

A Problem with Presidents: A Data to Information Challenge

Nikhil Yerra

August 10, 2024

Introduction

The analysis of U.S. Presidents' lifespans provides valuable insights into the longevity of these historical figures. By examining the data on their birth and death dates, we can compute various measures of lifespan, including the mean, median, mode, and standard deviation. This project aims to uncover trends and patterns in the lifespans of U.S. Presidents and to highlight both the shortest and longest-lived individuals in this group. The results of this analysis can offer a deeper understanding of presidential longevity.

Summary

- The average lifespan is approximately 72.5 years (26,504 days). The variation in lifespan is considerable, ranging from a minimum of 46.4 years (16,978 days) to a maximum of 99.8 years (36,473 days).
- The analysis highlights both the shortest and longest-lived presidents, with John F. Kennedy and Jimmy Carter representing the extremes of this spectrum.

Data Preparation

Parsing Dates and Calculating Longevity

The first step in the analysis was to parse the birth and death dates from the dataset provided. The dates were then used to calculate the number of years, months, and days each president lived.

Calculations

- **Lived Years, Months, and Days:** The difference between the birth and death dates was calculated in terms of years, months, and days. This calculation accounted for leap years and varying month lengths by using averages (365.25 days per year and 30.5 days per month).
- **Statistical Measures:** Key statistical measures were computed for the lifespan of presidents in days:
 - **Mean:** The average number of days lived by the presidents.
 - **Weighted Average:** Given that all presidents have equal importance in this analysis, the weighted average was calculated as the arithmetic mean.

- **Median:** The middle value when the days lived were sorted.
- **Mode:** The most frequently occurring number of days lived.
- **Maximum:** The highest number of days lived by a president.
- **Minimum:** The lowest number of days lived by a president.
- **Standard Deviation:** The measure of variability in the number of days lived.

| US Presidents Dataset | | | | | | | | | |
|-----------------------|------------------------|------------|-------------------------|------------|-----------------------|---------------|-------------|--------------|------------|
| | PRESIDENT | BIRTH DATE | BIRTH PLACE | DEATH DATE | LOCATION OF DEATH | YEAR OF BIRTH | LIVED YEARS | LIVED MONTHS | LIVED DAYS |
| 0 | George Washington | 1732-02-22 | Westmoreland Co., Va. | 1799-12-14 | Mount Vernon, Va. | 1732 | 67.8 | 812.0 | 24767 |
| 1 | John Adams | 1735-10-30 | Quincy, Mass. | 1826-07-04 | Quincy, Mass. | 1735 | 90.7 | 1085.9 | 33119 |
| 2 | Thomas Jefferson | 1743-04-13 | Albemarle Co., Va. | 1826-07-04 | Albemarle Co., Va. | 1743 | 83.2 | 996.6 | 30397 |
| 3 | James Madison | 1751-03-16 | Port Conway, Va. | 1836-06-28 | Orange Co., Va. | 1751 | 85.3 | 1021.3 | 31150 |
| 4 | James Monroe | 1758-04-28 | Westmoreland Co., Va. | 1831-07-04 | New York, New York | 1758 | 73.2 | 876.4 | 26729 |
| 5 | John Quincy Adams | 1767-07-11 | Quincy, Mass. | 1848-02-23 | Washington, D.C. | 1767 | 80.6 | 965.4 | 29446 |
| 6 | Andrew Jackson | 1767-03-15 | Waxhaws, No/So Carolin | 1845-06-08 | Nashville, Tennessee | 1767 | 78.2 | 936.9 | 28574 |
| 7 | Martin Van Buren | 1782-12-05 | Kinderhook, New York | 1862-07-24 | Kinderhook, New York | 1782 | 79.6 | 953.6 | 29085 |
| 8 | William Henry Harrison | 1773-02-09 | Charles City Co., Va. | 1841-04-04 | Washington, D.C. | 1773 | 68.1 | 816.1 | 24890 |
| 9 | John Tyler | 1790-03-29 | Charles City Co., Va. | 1862-01-18 | Richmond, Va. | 1790 | 71.8 | 859.9 | 26227 |
| 10 | James K. Polk | 1795-11-02 | Mecklenburg Co., N.C. | 1849-06-15 | Nashville, Tennessee | 1795 | 53.6 | 642.1 | 19583 |
| 11 | Zachary Taylor | 1784-11-24 | Orange County, Va. | 1850-07-09 | Washington, D.C | 1784 | 65.6 | 785.8 | 23967 |
| 12 | Millard Fillmore | 1800-01-07 | Cayuga Co., New York | 1874-03-08 | Buffalo, New York | 1800 | 74.2 | 888.1 | 27088 |
| 13 | Franklin Pierce | 1804-11-23 | Hillsborough, N.H. | 1869-10-08 | Concord, New Hamp. | 1804 | 64.9 | 776.9 | 23695 |
| 14 | James Buchanan | 1791-04-23 | Cove Gap, Pa. | 1868-06-01 | Lancaster, Pa. | 1791 | 77.1 | 923.4 | 28163 |
| 15 | Abraham Lincoln | 1809-02-12 | LaRue Co., Kentucky | 1865-04-15 | Washington, D.C. | 1809 | 56.2 | 672.7 | 20516 |
| 16 | Andrew Johnson | 1808-12-29 | Raleigh, North Carolina | 1875-07-31 | Elizabethhton, Tenn. | 1808 | 66.6 | 797.4 | 24320 |
| 17 | Ulysses S. Grant | 1822-04-27 | Point Pleasant, Ohio | 1885-07-23 | Wilton, New York | 1822 | 63.2 | 757.3 | 23098 |
| 18 | Rutherford B. Hayes | 1822-10-04 | Delaware, Ohio | 1893-01-17 | Fremont, Ohio | 1822 | 70.3 | 841.7 | 25673 |
| 19 | James A. Garfield | 1831-11-19 | Cuyahoga Co., Ohio | 1881-09-19 | Elberon, New Jersey | 1831 | 49.8 | 596.8 | 18202 |
| 20 | Chester Arthur | 1829-10-05 | Fairfield, Vermont | 1886-11-18 | New York, New York | 1829 | 57.1 | 684.0 | 20863 |
| 21 | Grover Cleveland | 1837-03-18 | Caldwell, New Jersey | 1908-06-24 | Princeton, New Jersey | 1837 | 71.3 | 853.4 | 26030 |
| 22 | Benjamin Harrison | 1833-08-20 | North Bend, Ohio | 1901-03-13 | Indianapolis, Indiana | 1833 | 67.6 | 809.0 | 24676 |
| 23 | William McKinley | 1843-01-29 | Niles, Ohio | 1901-09-14 | Buffalo, New York | 1843 | 58.6 | 702.0 | 21412 |
| 24 | Theodore Roosevelt | 1858-10-27 | New York, New York | 1919-01-06 | Oyster Bay, New York | 1858 | 60.2 | 720.8 | 21985 |
| 25 | William Howard Taft | 1857-09-15 | Cincinnati, Ohio | 1930-03-08 | Washington, D.C. | 1857 | 72.5 | 867.9 | 26471 |
| 26 | Woodrow Wilson | 1856-12-28 | Staunton, Virginia | 1924-02-03 | Washington, D.C. | 1856 | 67.1 | 803.5 | 24507 |
| 27 | Warren G. Harding | 1865-11-02 | Morrow County, Ohio | 1923-08-02 | San Francisco, Cal. | 1865 | 57.7 | 691.5 | 21091 |
| 28 | Calvin Coolidge | 1872-07-04 | Plymouth, Vermont | 1933-01-05 | Northampton, Mass. | 1872 | 60.5 | 724.6 | 22099 |
| 29 | Herbert Hoover | 1874-08-10 | West Branch, Iowa | 1964-10-20 | New York, New York | 1874 | 90.2 | 1080.1 | 32943 |
| 30 | Franklin Roosevelt | 1882-01-30 | Hyde Park, New York | 1945-04-12 | Warm Springs, Ga. | 1882 | 63.2 | 756.8 | 23082 |
| 31 | Harry S. Truman | 1884-05-08 | Lamar, Missouri | 1972-12-26 | Kansas City, Missouri | 1884 | 88.6 | 1061.4 | 32373 |
| 32 | Dwight Eisenhower | 1890-10-14 | Denison, Texas | 1969-03-28 | Washington, D.C. | 1890 | 78.5 | 939.5 | 28654 |
| 33 | John F. Kennedy | 1917-05-29 | Brookline, Mass. | 1963-11-22 | Dallas, Texas | 1917 | 46.5 | 556.7 | 16978 |
| 34 | Lyndon B. Johnson | 1908-08-27 | Gillespie Co., Texas | 1973-01-22 | Gillespie Co., Texas | 1908 | 64.4 | 771.3 | 23524 |
| 35 | Richard Nixon | 1913-01-09 | Yorba Linda, Cal. | 1994-04-22 | New York, New York | 1913 | 81.3 | 973.4 | 29688 |
| 36 | Gerald Ford | 1913-07-14 | Omaha, Nebraska | 2006-12-26 | Rancho Mirage, Cal. | 1913 | 93.5 | 1119.1 | 34133 |
| 37 | Jimmy Carter | 1924-10-01 | Plains, Georgia | 2024-08-10 | nan | 1924 | 99.9 | 1195.8 | 36473 |
| 38 | Ronald Reagan | 1911-02-06 | Tampico, Illinois | 2004-06-05 | Los Angeles, Cal. | 1911 | 93.3 | 1117.6 | 34088 |
| 39 | George Bush | 1924-06-12 | Milton, Mass. | 2018-11-30 | Houston, Texas | 1924 | 94.5 | 1131.3 | 34504 |
| 40 | Bill Clinton | 1946-08-19 | Hope, Arkansas | 2024-08-10 | nan | 1946 | 78.0 | 933.8 | 28481 |
| 41 | George W. Bush | 1946-07-06 | New Haven, Conn. | 2024-08-10 | nan | 1946 | 78.1 | 935.2 | 28525 |
| 42 | Barack Obama | 1961-08-04 | Honolulu, Hawaii | 2024-08-10 | nan | 1961 | 63.0 | 754.7 | 23017 |
| 43 | Donald Trump | 1946-06-14 | New York, New York | 2024-08-10 | nan | 1946 | 78.2 | 936.0 | 28547 |
| 44 | Joe Biden | 1942-11-20 | Scranton, Pa. | 2024-08-10 | nan | 1942 | 81.7 | 978.7 | 29849 |

Figure 1: US President Dataset

Results

Longest and Shortest Lived Presidents

Two tables were generated to rank the U.S. Presidents by their lifespan:

Longest Lived Presidents

The top 10 longest-lived presidents are presented in descending order of the number of years they lived. This table helps identify those who had the greatest longevity.

| PRESIDENT | BIRTH DATE | BIRTH PLACE | DEATH DATE | LOCATION OF DEATH | YEAR OF BIRTH | LIVED YEARS | LIVED MONTHS | LIVED DAYS |
|------------------|------------|--------------------|------------|-----------------------|---------------|-------------|--------------|------------|
| Jimmy Carter | 1924-10-01 | Plains, Georgia | 2024-08-10 | nan | 1924 | 99.9 | 1195.8 | 36473.0 |
| George Bush | 1924-06-12 | Milton, Mass. | 2018-11-30 | Houston, Texas | 1924 | 94.5 | 1131.3 | 34504.0 |
| Gerald Ford | 1913-07-14 | Omaha, Nebraska | 2006-12-26 | Rancho Mirage, Cal. | 1913 | 93.5 | 1119.1 | 34133.0 |
| Ronald Reagan | 1911-02-06 | Tampico, Illinois | 2004-06-05 | Los Angeles, Cal. | 1911 | 93.3 | 1117.6 | 34088.0 |
| John Adams | 1735-10-30 | Quincy, Mass. | 1826-07-04 | Quincy, Mass. | 1735 | 90.7 | 1085.9 | 33119.0 |
| Herbert Hoover | 1874-08-10 | West Branch, Iowa | 1964-10-20 | New York, New York | 1874 | 90.2 | 1080.1 | 32943.0 |
| Harry S. Truman | 1884-05-08 | Lamar, Missouri | 1972-12-26 | Kansas City, Missouri | 1884 | 88.6 | 1061.4 | 32373.0 |
| James Madison | 1751-03-16 | Port Conway, Va. | 1836-06-28 | Orange Co., Va. | 1751 | 85.3 | 1021.3 | 31150.0 |
| Thomas Jefferson | 1743-04-13 | Albemarle Co., Va. | 1826-07-04 | Albemarle Co., Va. | 1743 | 83.2 | 996.6 | 30397.0 |
| Joe Biden | 1942-11-20 | Scranton, Pa. | 2024-08-10 | nan | 1942 | 81.7 | 978.7 | 29849.0 |

Figure 2: Top 10 Presidents from longest lived to shortest lived

Shortest Lived Presidents

The top 10 shortest-lived presidents are presented in ascending order of the number of years they lived. This table highlights those who had the shortest lifespan.

| PRESIDENT | BIRTH DATE | BIRTH PLACE | DEATH DATE | LOCATION OF DEATH | YEAR OF BIRTH | LIVED YEARS | LIVED MONTHS | LIVED DAYS |
|--------------------|------------|-----------------------|------------|----------------------|---------------|-------------|--------------|------------|
| John F. Kennedy | 1917-05-29 | Brookline, Mass. | 1963-11-22 | Dallas, Texas | 1917 | 46.5 | 556.7 | 16978.0 |
| James A. Garfield | 1831-11-19 | Cuyahoga Co., Ohio | 1881-09-19 | Elberon, New Jersey | 1831 | 49.8 | 596.8 | 18202.0 |
| James K. Polk | 1795-11-02 | Mecklenburg Co., N.C. | 1849-06-15 | Nashville, Tennessee | 1795 | 53.6 | 642.1 | 19583.0 |
| Abraham Lincoln | 1809-02-12 | LaRue Co., Kentucky | 1865-04-15 | Washington, D.C. | 1809 | 56.2 | 672.7 | 20516.0 |
| Chester Arthur | 1829-10-05 | Fairfield, Vermont | 1886-11-18 | New York, New York | 1829 | 57.1 | 684.0 | 20863.0 |
| Warren G. Harding | 1865-11-02 | Morrow County, Ohio | 1923-08-02 | San Francisco, Cal. | 1865 | 57.7 | 691.5 | 21091.0 |
| William McKinley | 1843-01-29 | Niles, Ohio | 1901-09-14 | Buffalo, New York | 1843 | 58.6 | 702.0 | 21412.0 |
| Theodore Roosevelt | 1858-10-27 | New York, New York | 1919-01-06 | Oyster Bay, New York | 1858 | 60.2 | 720.8 | 21985.0 |
| Calvin Coolidge | 1872-07-04 | Plymouth, Vermont | 1933-01-05 | Northampton, Mass. | 1872 | 60.5 | 724.6 | 22099.0 |
| Barack Obama | 1961-08-04 | Honolulu, Hawaii | 2024-08-10 | nan | 1961 | 63.0 | 754.7 | 23017.0 |

Figure 3: Top 10 presidents from shortest lived to longest lived

Statistical Summary Table

A summary table was constructed to present the calculated statistics for the number of days lived by U.S. Presidents:

| | Measure | Lived Days |
|---|--------------------|--------------|
| 0 | Mean | 26504.044444 |
| 1 | Weighted Average | 26504.044444 |
| 2 | Median | 26227.000000 |
| 3 | Mode | 16978.000000 |
| 4 | Max | 36473.000000 |
| 5 | Min | 16978.000000 |
| 6 | Standard Deviation | 4632.047677 |

Figure 4: Statistics of the dataset

Visualization: Distribution of Days Lived

A histogram was created to visualize the distribution of the number of days lived by U.S. Presidents. The histogram includes a density plot overlay to better understand the data's distribution. The plot shows that most presidents lived between 22,500 and 30,000 days, with a noticeable skew towards longer lifespans.

Methodology

Data Handling

- **Date Parsing:** The dates were parsed using two formats: abbreviated and full month names. Any unparseable dates were converted to NaT (Not a Time).
- **Calculation of Time Differences:** The time differences were calculated in days, then converted to years and months using appropriate average values.
- **Handling Missing Data:** Presidents with unknown death dates were assumed to have lived until August 10, 2024, for the purpose of this analysis. This approach was applied to a very small number of entries and allows for a more complete dataset while acknowledging that these individuals are still alive.

Assumptions

- **Leap Years:** An average of 365.25 days per year was used to account for leap years.
- **Month Length:** An average of 30.5 days per month was used, considering months with different lengths.
- **Equal Weights:** All presidents were given equal importance when calculating the weighted average.

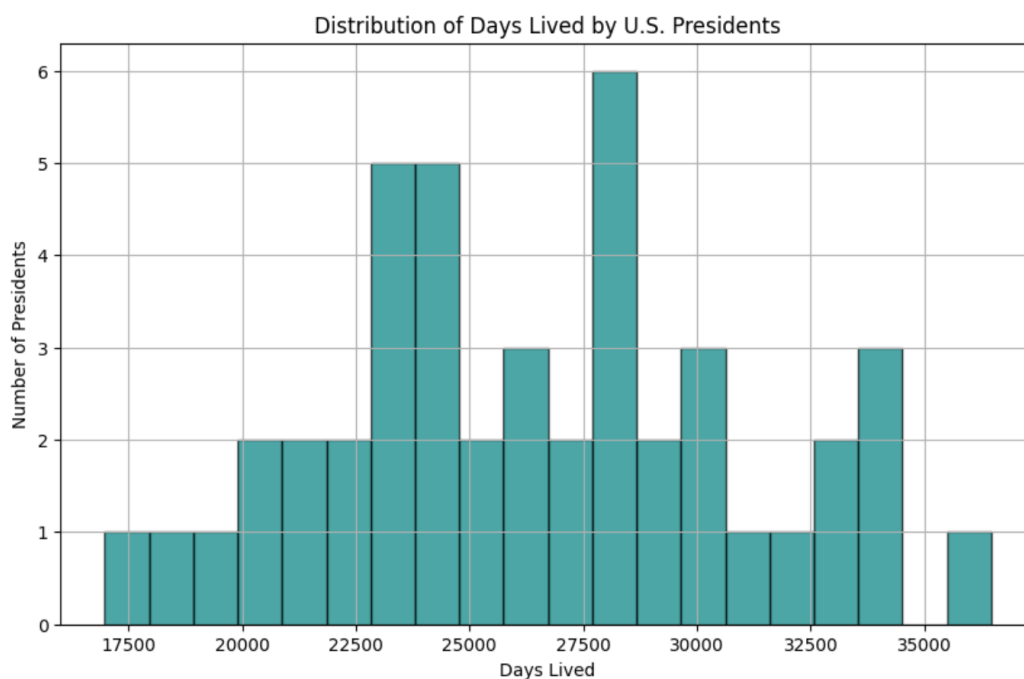


Figure 5: Distribution of Data

Conclusion

The analysis of U.S. Presidents' lifespans provides valuable insights into the longevity of these historical figures. The statistical summary indicates that the mean and weighted average lifespan is approximately 26,504 days, or around 72.5 years. This suggests that most presidents have lived into their seventies, with a few notable outliers on both ends of the spectrum.

The shortest-lived presidents include John F. Kennedy, who lived for just 16,978 days (46.4 years), while the longest-lived president, Jimmy Carter, has surpassed 36,000 days (99.8 years) and is still alive as of the analysis date. The histogram plot vividly illustrates this distribution, highlighting that although most presidents' lifespans cluster around the seventies, there are significant variations.