

A Problem with Presidents

Author: Dharani Kanchanapalli

Date: 27th July, 2023

Introduction:

We have a dataset with 44 presidents of the United States containing data about their births and deaths. In order to gather meaningful insights from this data, we perform steps such as data preprocessing, feature engineering, sorting and data visualisation. This report summarises all my findings after performing the necessary analysis and visualisations

Technologies & Libraries used:

I have used python and Jupyter notebooks to solve this problem. I have also leveraged libraries such as pandas & numpy to perform data preprocessing such as adding missing values, dropping rows with irrelevant data, and converting columns into different data types. I have also added new features (columns) to this dataset and computed values in those columns , sorted and filtered data. I used matplotlib and seaborn libraries to plot visually appealing graphs that enable us make conclusions from this data.

Tables & Graphs:

Top 10 Presidents from Shortest to Longest Lived

	PRESIDENT	Birth Year	Age	Days Lived
0	John F. Kennedy	1917	46	16978
1	James A. Garfield	1831	50	18202
2	James K. Polk	1795	54	19583
3	Abraham Lincoln	1809	56	20516
4	Chester Arthur	1829	57	20863
5	Warren G. Harding	1865	58	21091
6	William McKinley	1843	58	21412
7	Theodore Roosevelt	1858	61	21985
8	Calvin Coolidge	1872	61	22099
9	Barack Obama	1961	62	22637

TOP 10 Deceased Presidents from shortest to longest lived

	PRESIDENT	Birth Year	Age	Days Lived
0	John F. Kennedy	1917	46	16978
1	James A. Garfield	1831	50	18202
2	James K. Polk	1795	54	19583
3	Abraham Lincoln	1809	56	20516
4	Chester Arthur	1829	57	20863
5	Warren G. Harding	1865	58	21091
6	William McKinley	1843	58	21412
7	Theodore Roosevelt	1858	61	21985
8	Calvin Coolidge	1872	61	22099
9	Barack Obama	1961	62	22637

TOP 10 Presidents from longest to shortest lived

	PRESIDENT	Birth Year	Age	Days Lived
0	Jimmy Carter	1924	99	36093
1	George Bush	1924	94	34504
2	Gerald Ford	1913	93	34133
3	Ronald Reagan	1911	93	34088
4	John Adams	1735	91	33119
5	Herbert Hoover	1874	90	32943
6	Harry S. Truman	1884	88	32373
7	James Madison	1751	85	31150
8	Thomas Jefferson	1743	83	30397
9	Richard Nixon	1913	81	29688

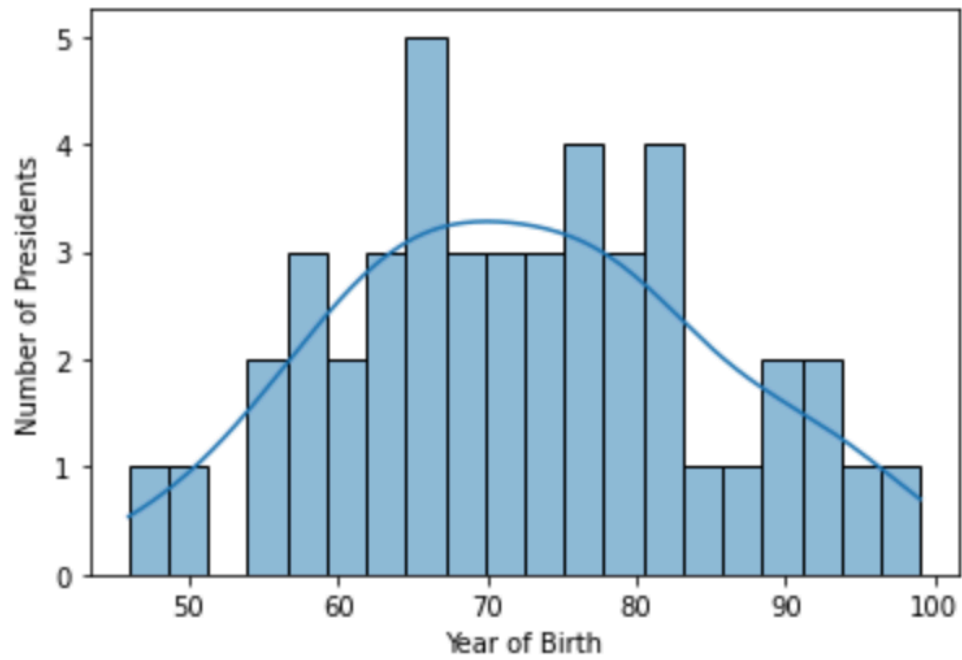
Longest to Shortest LIVING presidents

	PRESIDENT	Birth Year	Age	Days Lived
0	Jimmy Carter	1924	99	36093
1	Joe Biden	1942	81	29469
2	Donald Trump	1946	77	28167
3	George W. Bush	1946	77	28145
4	Bill Clinton	1946	77	28101
5	Barack Obama	1961	62	22637

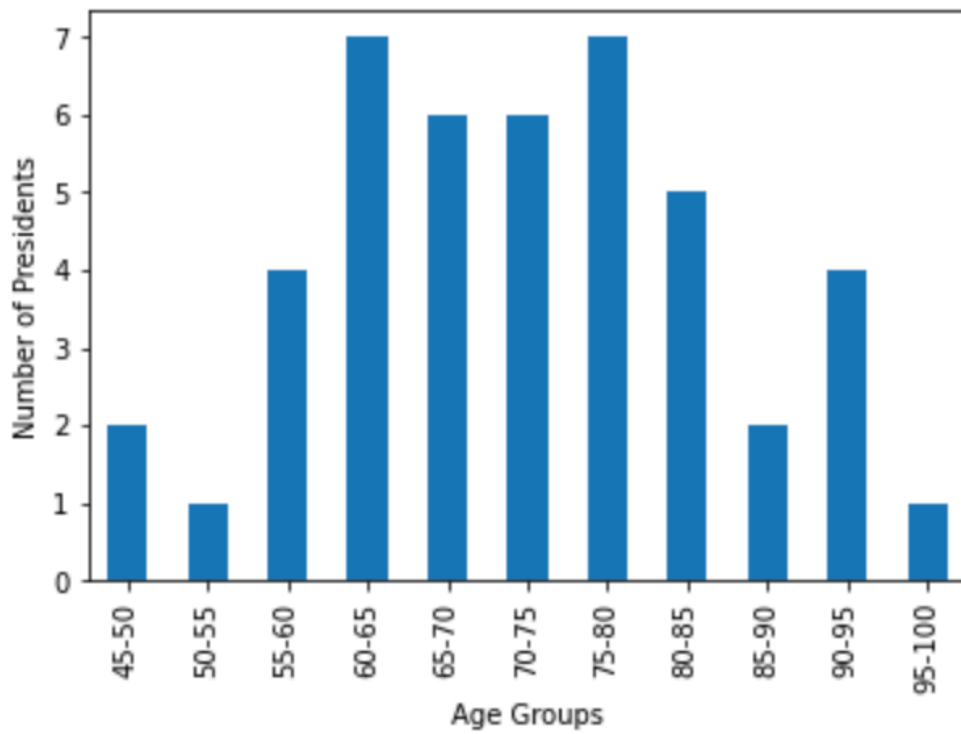
Statistical Analysis:

	Metric	Value
0	Mean	26453.38
1	Weighted Average	27231.46
2	Median	26227.00
3	Mode	16978.00
4	Maximum Value	36093.00
5	Minimum Value	16978.00
6	Standard Deviation	4604.22

Kernel Density Estimation of the Data:



Frequency Distribution of ages over different age groups:



Observations:

- The oldest living president is still alive :).
- John F Kennedy & George Bush are the youngest & oldest deceased presidents respectively.
- Barack Obama is the youngest amongst all living presidents.
- The average life expectancy of a president is around 72 years.
- The weighted average age of presidents is around 74 years.
- KDE is used to create a smooth curve over the given data and it helps us visualise the “shape” of the data.
- According to the frequency distribution, most presidents have died between the ages of 60-65 and 75-80.

Future Scope:

As we get more data on presidents, we can make more conclusions about the data. Also having more features such as data about health, lifestyle of presidents would enable us to gather more insights about their life expectancies.