

# A Problem with Presidents: Converting Data to Information

The objective of this project is to demonstrate the process of converting data into information using Python and mathematics. I analyzed the data related to the lifespan of US Presidents and perform basic exploratory data analysis with the dataset given and broken down into the three figures. Each of these figures shows what I have found in terms of solving this problem.

Figure 1 shows that the top 10 presidents from longest lived to shortest lived and the top 10 presidents from shortest lived to longest lived. I've found that President George Bush lived the longest, with a lifespan of 94 years and 84 days. While President John F. Kennedy had the shortest lifespan, with 46 years and 183 days. Both presidents are statistically recorded as Maximum Value and Minimum Value respectively in the next figure.

Figure 2 shows the statistical table of the average lifespan of US Presidents. The Mean is calculated from adding all the lived\_days variable and divide it by the total number of presidents. The Weighted Average is similar to the mean but add significance for every year the Presidents have lived. The Standard Deviation for this particular dataset is 4596 (around 12 years). It is also important to note that there are living presidents (such as Jimmy Carter) are not part of this data.

Figure 3 shows a histogram of how many frequently range of values for every president have lived. It revealed that the distribution of the data peaked in both 21000 and 25000 days where presidents lived in those two values are the most frequent. The right end of the Histogram is rising in frequency as an average lifespan of healthy Presidents that last longer while the left end of the Histogram listing Presidents who are mostly assassinated.

Concluding this report using Python to parse the given dataset of US presidents, my analysis of the US Presidents' lifespan data revealed insights that can be used to gain a better understanding of this historical aspect of the US Presidency. The use of descriptive statistics and visualization techniques allowed us to communicate the findings in a way that is easily understandable to a broad audience.

## References

[1] Woolf, M. (n.d.). Statistics and Probability Background. LibreTexts.  
[https://eng.libretexts.org/Bookshelves/Industrial\\_and\\_Systems\\_Engineering/Book%3A\\_Chemical\\_Process\\_Dynamics\\_and\\_Controls\\_\(Woolf\)/13%3A\\_Statistics\\_and\\_Probability\\_Background/13.01%3A\\_Basic\\_statistics-\\_mean%2C\\_median%2C\\_average%2C\\_standard\\_deviation%2C\\_z-scores%2C\\_and\\_p-value](https://eng.libretexts.org/Bookshelves/Industrial_and_Systems_Engineering/Book%3A_Chemical_Process_Dynamics_and_Controls_(Woolf)/13%3A_Statistics_and_Probability_Background/13.01%3A_Basic_statistics-_mean%2C_median%2C_average%2C_standard_deviation%2C_z-scores%2C_and_p-value).

```
=====
Table of Top 10 Presidents
=====
Top 10 Presidents from longest lived to shortest lived:
PRESIDENT YEAR_OF_BIRTH LIVED_YEARS LIVED_MONTHS LIVED_DAYS
39 George Bush 1924 94 1,134 34,504
36 Gerald Ford 1913 93 1,121 34,133
38 Ronald Reagan 1911 93 1,120 34,088
1 John Adams 1735 91 1,088 33,119
29 Herbert Hoover 1874 90 1,082 32,943
31 Harry S. Truman 1884 89 1,064 32,373
3 James Madison 1751 85 1,023 31,150
2 Thomas Jefferson 1743 83 999 30,397
35 Richard Nixon 1913 81 975 29,688
5 John Quincy Adams 1767 81 967 29,446
Top 10 presidents from shortest lived to longest lived:
PRESIDENT YEAR_OF_BIRTH LIVED_YEARS LIVED_MONTHS LIVED_DAYS
33 John F. Kennedy 1917 46 558 16,978
19 James A. Garfield 1831 50 598 18,202
10 James K. Polk 1795 54 643 19,583
15 Abraham Lincoln 1809 56 674 20,516
20 Chester Arthur 1829 57 685 20,863
27 Warren G. Harding 1865 58 693 21,091
23 William McKinley 1843 59 703 21,412
24 Theodore Roosevelt 1858 60 722 21,985
28 Calvin Coolidge 1872 61 726 22,099
30 Franklin Roosevelt 1882 63 758 23,082
```

Figure 1 – The Top 10 Lists

```
=====
Statistics for Lived Days
=====
Mean: 26,097
Weighted Average: 26,886
Median: 25,673
Mode: nan
Max: 34,504
Min: 16,978
Standard Deviation: 4,596
```

Figure 2 – The Statistics

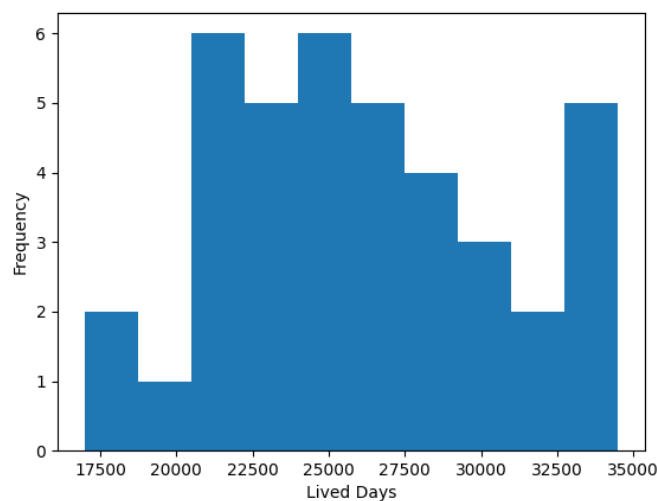


Figure 3 – The Histogram