In [1]:
```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns

#Descriptive Statistical Analysis
df = pd.read_csv('C:/Users/Jalay/OneDrive/Desktop/indexes by year.csv')
descriptive_stats = df.describe()
print(descriptive_stats)
```

```
              Rank  Quality of Life Index  Purchasing Power Index  \
count  361.000000             361.000000              361.000000
mean    37.091413             136.973047               66.129917
std     21.694596              35.177825               29.997066
min      1.000000               0.000000                3.350000
25%     19.000000             108.470000               39.850000
50%     37.000000             139.140000               61.160000
75%     55.000000             165.930000               91.020000
max     83.000000             198.570000              138.290000

       Safety Index  Health Care Index  Cost of Living Index  \
count    361.000000         361.000000            361.000000
mean      61.020970          65.594737             55.068449
std       13.629872           9.868366             21.116408
min       14.720000          36.900000             20.400000
25%       53.210000          58.010000             38.260000
50%       60.330000          66.590000             50.180000
75%       71.430000          73.300000             71.790000
max       88.140000          86.710000            131.750000

       Property Price to Income Ratio  Traffic Commute Time Index  \
count                      361.000000                  361.000000
mean                        13.211828                   35.910360
std                         12.086857                    7.528254
min                          2.570000                   19.740000
25%                          8.520000                   30.170000
50%                         10.900000                   34.760000
75%                         14.210000                   40.330000
max                        202.070000                   61.680000

       Pollution Index  Climate Index         Year
count       361.000000     361.000000   361.000000
mean         51.990360      72.669806  2019.144044
std          20.666634      23.599924     1.426446
min          11.550000     -79.430000  2017.000000
25%          33.590000      65.320000  2018.000000
50%          54.530000      77.560000  2019.000000
75%          67.080000      88.740000  2020.000000
max          89.350000      99.790000  2021.000000
```

In [2]:
```
#Correlation Analysis
correlation_matrix = df.corr()
print(correlation_matrix)
```

	Rank	Quality of Life Index
Rank	1.000000	-0.966842
Quality of Life Index	-0.966842	1.000000
Purchasing Power Index	-0.799841	0.794155
Safety Index	-0.518167	0.544358
Health Care Index	-0.558367	0.562061
Cost of Living Index	-0.692935	0.675878
Property Price to Income Ratio	0.391968	-0.492405
Traffic Commute Time Index	0.655114	-0.662865
Pollution Index	0.868033	-0.852932
Climate Index	-0.130091	0.222355
Year	0.170300	-0.079711

	Purchasing Power Index	Safety Index
Rank	-0.799841	-0.518167
Quality of Life Index	0.794155	0.544358
Purchasing Power Index	1.000000	0.431001
Safety Index	0.431001	1.000000
Health Care Index	0.556022	0.364873
Cost of Living Index	0.737592	0.402920
Property Price to Income Ratio	-0.348384	-0.170472
Traffic Commute Time Index	-0.398268	-0.488256
Pollution Index	-0.609125	-0.398633
Climate Index	-0.118776	-0.179221
Year	-0.276772	-0.027643

	Health Care Index	Cost of Living Index
Rank	-0.558367	-0.692935
Quality of Life Index	0.562061	0.675878
Purchasing Power Index	0.556022	0.737592
Safety Index	0.364873	0.402920
Health Care Index	1.000000	0.561195
Cost of Living Index	0.561195	1.000000
Property Price to Income Ratio	-0.168530	-0.150445
Traffic Commute Time Index	-0.174141	-0.357455
Pollution Index	-0.468267	-0.646175
Climate Index	-0.013718	0.066195
Year	-0.066581	-0.058949

	Property Price to Income Ratio
Rank	0.391968
Quality of Life Index	-0.492405
Purchasing Power Index	-0.348384
Safety Index	-0.170472
Health Care Index	-0.168530
Cost of Living Index	-0.150445
Property Price to Income Ratio	1.000000
Traffic Commute Time Index	0.293700
Pollution Index	0.267345
Climate Index	0.012705
Year	-0.014585

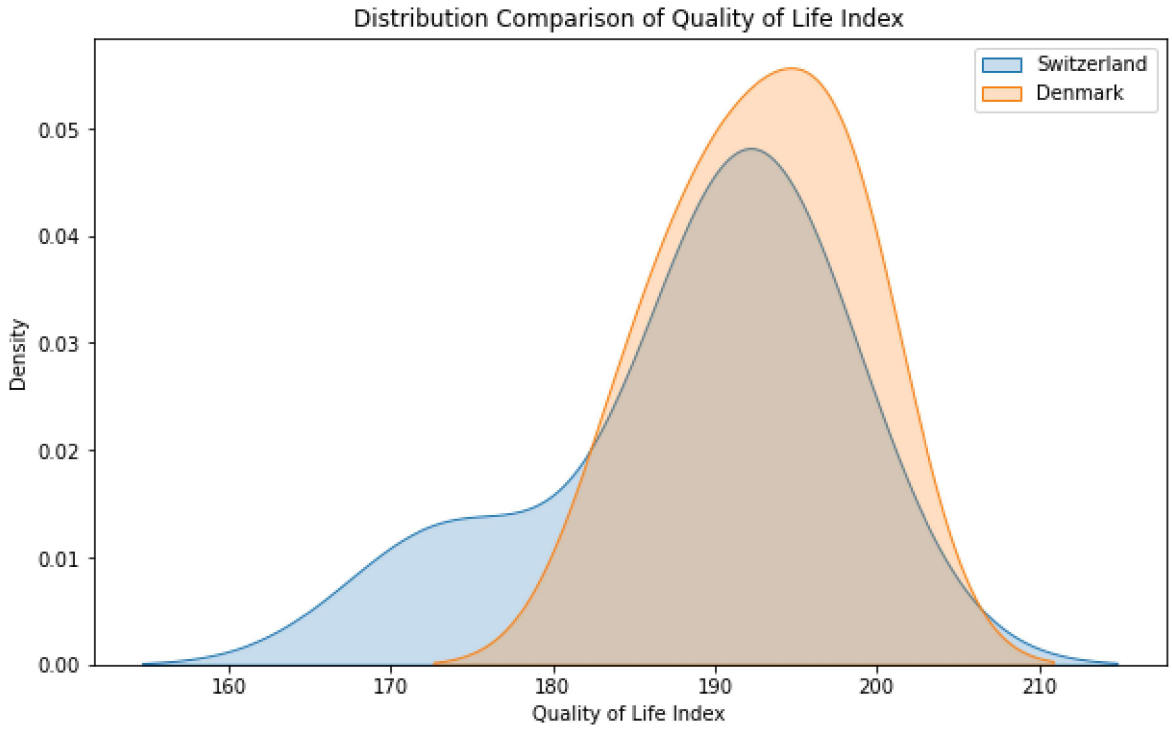
	Traffic Commute Time Index	Pollution Index
Rank	0.655114	0.868033
Quality of Life Index	-0.662865	-0.852932
Purchasing Power Index	-0.398268	-0.609125

```
Safety Index                                -0.488256        -0.398633
Health Care Index                           -0.174141        -0.468267
Cost of Living Index                        -0.357455        -0.646175
Property Price to Income Ratio               0.293700         0.267345
Traffic Commute Time Index                   1.000000         0.565534
Pollution Index                              0.565534         1.000000
Climate Index                               -0.058348        -0.088464
Year                                        -0.011154         0.032210

                              Climate Index      Year
Rank                              -0.130091  0.170300
Quality of Life Index              0.222355 -0.079711
Purchasing Power Index            -0.118776 -0.276772
Safety Index                      -0.179221 -0.027643
Health Care Index                 -0.013718 -0.066581
Cost of Living Index               0.066195 -0.058949
Property Price to Income Ratio     0.012705 -0.014585
Traffic Commute Time Index        -0.058348 -0.011154
Pollution Index                   -0.088464  0.032210
Climate Index                      1.000000  0.305165
Year                               0.305165  1.000000
```

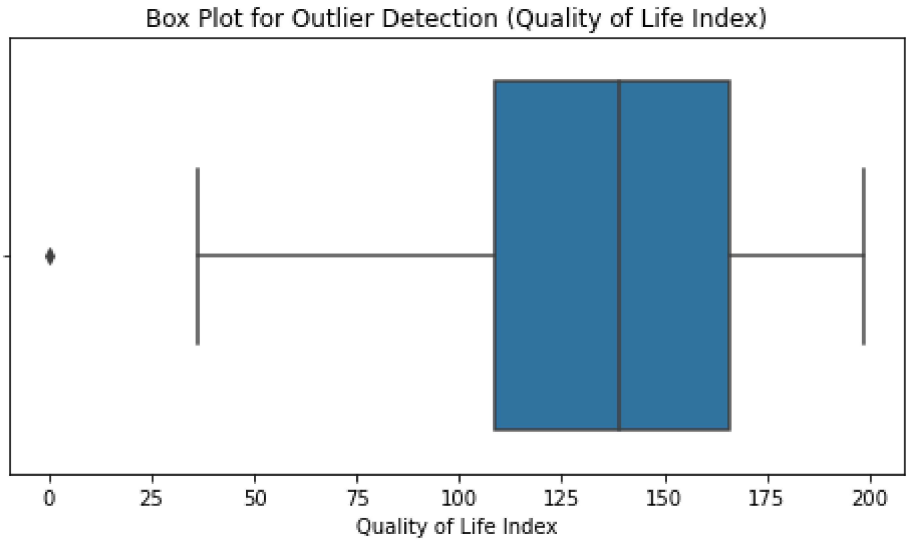
In [27]:
```
#Grouped Analysis
grouped_data = df.groupby('Country').mean().head(5)
print(grouped_data)
```

```
            Rank  Quality of Life Index  Purchasing Power Index  Safety I
ndex  \
Country
Argentina   46.6                123.948                  51.066        3
7.590
Australia    6.0                184.820                 110.874        5
7.694
Austria      4.4                187.302                  90.332        7
8.002
Azerbaijan  66.0                102.880                  25.770        6
8.340
Bangladesh  80.0                 67.650                  28.640        3
5.920

            Health Care Index  Cost of Living Index  \
Country
Argentina              70.366                42.044
Australia              76.386                78.944
Austria                79.206                74.130
```

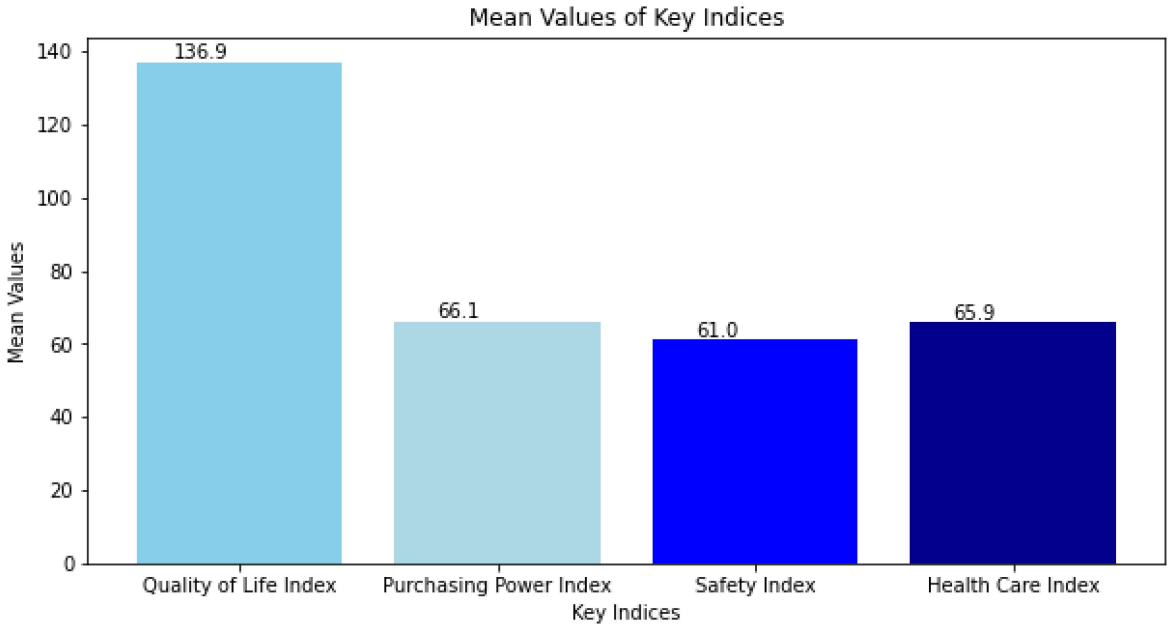
In [22]:
```
#Distribution Comparison of QLI
plt.figure(figsize=(10, 6))
sns.kdeplot(df[df['Country'] == 'Switzerland']['Quality of Life Index'], label
sns.kdeplot(df[df['Country'] == 'Denmark']['Quality of Life Index'], label='De
plt.title('Distribution Comparison of Quality of Life Index')
plt.xlabel('Quality of Life Index')
plt.ylabel('Density')
plt.legend()
plt.show()
```



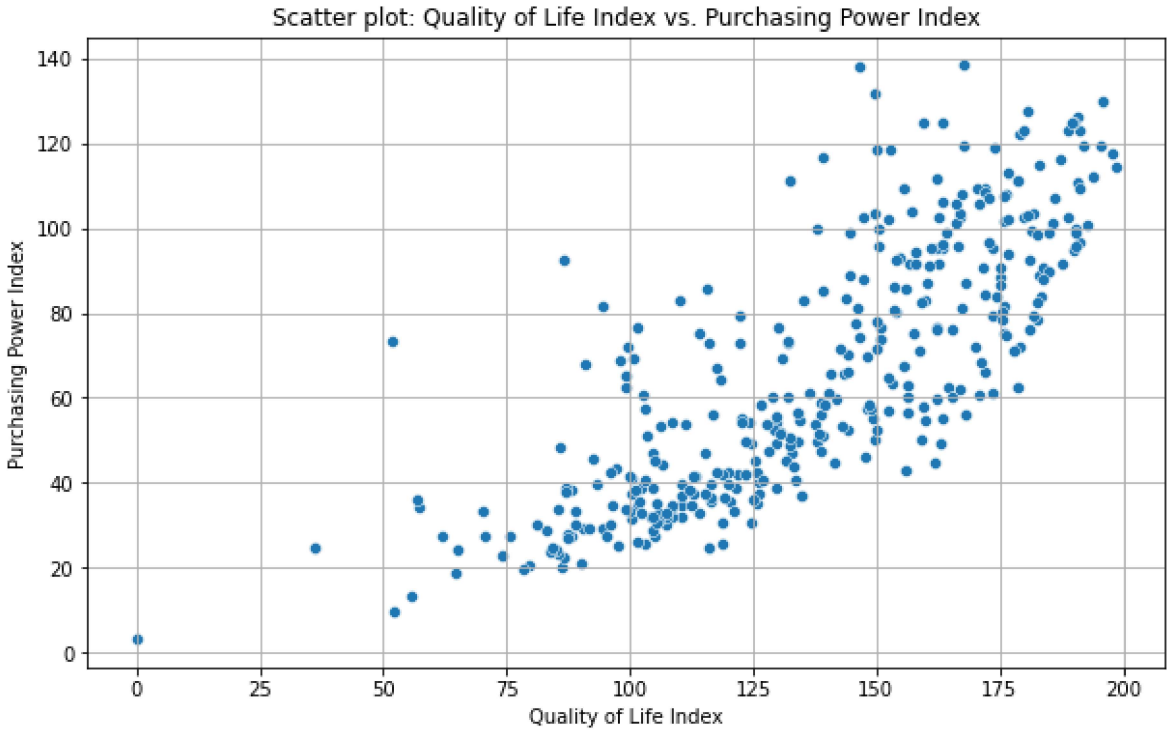
In [15]:
```
#Outliers Detection of QLI
plt.figure(figsize=(8, 4))
sns.boxplot(x=df['Quality of Life Index'])
plt.title('Box Plot for Outlier Detection (Quality of Life Index)')
plt.show()
```



In [24]:
```
#Mean of key indices
key_indices = ['Quality of Life Index', 'Purchasing Power Index', 'Safety Inde
mean_values = [136.9 , 66.1, 61.0, 65.9]
colors = ['skyblue', 'lightblue', 'blue', 'darkblue']
plt.figure(figsize=(10, 5))
bars = plt.bar(key_indices, mean_values, color=colors)
plt.title('Mean Values of Key Indices')
plt.xlabel('Key Indices')
plt.ylabel('Mean Values')
for bar, value in zip(bars, mean_values):
    plt.text(bar.get_x() + bar.get_width() / 2 - 0.15, bar.get_height() + 1,
plt.show()
```

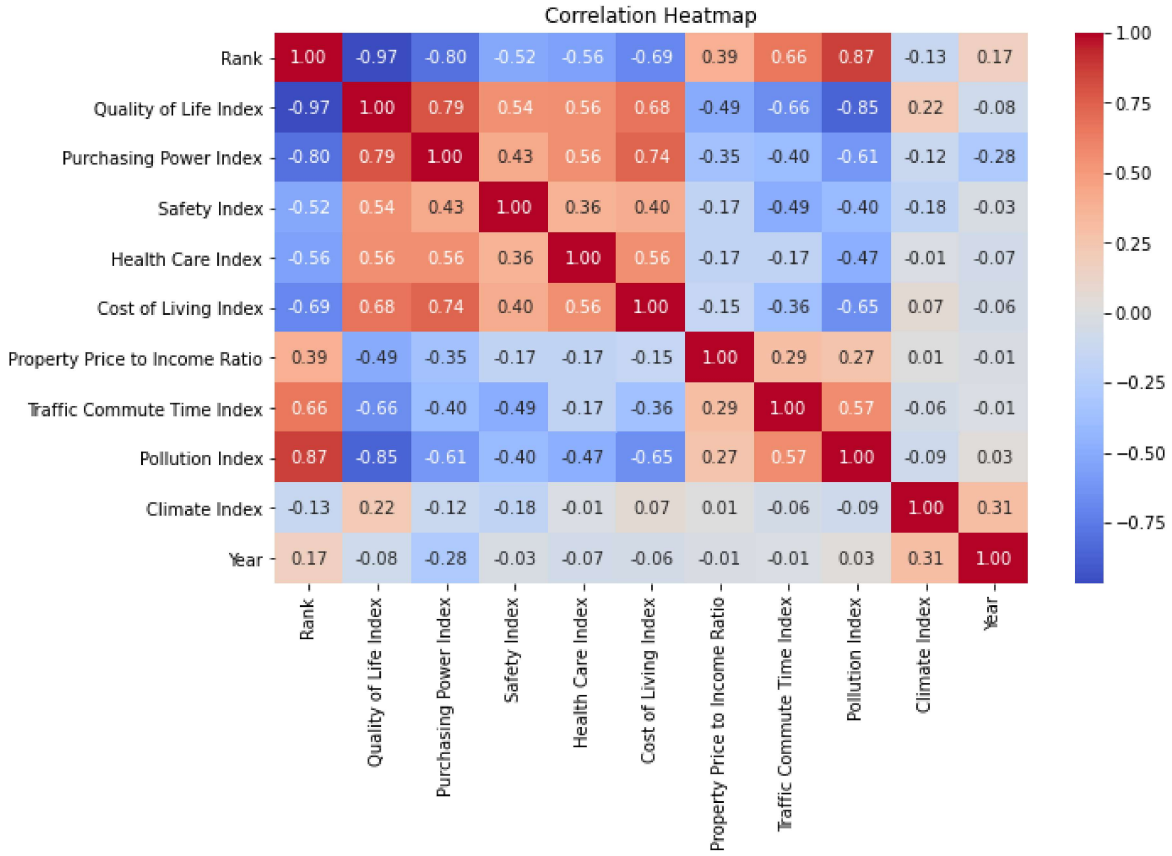


In [11]:
```
#Correlation coefficient between QLI and PPI
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Quality of Life Index', y='Purchasing Power Index', data=d
plt.title('Scatter plot: Quality of Life Index vs. Purchasing Power Index')
plt.xlabel('Quality of Life Index')
plt.ylabel('Purchasing Power Index')
plt.grid(True)
plt.show()
correlation_coefficient = df['Quality of Life Index'].corr(df['Purchasing Powe
print(f"The correlation coefficient between QLI and PPI is: {correlation_coef
```



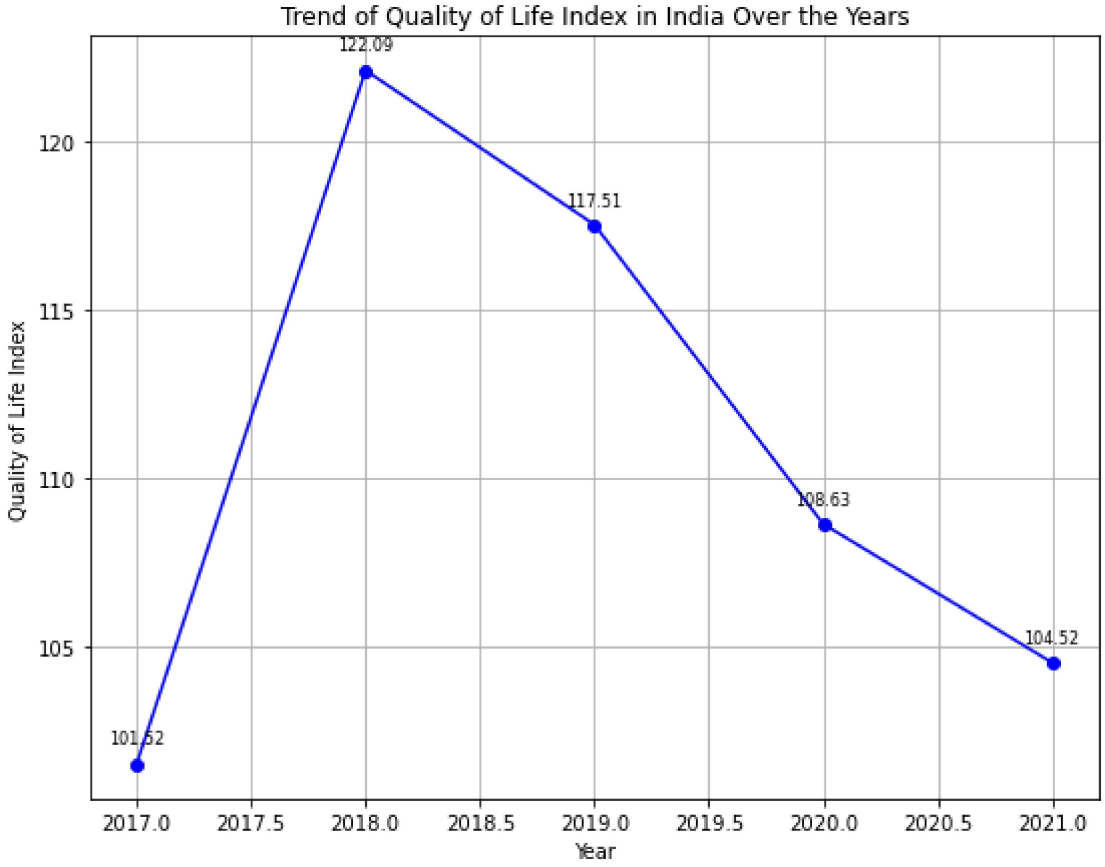

```
The correlation coefficient between QLI and PPI is: 0.79
```

In [12]:
```
#Correlation heatmap
correlation_matrix = df.corr()
plt.figure(figsize=(10, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Heatmap')
plt.show()
```

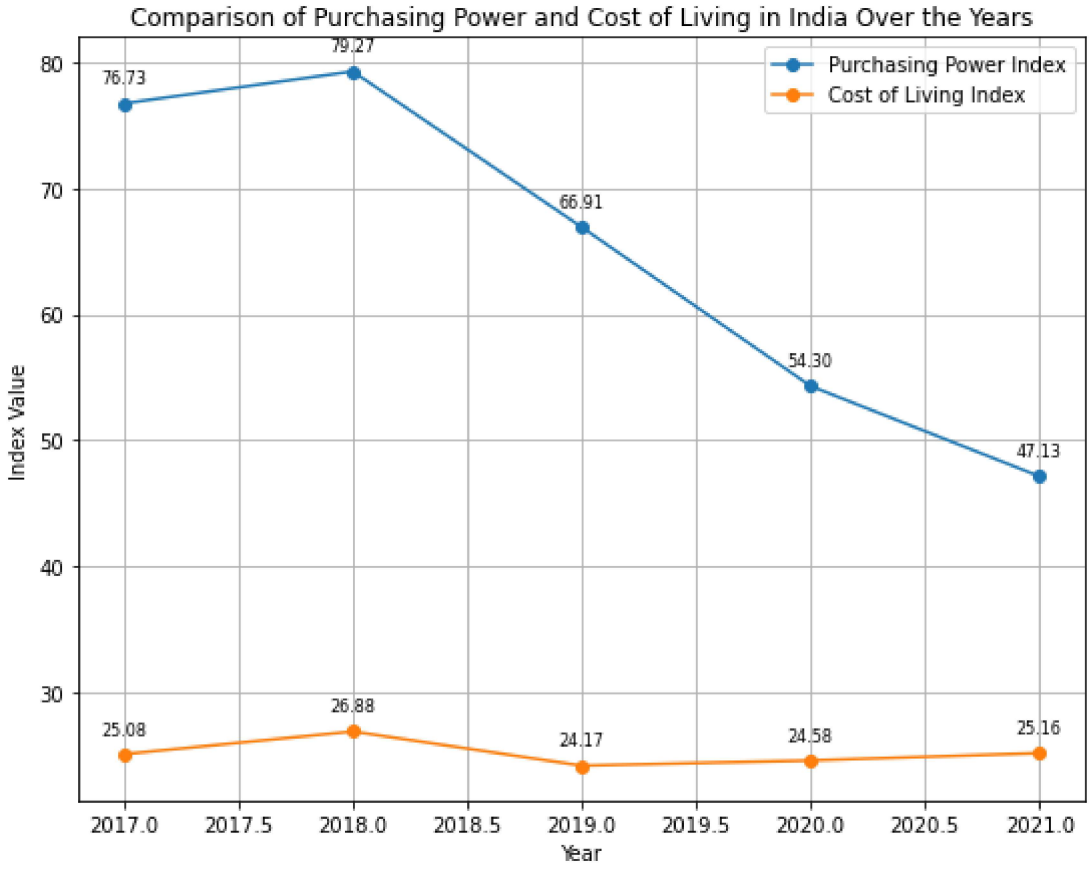


statistical analysis of india

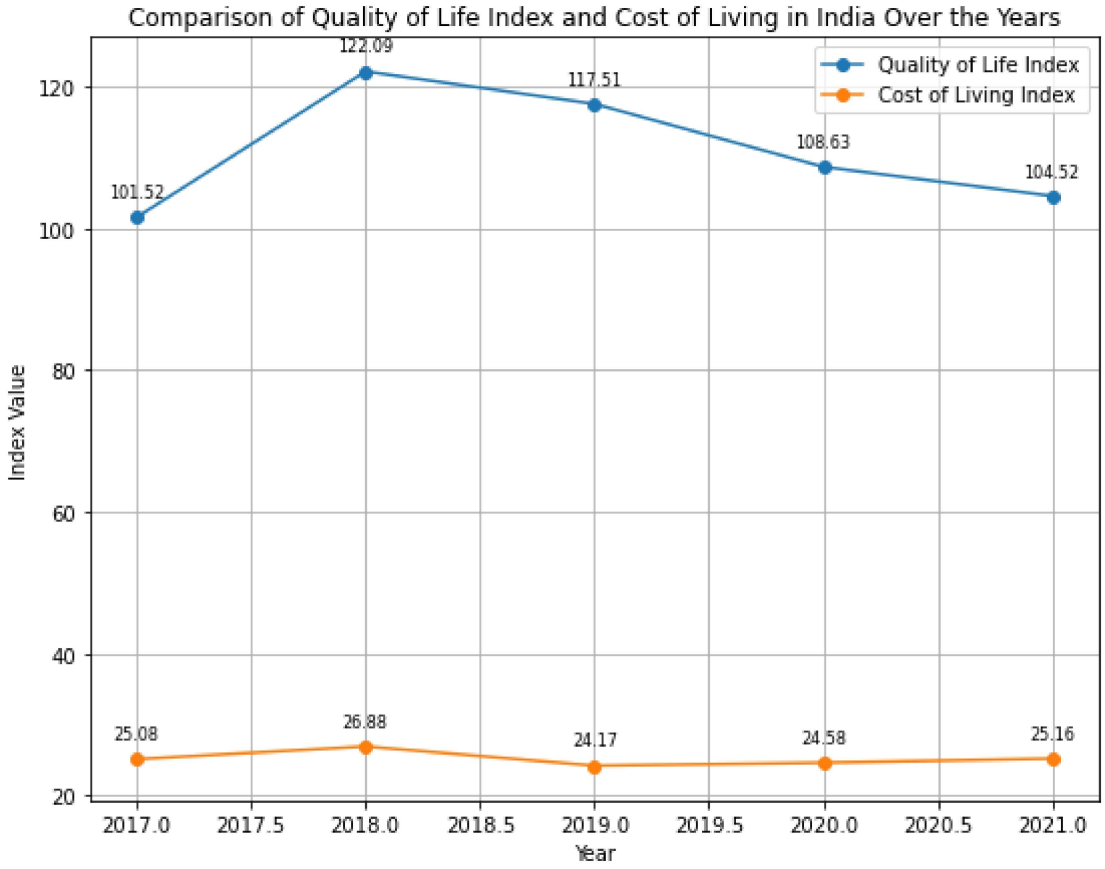
In [50]:
```
#Trend of india's QLI over the years(2017 to 2021)
india_data = df[df['Country'] == 'India']
plt.figure(figsize=(9, 7))
plt.plot(india_data['Year'], india_data['Quality of Life Index'], marker='o',
plt.title('Trend of Quality of Life Index in India Over the Years')
plt.xlabel('Year')
plt.ylabel('Quality of Life Index')
plt.grid(True)
for x, y in zip(india_data['Year'], india_data['Quality of Life Index']):
    label = f'{y:.2f}'
    plt.annotate(label, (x, y), textcoords="offset points", xytext=(0,10), ha=
plt.show()
```



In [54]:
```
india_data = df[df['Country'] == 'India']
indices_to_compare = ['Purchasing Power Index', 'Cost of Living Index']
plt.figure(figsize=(9, 7))
for index in indices_to_compare:
    plt.plot(india_data['Year'], india_data[index], marker='o', label=index)

    for x, y in zip(india_data['Year'], india_data[index]):
        label = f'{y:.2f}'
        plt.annotate(label, (x, y), textcoords="offset points", xytext=(0,10)

plt.title('Comparison of Purchasing Power and Cost of Living in India Over the
plt.xlabel('Year')
plt.ylabel('Index Value')
plt.legend()
plt.grid(True)
plt.show()
```



In [56]:
```
india_data = df[df['Country'] == 'India']
indices_to_compare = ['Quality of Life Index', 'Cost of Living Index']
plt.figure(figsize=(9, 7))
for index in indices_to_compare:
    plt.plot(india_data['Year'], india_data[index], marker='o', label=index)

    for x, y in zip(india_data['Year'], india_data[index]):
        label = f'{y:.2f}'
        plt.annotate(label, (x, y), textcoords="offset points", xytext=(0,10)

plt.title('Comparison of Quality of Life Index and Cost of Living in India Ove
plt.xlabel('Year')
plt.ylabel('Index Value')
plt.legend()
plt.grid(True)
plt.show()
```



In [ ]: