**A Problem with Presidents**

**Mustafa Gangardiwala**

**June 2, 2023**

1. **Introduction**

The given dataset with the problem statements contains information about each president starting from first president George Washington to current president Joe Biden. The information is birth date, birthplace, death date and location of death.

1. **Implementation**
   1. **Technologies**
      1. I have used Python for the problem and performing to task asked in the problem statement.
      2. The data analysis part is done in a Jupyter Notebook.
   2. **Libraires**
      1. Pandas
      2. NumPy
      3. Datetime
      4. Matplotlib
      5. Seaborn
   3. **Assumptions**
      1. The number of years, months, and days a president has lived is calculated from their birthday to present date which is June 2, 2023.
      2. While calculating weighted average I have used president’s lived years as weights
   4. **Running Instruction**
      1. Requirement.txt contains all the libraries that are used in the analysis to and can be ran using “*pip install -r requirements.txt*”
      2. The analysis is in the “Solution.ipynb” file which is Jupyter Notebook.
2. **Summary**

The dataset contains 47 rows, with 5 columns out which last row is a reference which I removed. The longest-lived president is Jimmy Carter aged 99 years old currently. The shortest-lived president is John F. Kennedy who died at the age of 46 years old. The average life span of a US president is 72.467 years which comes out to be 26446 days and median is 72 years which is 26227 days.

1. **Observations**
   1. **Top 10 Presidents from longest lived to shortest lived**

**A picture containing text, screenshot, black and white, black

Description automatically generated**

Figure 1 Top 10 Longest lived presidents

* 1. **Top 10 Presidents from shortest lived to longest lived**

**A picture containing text, screenshot, black and white, black

Description automatically generated**

Figure 2 Top 10 Shortest lived presidents

* 1. **A screenshot of a computer

     Description automatically generated with medium confidenceStatistical Values**

Figure 3 Statistical Data about "lived\_days"

**A picture containing text, screenshot, font

Description automatically generated**

Figure 4 Statistical Data about newly added columns

**A picture containing text, diagram, screenshot, plot

Description automatically generated**

Figure 5 Box Plot with histogram showing the desnsity of presidents lived days

I decided to go with a box plot because it the way to show distribution of data.

1. **Conclusion**

Based on the above analysis it can be concluded that the average life-span of a US president is 72.467 years with the range being between 46 years to 99 years. There is a need for more data for further analysis.

1. **Future Scope**

A heat maps of US states in which presidents were born and died can created for further analysis. Apart, from that we can also include the data about the causes of deaths of US presidents. Another, thing I want to do in the analysis is to see the average life-span of a president after leaving office.