2.2 Exercise. DSC550 - Jennifer Barrera Conde

March 19, 2024

1 2.2 Exercise

1.1 DSC550-T301 Data Mining

1.2 Jennifer Barrera Conde

1.2.1 Cities with the Best Work-Life Balance 2022

Questions to be explored in the data set:

1. How does the work-life balance score vary among different cities?
2. Is there any correlation between affordability and the total score?
3. How does the Happiness, Culture & Leisure affect the total score?

Create at least three graphs that help answer these questions. Make sure your graphs are clearly readable and are labeled appropriately and professionally:

[1]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

[2]:
```
data = pd.read_csv("cities_with_best_worklife_balance_2022.csv")
```

[3]:
```
print(data.columns)
```

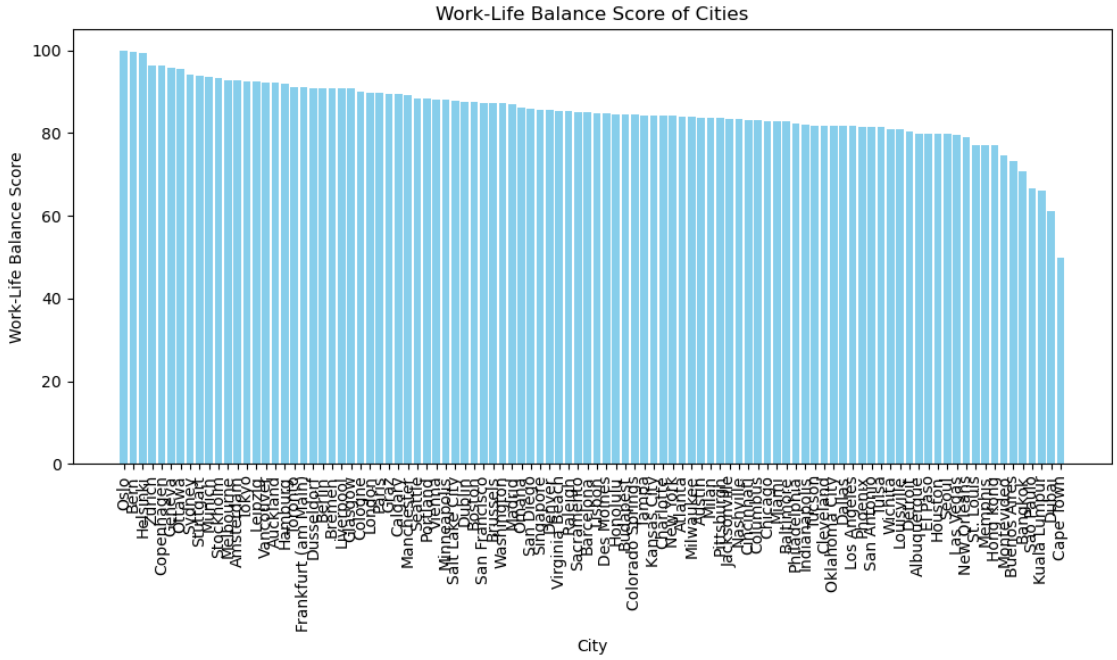
```
Index(['2022', '2021', 'City', 'Country', 'Remote Jobs',
       'Overworked Population', 'Minimum Vacations Offered (Days)',
       'Vacations Taken (Days)', 'Unemployment', 'Multiple Jobholders',
       'Inflation', 'Paid Parental Leave (Days)', 'Covid Impact',
       'Covid Support', 'Healthcare', 'Access to Mental Healthcare',
       'Inclusivity & Tolerance', 'Affordability',
       'Happiness, Culture & Leisure', 'City Safety', 'Outdoor Spaces',
       'Air Quality', 'Wellness and Fitness', 'TOTAL SCORE'],
      dtype='object')
```

The previous index helped me visualize my data column options and helped me set a path as to what I would like to achieve with this exercise.

[ ]:
```

```

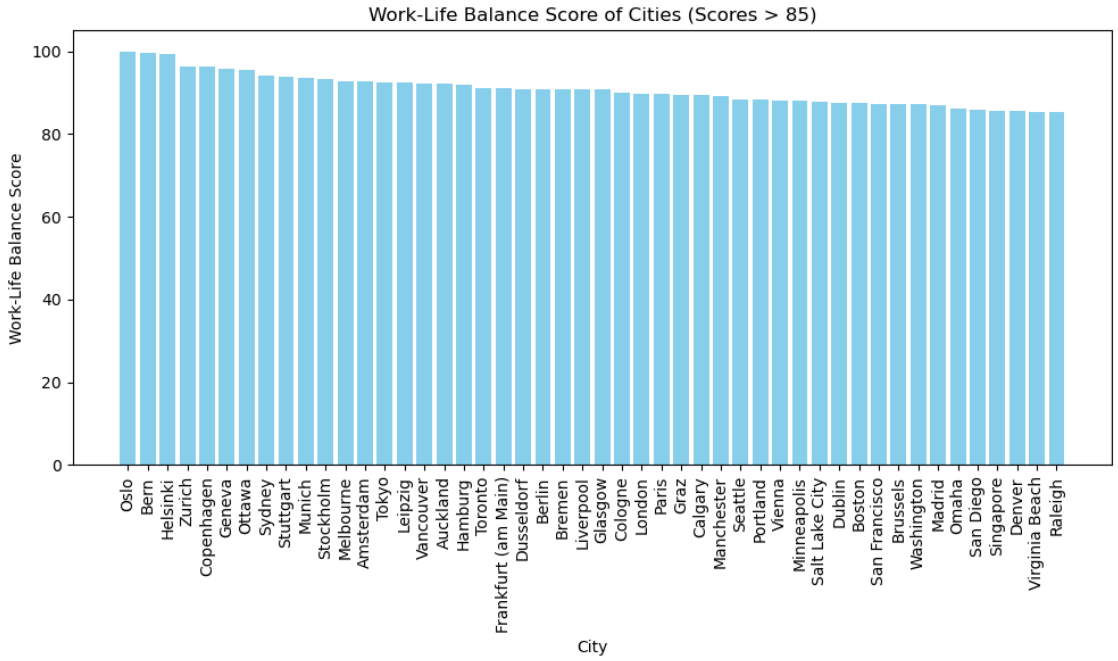
The first chart is a bar chart that displays the Total Scores of various cities. By visualizing this data, we can identify which cities offer the best and worst work-life balance.

[4]:
```
# Bar Chart - Total Scores of Cities
plt.figure(figsize=(10, 6))
plt.bar(data['City'], data['TOTAL SCORE'], color='skyblue')
plt.xticks(rotation=90)
plt.xlabel('City')
plt.ylabel('Work-Life Balance Score')
plt.title('Work-Life Balance Score of Cities')
plt.tight_layout()
plt.show()
```



There is too much data making the graph hard to read, I decided to only use Total Scores higher than 85

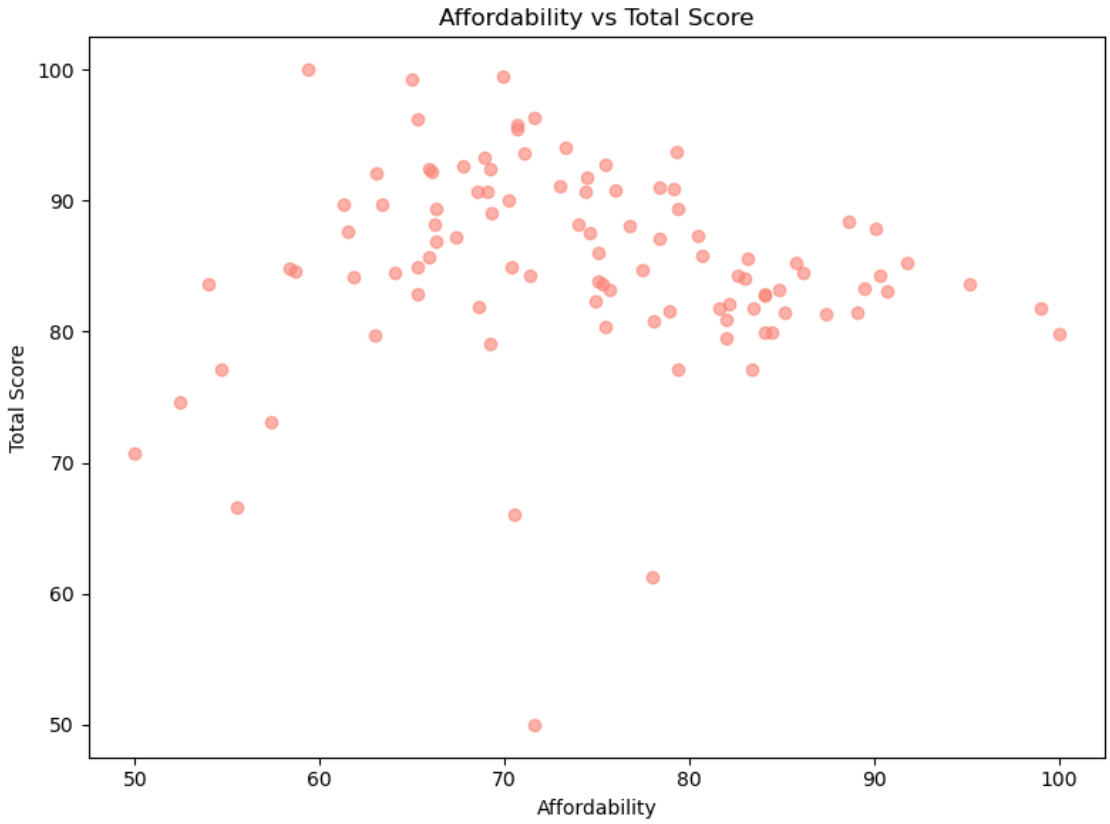
[5]:
```
high_score_data = data[data['TOTAL SCORE'] > 85]

# Bar Chart - Total scores of Cities (Scores > 85)
plt.figure(figsize=(10, 6))
plt.bar(high_score_data['City'], high_score_data['TOTAL SCORE'],␣
 ↪color='skyblue')
plt.xticks(rotation=90)
plt.xlabel('City')
plt.ylabel('Work-Life Balance Score')
plt.title('Work-Life Balance Score of Cities (Scores > 85)')
plt.tight_layout()
plt.show()
```



[ ]:

The second chart is a scatter plot that displays the Affordability vs Total Score. Analyzing this graph can help us understand if there is any correlation between Affordability and Total Score. There seems to be little to no correlation, but one can see that having a score between 60 and 80 in affordability leads to some positive correlation to Total Cost.

[6]:
```
# Scatter Plot - Affordability vs Total Score
plt.figure(figsize=(8, 6))
plt.scatter(data['Affordability'], data['TOTAL SCORE'], color='salmon', alpha=0.
↪6)
plt.xlabel('Affordability')
plt.ylabel('Total Score')
plt.title('Affordability vs Total Score')
plt.tight_layout()
plt.show()
```

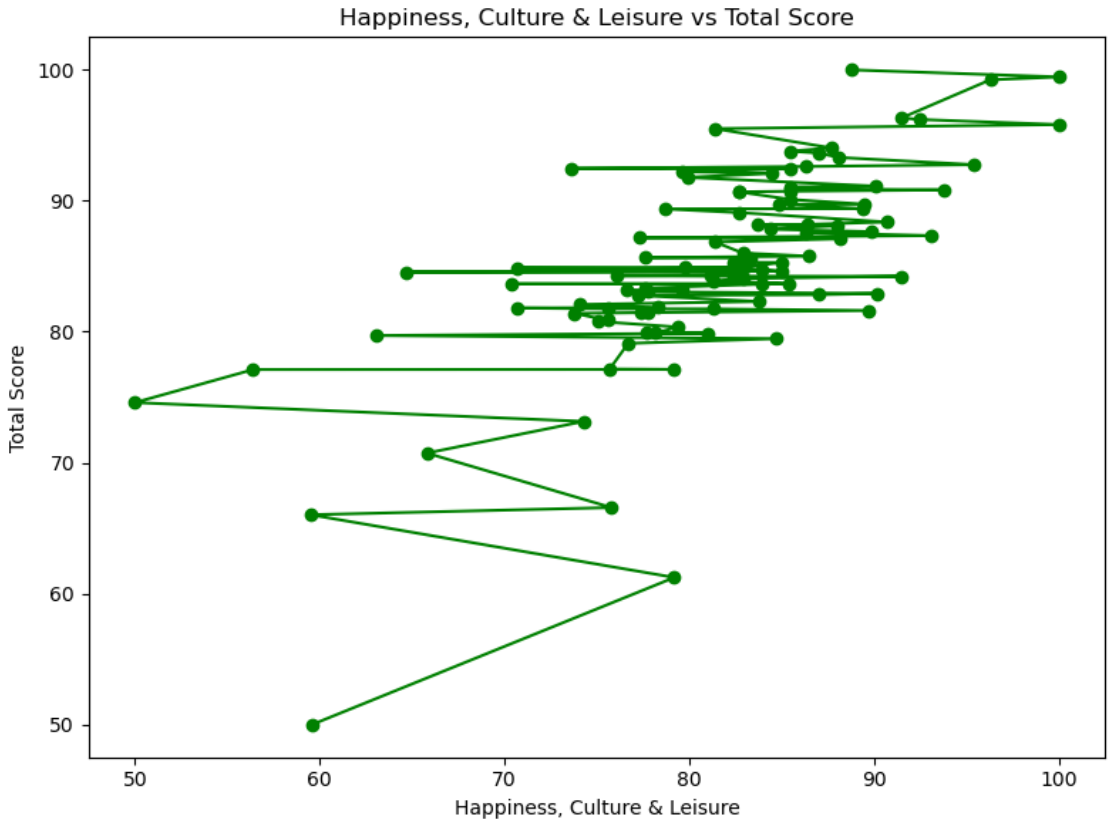


[ ]:

The third graph is a line graph that displays the Happiness, Culture & Leisure vs Total Score. Examining this graph allows us to determine if higher Happiness, Culture & Leisure lead to higher Total Score.

This graph shocked me the most, as although it was a test and not the most appealing, it showed a definite increase signifying that the Happiness, Culture & Leisure is positively correlated to the Total Scores.

[7]:
```
# Line Graph: 'Happiness, Culture & Leisure vs Total Score'
plt.figure(figsize=(8, 6))
plt.plot(data['Happiness, Culture & Leisure'], data['TOTAL SCORE'], marker='o',␣
↪color='green')
plt.xlabel('Happiness, Culture & Leisure')
plt.ylabel('Total Score')
plt.title('Happiness, Culture & Leisure vs Total Score')
plt.tight_layout()
plt.show()
```



[ ]:

Explain what you have learned from each of your graphs: From the bar chart we learned: Cities like Zurich, Copenhagen, and Oslo tend to have the highest work-life balance scores, indicating better overall balance between work and personal life. On the other hand, cities such as Cape Town and Dubai exhibit lower work-life balance scores, suggesting potential challenges in maintaining a healthy balance.

From the scatter plot we learned: The scores vary largely, we can see that the highest Total Score, had an affordability of about 60. This could mean that although it is not affordable there may be other implications that allow such locations to be viewed as one of the best Work-Life cities.

From the line graph we learned: The line graph illustrates that as Happiness, Culture & Leisure increase, there is a tendency for the work-life balance score to increase as well. This suggests that higher Happiness, Culture & Leisure may positively impact individuals' ability to balance their work and personal life (Total Score) effectively.

[ ]:

Write a conclusion that summarizes your findings: In conclusion, analyzing the dataset on cities with the best work-life balance for 2022 provides valuable insights into global trends regarding work-life balace (Total Score), Affordability, and Happiness, Culture & Leisure. Cities with higher

work-life balance scores tend to offer better overall quality of life, often characterized by a medium to high (60 to 80) on affordability, and high values for Happiness, Culture & Leisure. Understanding these dynamics can assist individuals and policymakers in making informed decisions about where to live and work to achieve a healthier work-life balance.

[ ]: