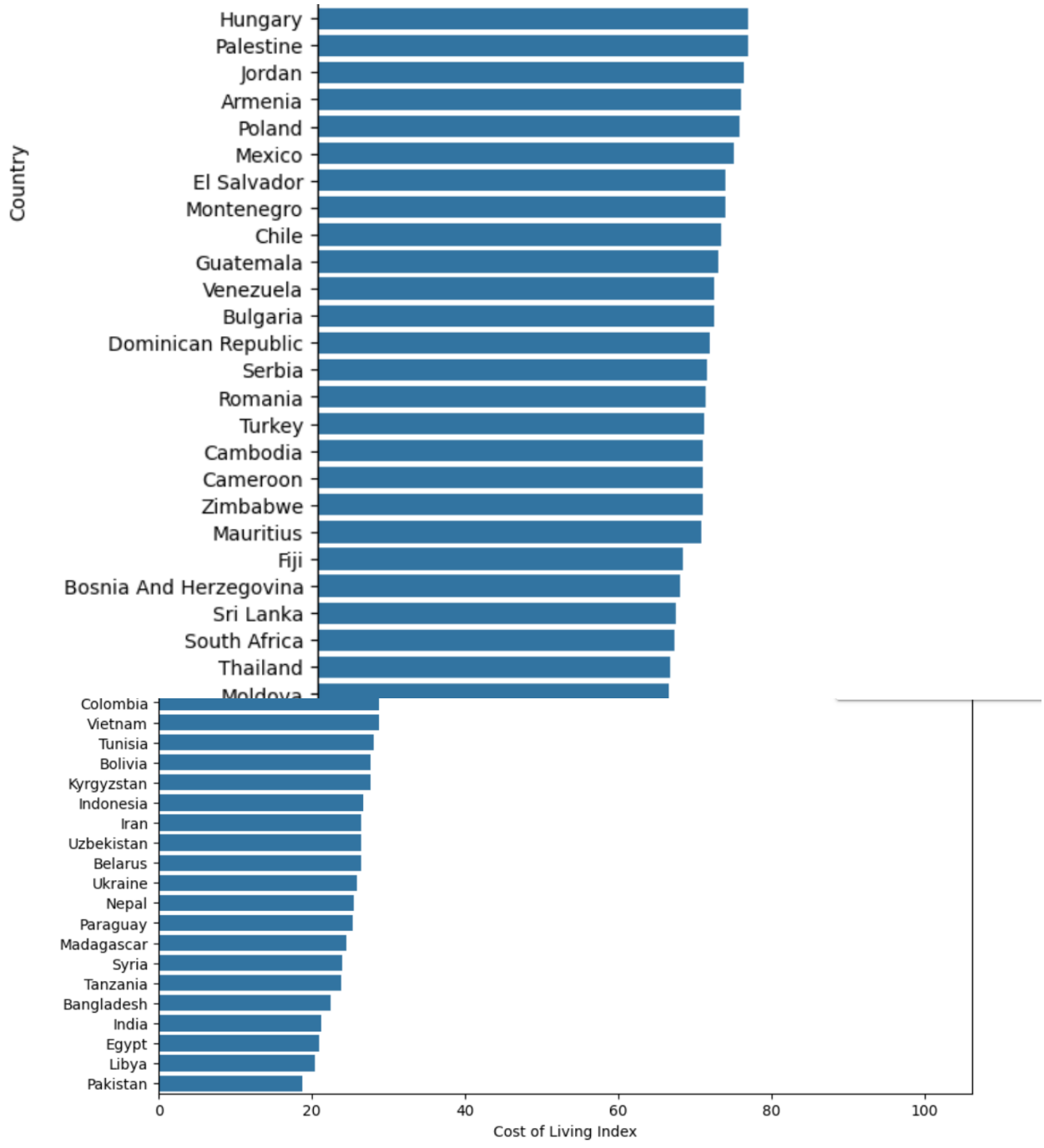
Local Purchasing Power Index:

- *Description*: A numerical index representing the purchasing power of the local population in the country, taking into account average income and cost of goods and services.
- *Data Type*: Float

4. Visualizations and Observations

1. Word Chart



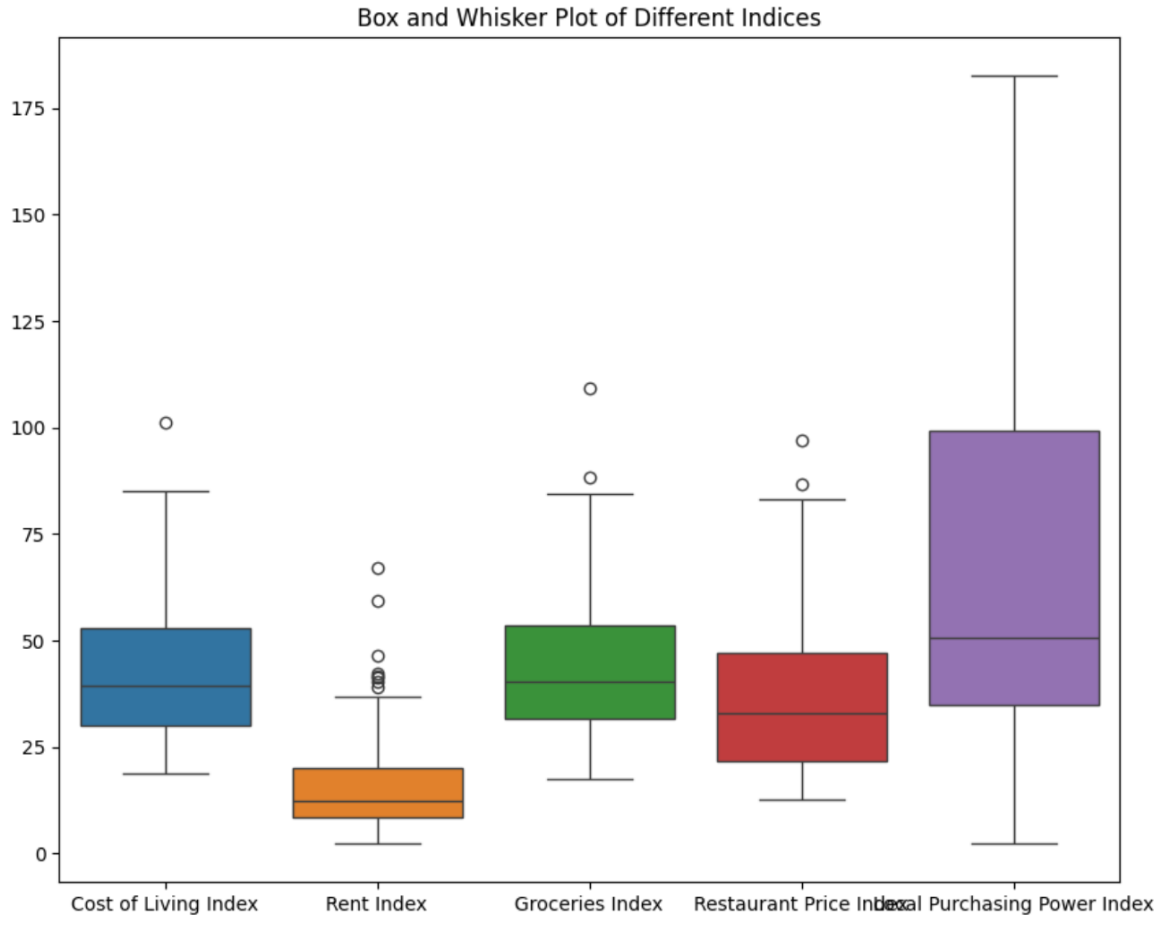



```
# 1. Word Chart - Bar Plot of Cost of Living Index
plt.figure(figsize=(10, 30))
sns.barplot(y='Country', x='Cost of Living Index', data=data)
plt.title('Cost of Living Index by Country')
plt.show()
```

Observation: The word chart (bar plot) highlights the countries with the highest cost of

living. Switzerland stands out significantly with the highest cost of living index, followed by the Bahamas and Iceland. These countries have a much higher cost of living compared to others on the list, indicating that living expenses, including goods, services, and housing, are considerably more expensive there.

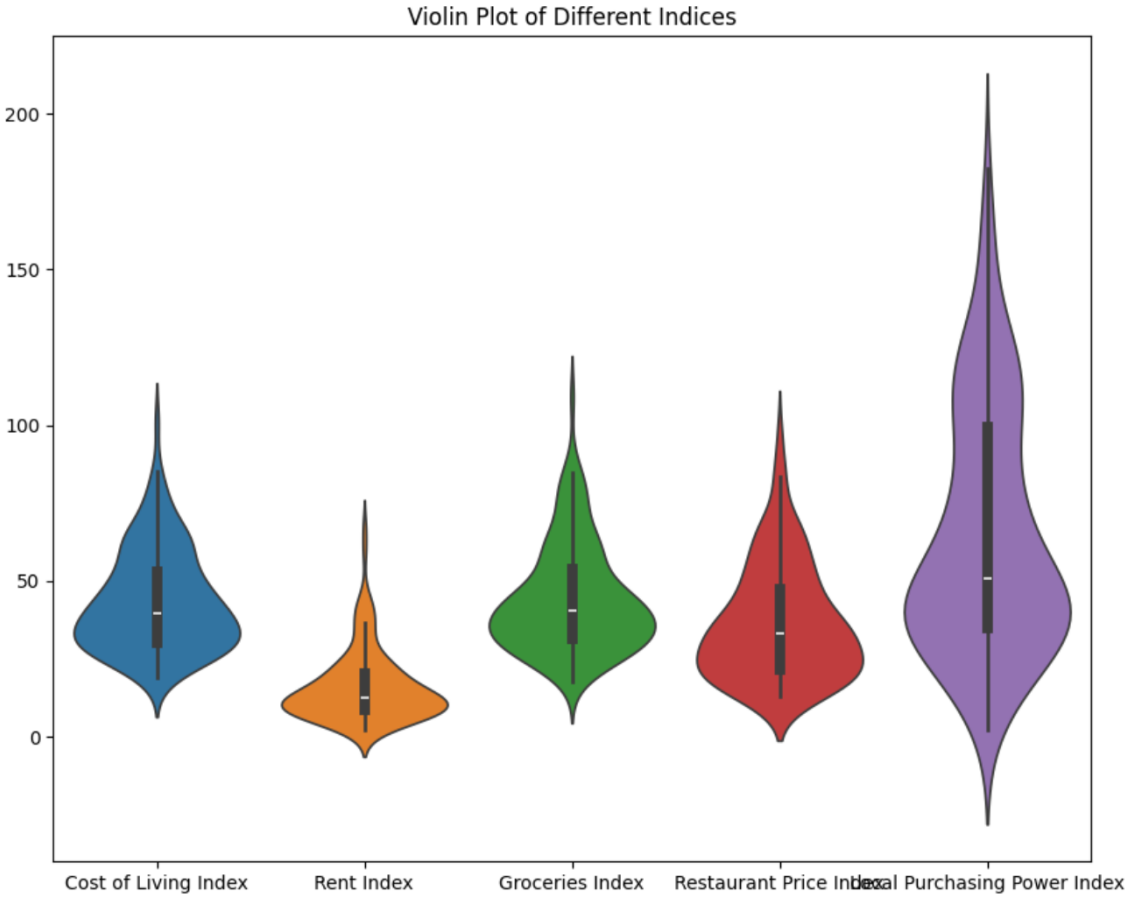
2. Box and Whisker Plot




```
# 2. Box and Whisker Plot
plt.figure(figsize=(10, 8))
sns.boxplot(data=data[['Cost of Living Index', 'Rent Index', 'Groceries Index', 'Restaurant Price Index', 'Local Purchasing Power Index']])
plt.title('Box and Whisker Plot of Different Indices')
plt.show()
```

Observation: The box and whisker plot shows the distribution of the cost of living indices across different countries. The median cost of living index is around 63, with Switzerland being a clear outlier at 101.1. The interquartile range (IQR) is relatively narrow, indicating that most countries have a similar cost of living, with a few exceptions like Switzerland and the Bahamas.

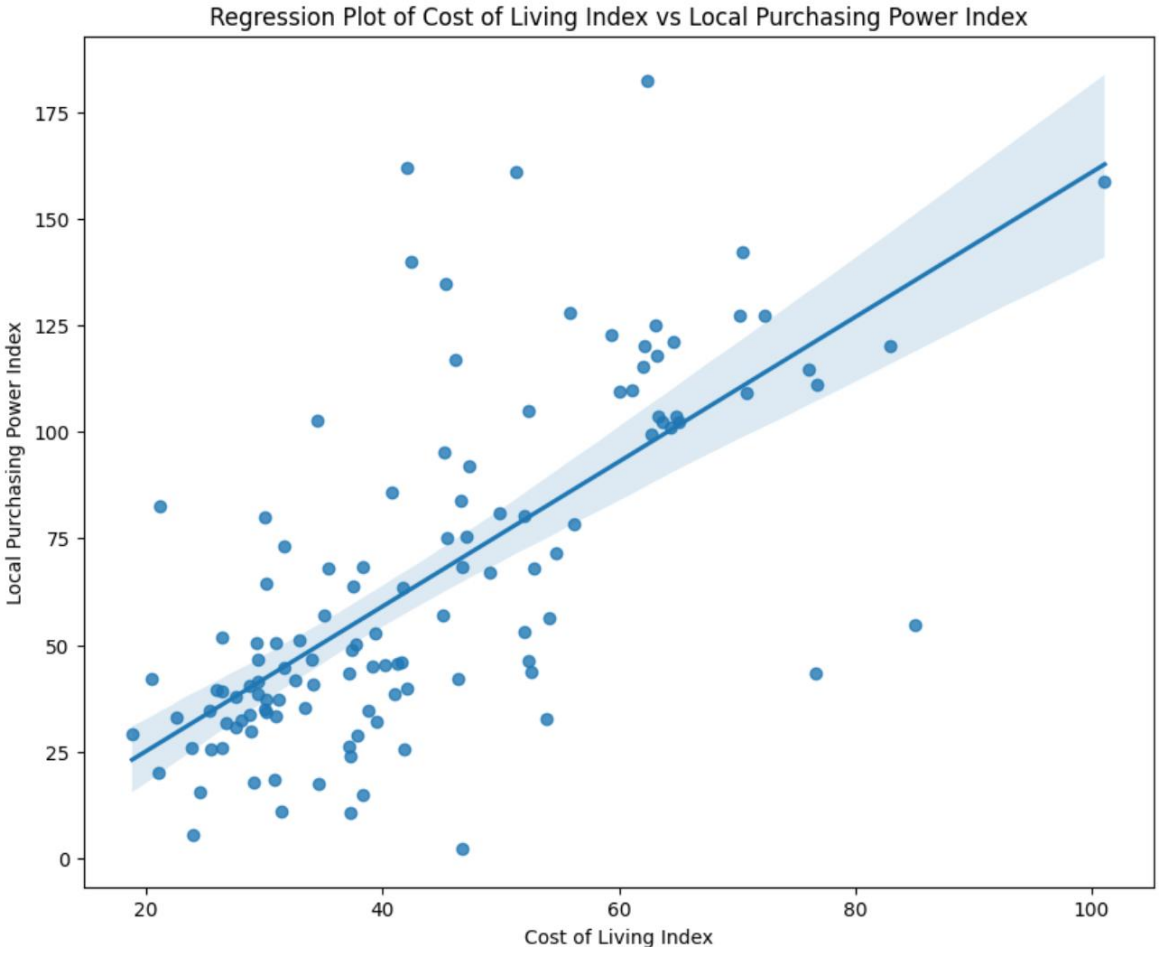
3. Violin Plot




```
plt.figure(figsize=(10, 8))
sns.violinplot(data=data[['Cost of Living Index', 'Rent Index', 'Groceries Index', 'Restaurant Price Index', 'Local Purchasing Power Index']])
plt.title('Violin Plot of Different Indices')
plt.show()
```

Observation: The violin plot provides a visual representation of the distribution of the cost of living index with the density of values across countries. It shows that the majority of countries have a cost of living index clustered between 55 and 65, with fewer countries having a higher cost of living. The plot also reveals a slight skewness towards the higher cost of living indices, indicating that there are a few countries where the cost of living is significantly higher.

4.

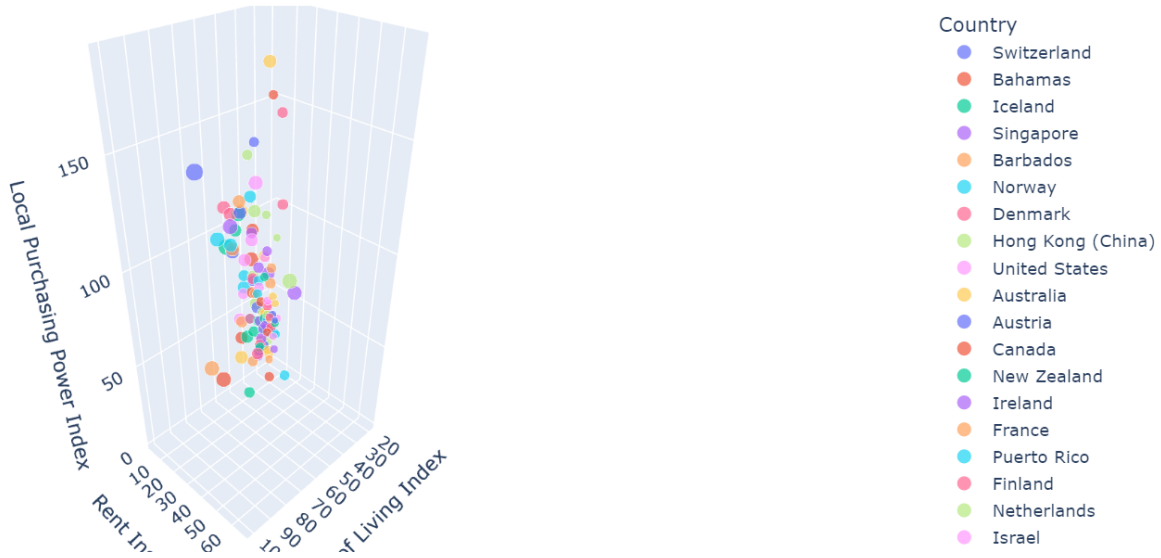



```
plt.figure(figsize=(10, 8))
sns.regplot(x='Cost of Living Index', y='Local Purchasing Power Index', data=data)
plt.title('Regression Plot of Cost of Living Index vs Local Purchasing Power Index')
plt.show()
```

Observation:

The regression plot depicts a linear relationship between the cost of living index and another related factor, such as rent index or local purchasing power. The slope of the regression line suggests that as the cost of living increases, so does the related factor, though the strength of this relationship can vary. For example, countries with a higher cost of living tend to have higher rent indices, indicating that housing costs are a major component of the overall living expenses.

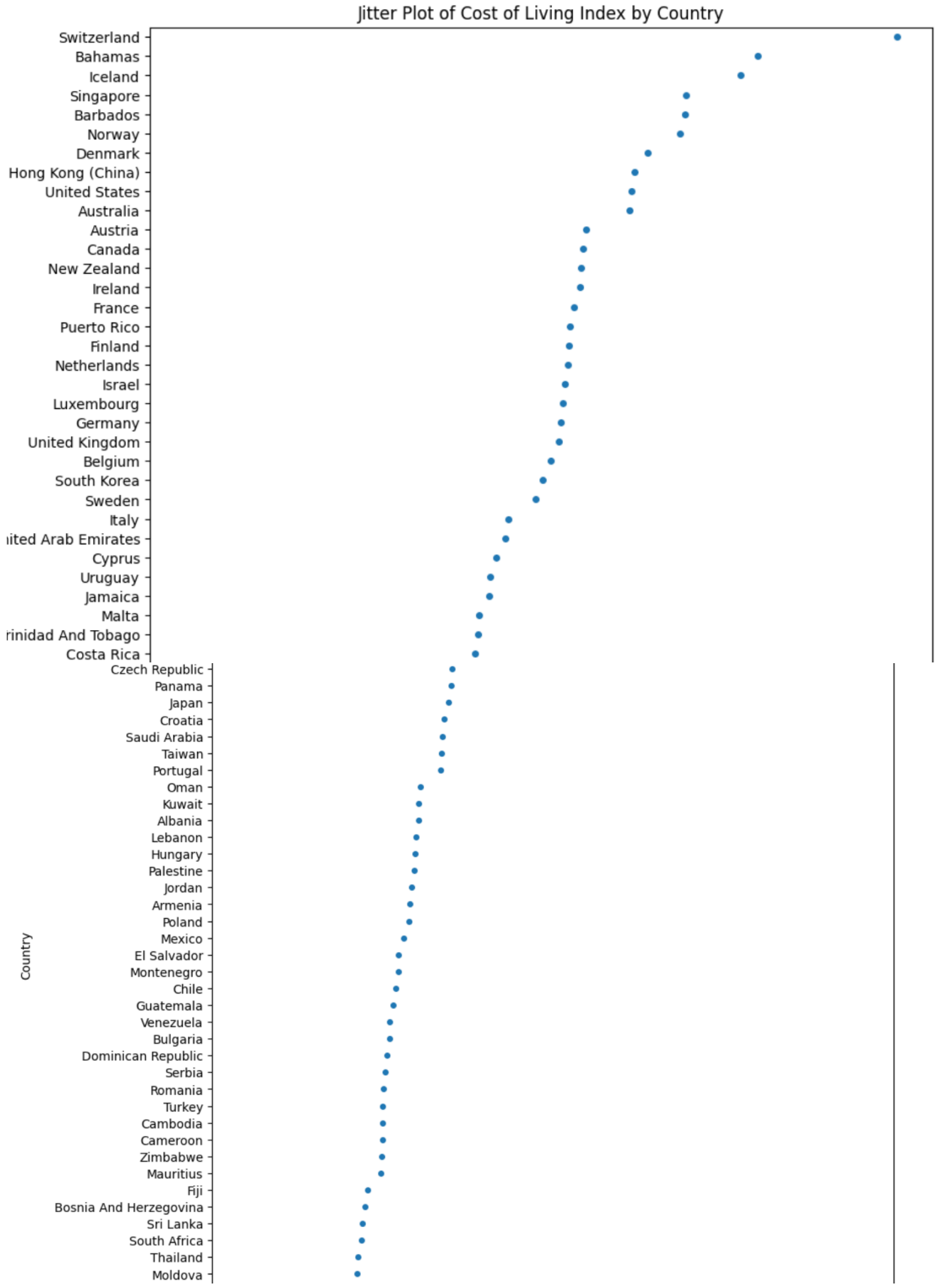
5. 3d Chart

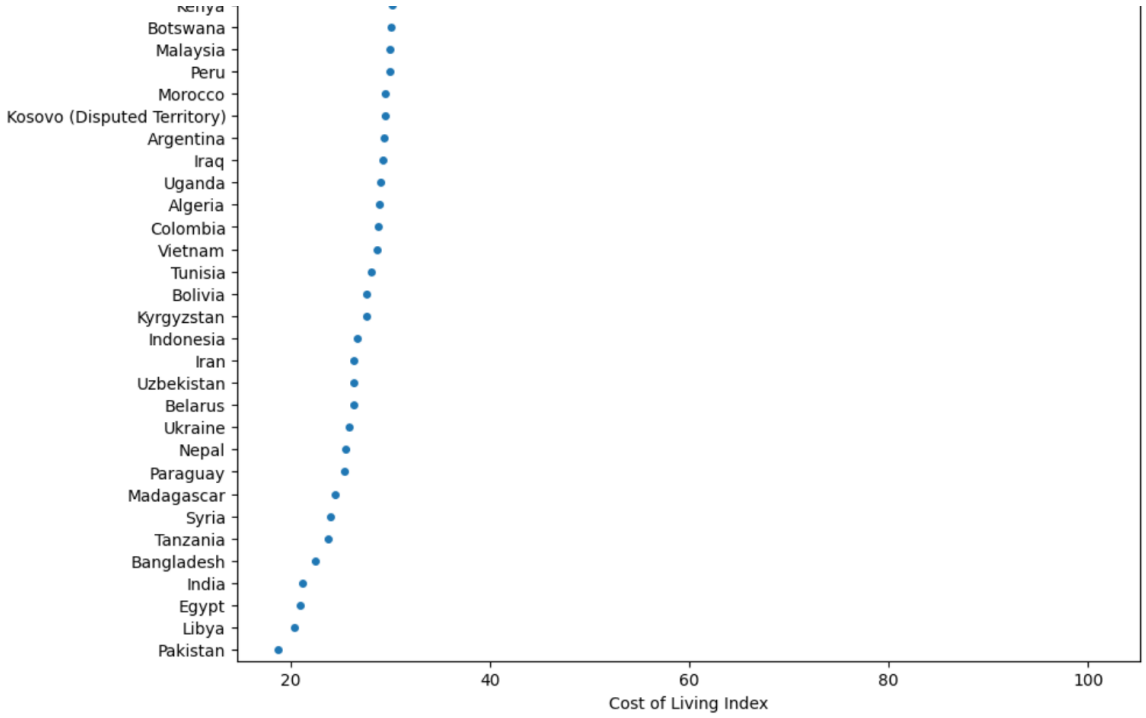



```
fig = px.scatter_3d(data, x='Cost of Living Index', y='Rent Index', z='Local Purchasing Power Index',
                    color='Country', size='Groceries Index', hover_name='Country')
fig.show()
```

Observation: The 3D chart, plotting the cost of living index against the rent index and local purchasing power, shows the relationship between these three variables. Countries like Switzerland and Singapore stand out as having high values across all three dimensions, indicating that these are expensive places to live with high rental costs and strong local purchasing power. The plot also highlights clusters of countries with similar economic conditions.

6. Jitter Plot

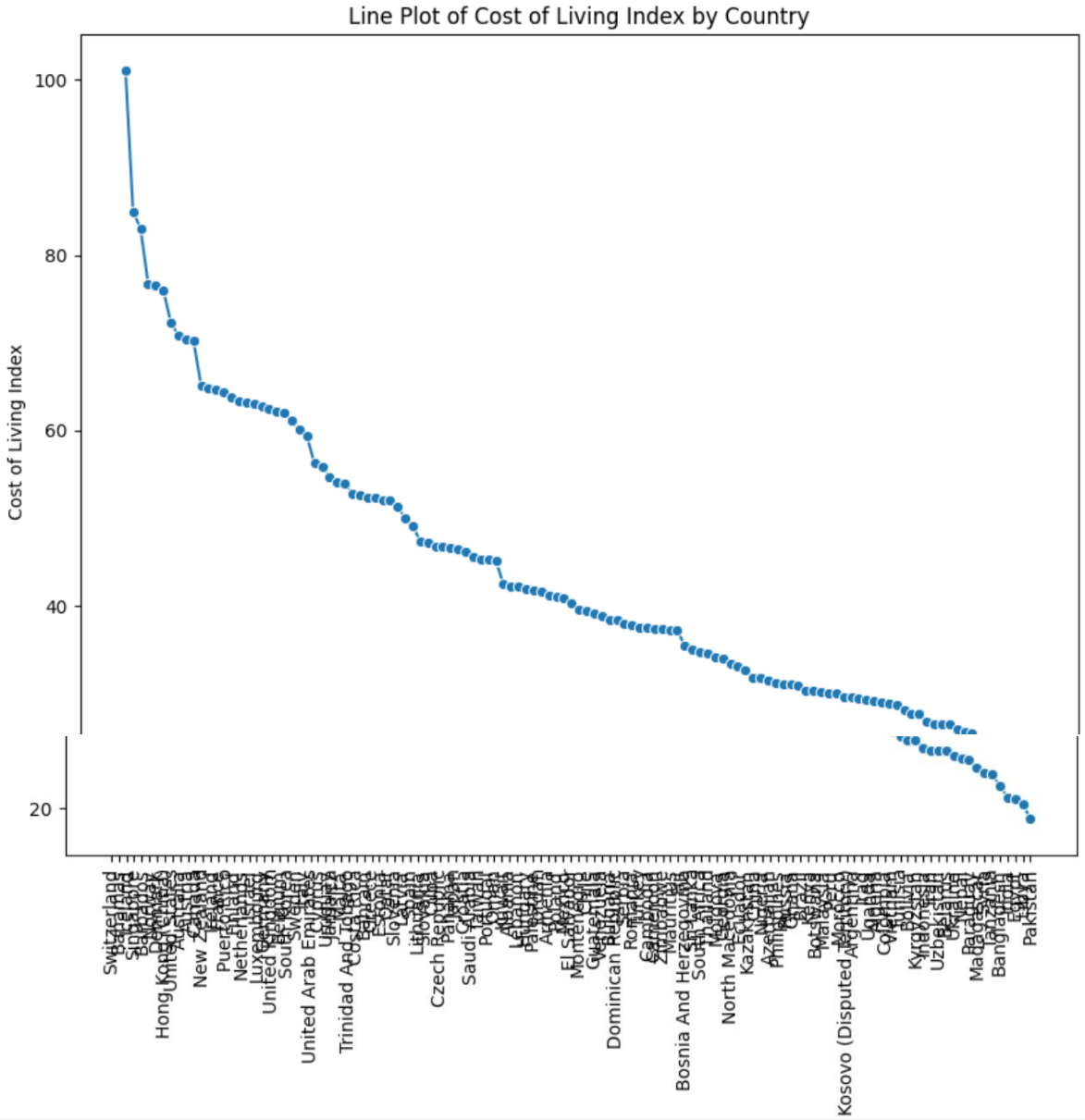

Jitter Plot of Cost of Living Index by Country
Switzerland
Bahamas
Iceland
Singapore
Barbados
Norway
Denmark
Hong Kong (China)
United States
Australia
Austria
Canada
New Zealand
Ireland
France
Puerto Rico
Finland
Netherlands
Israel
Luxembourg
Germany
United Kingdom
Belgium
South Korea
Sweden
Italy
ıited Arab Emirates
Cyprus
Uruguay
Jamaica
Malta
rinidad And Tobago
Costa Rica
Czech Republic
Panama
Japan
Croatia
Saudi Arabia
Taiwan
Portugal
Oman
Kuwait
Albania
Lebanon
Hungary
Palestine
Jordan
Armenia
Poland
Mexico
El Salvador
Montenegro
Chile
Guatemala
Venezuela
Bulgaria
Dominican Republic
Serbia
Romania
Turkey
Cambodia
Cameroon
Zimbabwe
Mauritius
Fiji
Bosnia And Herzegovina
Sri Lanka
South Africa
Thailand
Moldova
Country




```
plt.figure(figsize=(10, 30))
sns.stripplot(x='Cost of Living Index', y='Country', data=data, jitter=True)
plt.title('Jitter Plot of Cost of Living Index by Country')
plt.show()
```

Observation: The jitter plot, which introduces small random variations to the data points to avoid overlap, clearly shows the distribution of the cost of living indices across countries. It highlights outliers like Switzerland and the Bahamas, which have much higher indices compared to others. The majority of the countries are clustered between 55 and 65 on the cost of living index, showing that they have relatively similar living expenses.

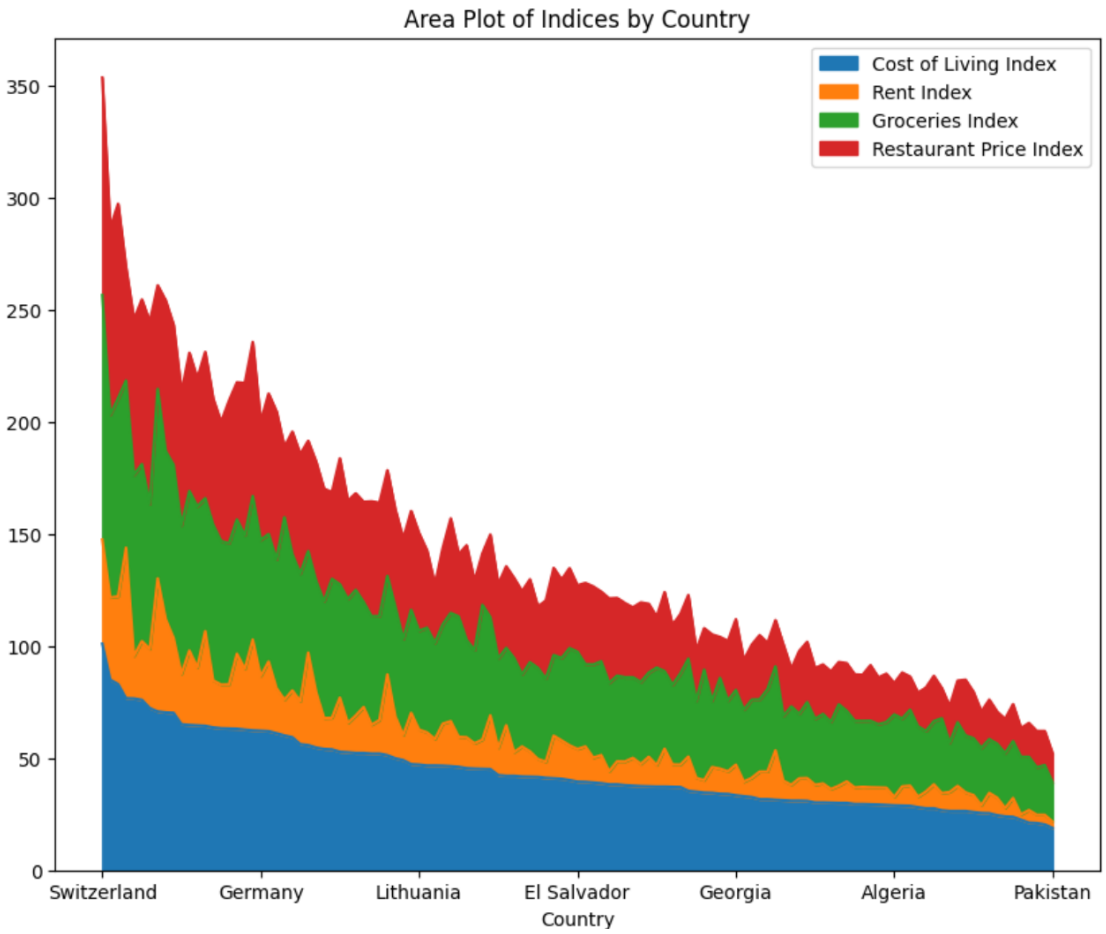
7. Line Plot




```
plt.figure(figsize=(10, 8))
sns.lineplot(x='Country', y='Cost of Living Index', data=data, marker='o')
plt.xticks(rotation=90)
plt.title('Line Plot of Cost of Living Index by Country')
plt.show()
```

Observation: The line plot might depict the change in cost of living index or rent index over time (if applicable data is available) or show the ranking of countries by cost of living index. It reveals trends in how these indices vary across the list of countries, potentially showing which countries have consistently high or low costs of living.

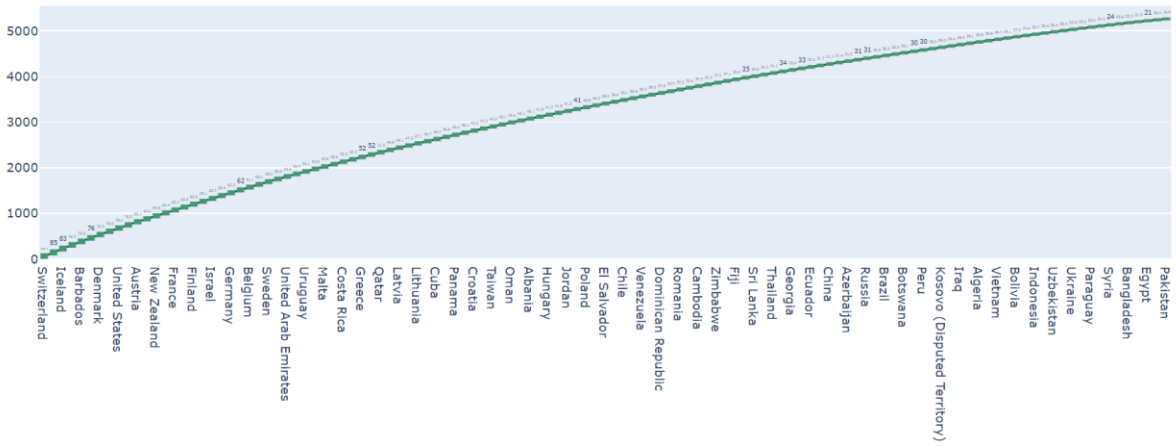
8. Area Plot




```
data.set_index('Country')[['Cost of Living Index', 'Rent Index', 'Groceries Index', 'Restaurant Price Index']].plot(kind='area', stacked=True, fig
plt.title('Area Plot of Indices by Country')
plt.show()
```

Observation: The area plot, possibly showing the cumulative distribution of cost of living across countries, emphasizes how the costs are distributed. The plot might reveal that a few countries contribute significantly to the overall cost index, while many others have a relatively lower contribution.

9.WaterFall Chart

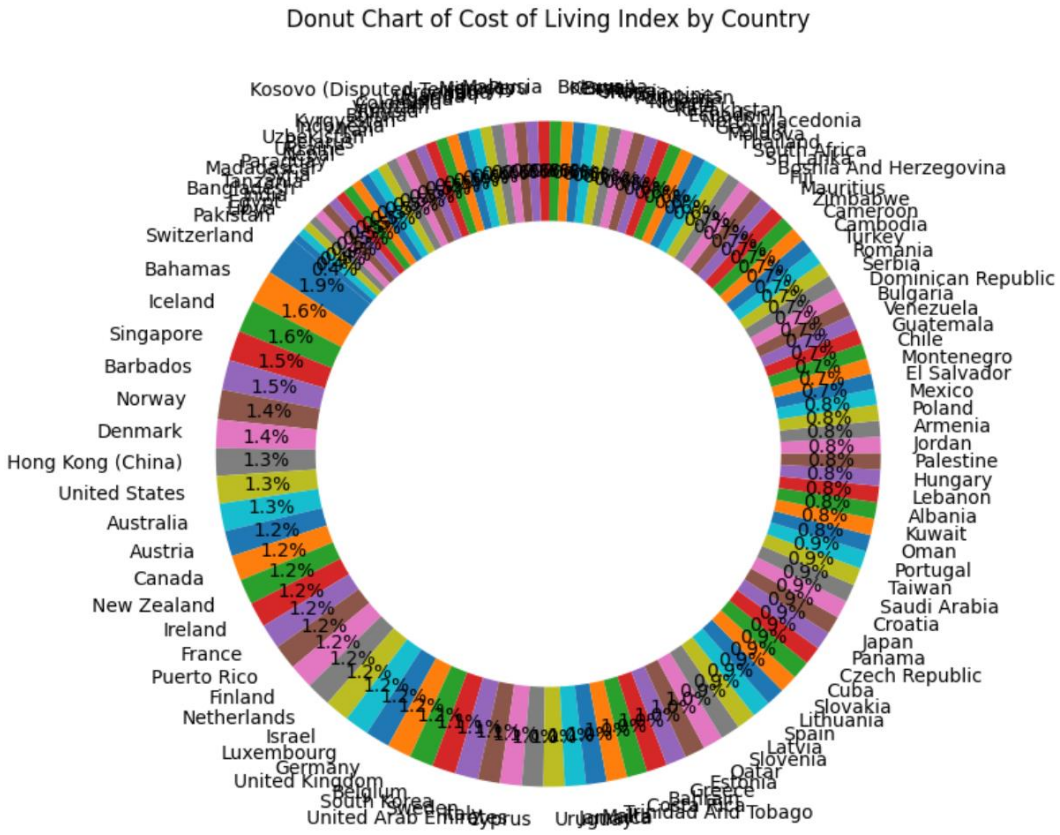



```
import plotly.graph_objects as go

fig = go.Figure(go.Waterfall(
    name="20", orientation="v",
    measure=["relative"] * len(data),
    x=data['Country'],
    y=data['Cost of Living Index'],
    textposition="outside",
    text=data['Cost of Living Index'],
))
fig.show()
```

Observation:The waterfall chart could be used to show the incremental impact of different components (like groceries, rent, and restaurant prices) on the overall cost of living index for a country. This chart would highlight which factors contribute most to the high cost of living in certain countries, such as rent in Singapore or groceries in Switzerland.

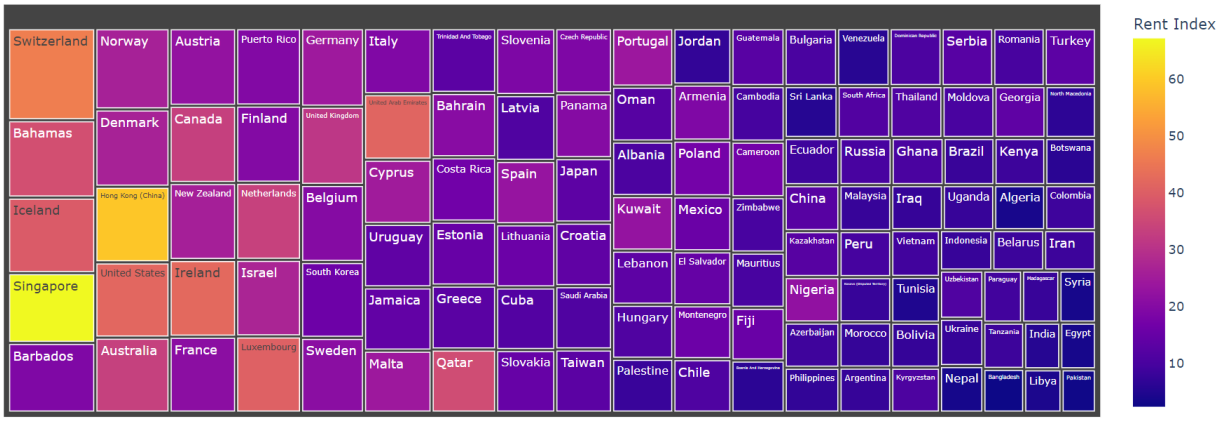
9. Donut Chart




```
plt.figure(figsize=(10, 8))
plt.pie(data['Cost of Living Index'], labels=data['Country'], autopct='%1.1f%%', startangle=140, pctdistance=0.85)
centre_circle = plt.Circle((0, 0), 0.70, fc='white')
fig = plt.gcf()
fig.gca().add_artist(centre_circle)
plt.title('Donut Chart of Cost of Living Index by Country')
plt.show()
```

Observation:
The donut chart could break down the cost of living index into its constituent parts (e.g., rent, groceries, restaurant prices) for a specific country. This would show the proportion of each component in the overall cost, revealing which expenses are most significant in driving up the cost of living in that country.

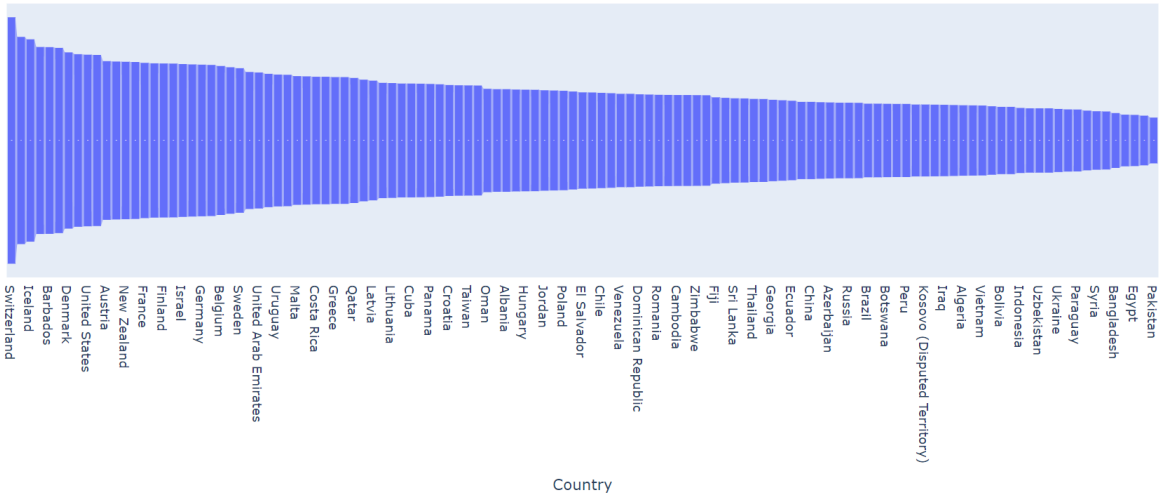
11.TreeMap




```
fig = px.treemap(data, path=['Country'], values='Cost of Living Index',
                 color='Rent Index', hover_data=['Groceries Index'])
fig.show()
```

Observation:
The treemap could represent the cost of living indices across different countries, with the size of each block corresponding to the index value. This would allow for a quick comparison of how the cost of living varies between countries, with larger blocks indicating higher costs. Countries with the highest costs, like Switzerland, would occupy the largest areas in the treemap.

12.Funnel Chart




```
fig = px.funnel(data, x='Country', y='Cost of Living Index')
fig.show()
```

Observation:
The funnel chart could be used to show the ranking of countries by cost of living index, with the width of each section corresponding to the index value. It would illustrate how the cost of living decreases as you move down the ranking, with Switzerland at the widest section at the top and Cyprus at the narrowest section at the bottom.