ASumbaraju_wk3_Graphs

June 26, 2021

1 DSC 550

2 Week 3 Assignment

3 Code and 300 word summary justifying ranking

```
[46]: import matplotlib.pyplot as plt
      import numpy as np
      import pandas as pd
      import math
      storms = pd.read_csv("storms_2000_2020.csv")
```

```
[23]: storms.head()
```

```
[23]:    Storm_Year      Storm_Type  severity  severity strength
      0        2001  Thunderstorms      Weak                  2
      1        2002  Thunderstorms  Moderate                  4
      2        2004  Thunderstorms      Weak                  2
      3        2006  Thunderstorms      Weak                  2
      4        2009  Thunderstorms  Moderate                  4
```

```
[48]: df = pd.DataFrame(storms)
      df['Storm_Year'] = df['Storm_Year'].astype(int)
      Storm_Year = df['Storm_Year']
      severity_strength = df['severity strength']
      severity = df['severity']
```

```
[49]: # Line Graph
      labels =df['Storm_Year']
      y_pos = np.arange(len(labels))
      severity_strength =df['severity strength']

      plt.plot(y_pos, severity_strength)
      plt.xticks(y_pos, labels,rotation='vertical')
      plt.xlabel("Year of ThunderStorms")
      plt.ylabel("Severity Scale 1-8")
      plt.title("Severity of ThunderStorms")
```

```
plt.show()
```

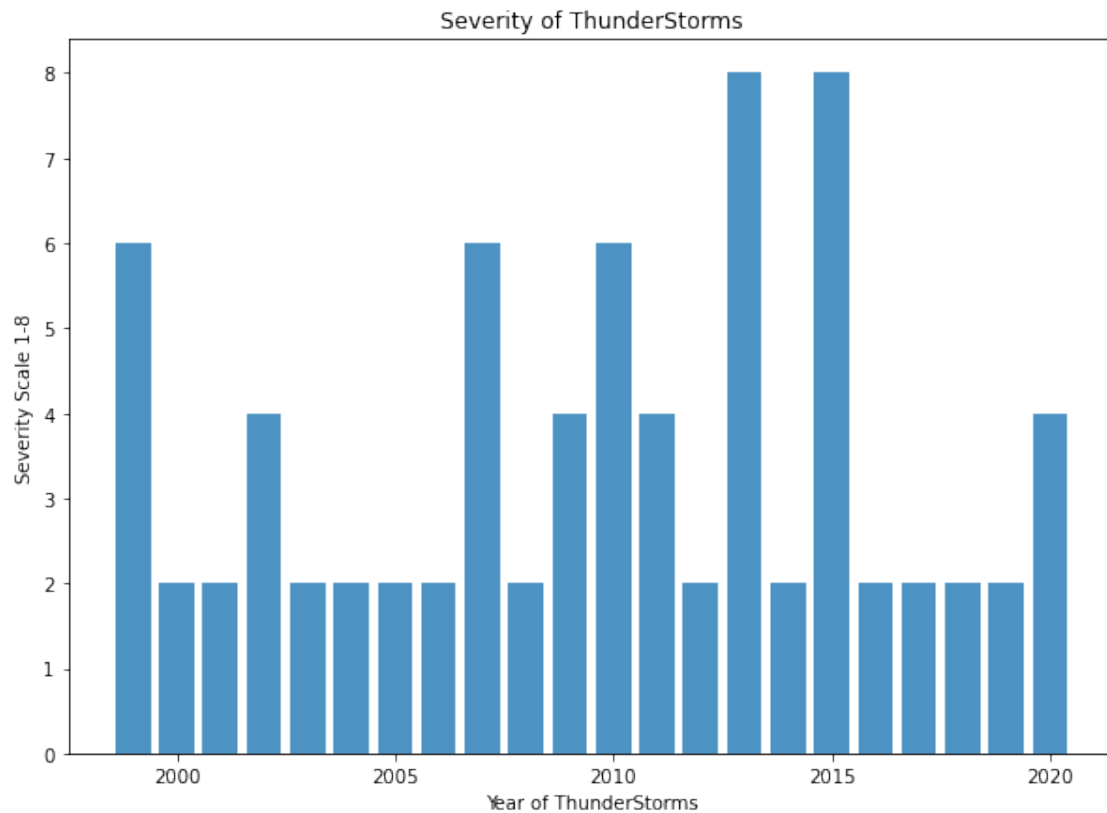## Severity of ThunderStorms
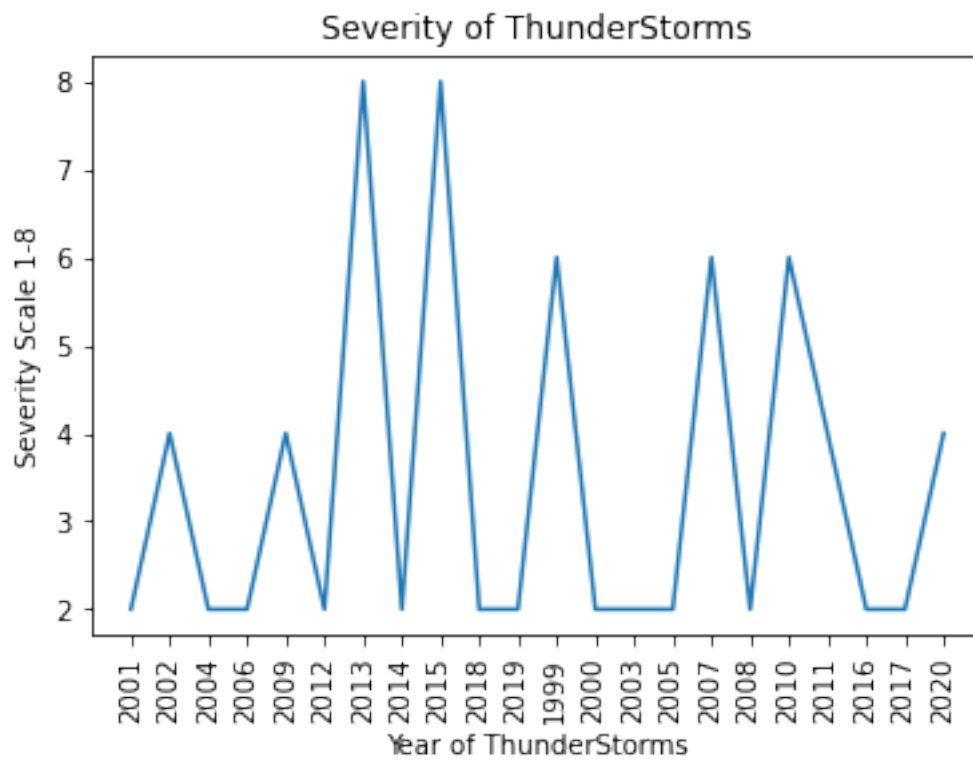

```
[51]:  # Figure Size
       fig = plt.figure(figsize =(10, 7))

       # Horizontal Bar Plot

       plt.bar(Storm_Year, severity_strength, align='center', alpha = 0.8)
       plt.xlabel('Year of ThunderStorms')
       plt.ylabel('Severity Scale 1-8')
       plt.title('Severity of ThunderStorms')


       # Show Plot
       plt.show()
```
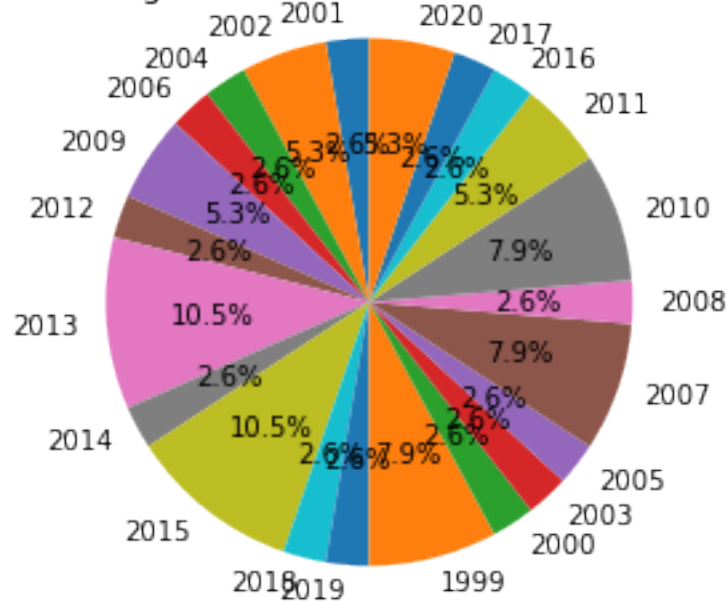
Severity of ThunderStorms

```
# Pie Chart

labels = df['Storm_Year']
sizes = df['severity strength']


fig1, ax1 = plt.subplots()
ax1.pie(sizes, labels=labels, autopct='%1.1f%%',
        shadow=False, startangle=90)
ax1.axis('equal')
plt.title("Severity Percentage of Occurrences of thunderstorms per year ")
plt.show()
```

Severity Percentage of Occurrences of thunderstorms per year

3.1 Graphs Summary

The Dataset storms_2000_2020.csv deals with the number of occurrences of the thunderstorms per Year w.r.t its severity. At the same time, it describes the severity of the storm as either weak, moderate, strong, or very strong, along with the strength being 2, 4, 6, or 8, respectively. For analysis, I chose to use the severity of the thunderstorms across the years that have occurred within the data.

3.2 Line Chart:

Line graphs are used to compare changes over the same period for more than one group. It is the most effective visual representation for this type of data in showing how the severity has grown and decreased over the years ranging from 2000-2020 contained within the data. The nature of data exemplifies "Trends" – hence I would like to assign the first rank to it.

3.3 Bar Chart:

Looking at the bar chart shows the strength of the years that the thunderstorm occurred and compared the severity levels w.r.t years. However, it does not show how the severity % variations changed throughout the years, which I feel is the most crucial aspect of the data. The line chart shows the same aspect as the bar in that it can represent the strength as well. The main difference and benefit are that it is easier to represent the trends. I have decided to rank this visual as the second effective.

3.4 Pie Chart:

Lastly, it is an excellent visual to express the occurrences of different strengths of the thunderstorms overall with the pie chart. I have managed to represent data by its severity percentage over the period of years.But this is specific to identify the percentage of stake and this is unlikely conceptuaized as "trend" graph. hence i would like to #Rank3.

I want to conclude that each visual has its strengths and weakness. Still, it will always depend on the input data, and most importantly, "storytelling" skills are essential on how best we present the data to the target audience.

[ ]: