Analysis of postwar presidents by debt increases 1949-2016

By Ayo Akinnawonu

**Overview**

The following dataset has the percentage increase of the US national debt since the beginning of the of modern era following the Second World War. The national debt has always been a fixture of the news regardless of which party is in power as it is often used to qualify how much spending the US government is doing at the time. The national debt has ballooned to tens of trillions of dollars as of 2019. I believe that it is important to contextualize the national debt in terms of when and under whose administration the debt tends to have the largest increases. The data set is from <https://data.world/kevinnayar/us-presidents-and-debt> and was created from info provided by the Office of Management and Budget.

**About the Data**

> presidential\_data<-read.csv(file.choose())

> View(presidential\_data)

The dataset contains information about all the postwar presidents from Harry Truman to Barack Obama. The data begins from 1949, presumably because that is the term in which Harry Truman was personally elected president of the United States and would also be a good starting point in order to distance the spending from WWII which would have been outliers in the data if the previous presidential term was included. The matrix contains 4 variables with 67 objects. The columns contain data on the year, the incumbent in the white house who would have made policy decisions, their party affiliation and finally, the change percentage in the debt that year.

**Analysis**

aggregate(annual\_debt\_increase\_pct~president, presidential\_data, mean)

president annual\_debt\_increase\_pct

1 Barack Obama 8.325000

2 Bill Clinton 4.287500

3 Dwight D. Eisenhower 1.262500

4 George H. W. Bush 11.825000

5 George W. Bush 7.387500

6 Gerald Ford 10.766667

7 Harry S. Truman 0.675000

8 Jimmy Carter 9.975000

9 John F. Kennedy 2.233333

10 Lyndon B. Johnson 2.600000

11 Richard Nixon 5.700000

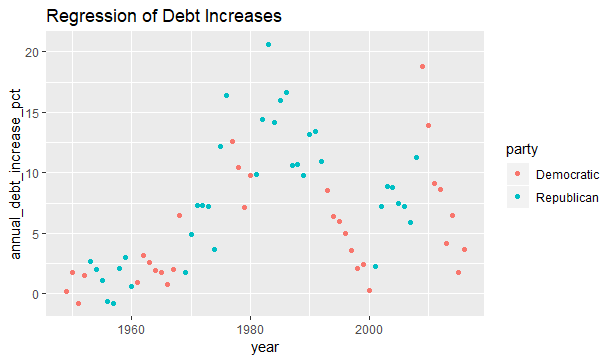
12 Ronald Reagan 14.125000

Using the aggregate function, I calculated the average percentage increase in the debt of in each of the postwar presidents. The above piece of code displays that the early presidents from Truman to Johnson had only small average increases from 0.6% to 2.6%. The averages raise significantly from Nixon to Reagan with Ronald Reagan being the high point in the percentage increases. It is important to note that the four highest averages are held by the back to back presidencies of Ford, Carter, Reagan and Bush Sr.

|  |
| --- |
| aggregate(annual\_debt\_increase\_pct~party, presidential\_data, mean)  party annual\_debt\_increase\_pct  1 Democratic 5.100000  2 Republican 8.063889 |
|  |
| |  | | --- | | > | |

The above piece shows the average of the postwar presidents when they are comparted to along party lines. It shows that the debt has its greatest increases during the administrations of Republican presidents with the top three increasers of the national debt all happening during Republican presidencies.

ggplot() + geom\_point(data=presidential\_data, aes(x=year, y=annual\_debt\_increase\_pct, color=party),) + stat\_smooth(method="lm", se=TRUE)



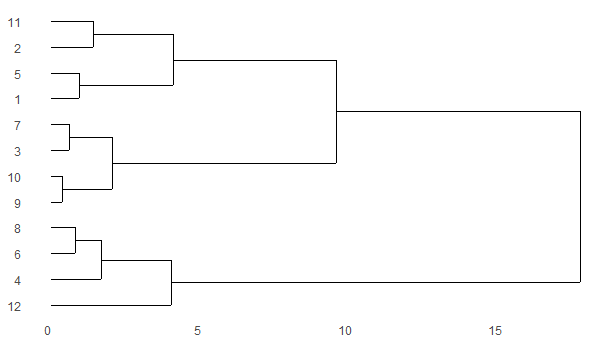
A linear regression of the percentage increases over the years, (the years are integers in the dataset), shows a trend oscillating between the two major parties. Under the republican presidents, the spending increases are being led by them with their democratic successors matching but not necessarily increasing the percentages. The plot points of the democratic president’s post 1980 have very sharp trends of declining percentages. The regression model shows that there is no unified trend with the postwar presidents. The different decisions in how their government choose to spend tax payer money shows in the oscillations in the plot. The real correlation with the data is split by the party. The graph shows that the debt has pronounced increases in years with the country being run by Republican presidents with the Democratic president’s sharp decline, likely in response to the spending of their political rivals. The linear regression supports the oscillating nature of politics that acts like a pendulum with the real variation coming from parties and not the individual presidents.

> model3<-aggregate(annual\_debt\_increase\_pct~president,presidential\_data,mean)

> model3.1<-hclust(dist(model3), method="ward.D2")

> ggdendrogram(model3.1, rotate=TRUE)

|  |
| --- |
|  |
| **1** | Barack Obama | 8.325000 |
| **2** | Bill Clinton | 4.287500 |
| **3** | Dwight D. Eisenhower | 1.262500 |
| **4** | George H. W. Bush | 11.825000 |
| **5** | George W. Bush | 7.387500 |
| **6** | Gerald Ford | 10.766667 |
| **7** | Harry S. Truman | 0.675000 |
| **8** | Jimmy Carter | 9.975000 |
| **9** | John F. Kennedy | 2.233333 |
| **10** | Lyndon B. Johnson | 2.600000 |
| **11** | Richard Nixon | 5.700000 |
| **12** | Ronald Reagan | 14.125000 |



The hierarchal clustering method used in this dendrogram is the ward method. The algorithm in which the presidents are clustered together sorts them along the lines of closest neighbor in terms of distance to each other. Once a pair is made then the cluster builds up and creates in the next connection. An interesting observation is that in the first round of the dendrogram, five clusters are made and for four of them, the presidents that connected served back to back to each other. Also, Presidents Truman, Eisenhower, Kennedy and Johnson all served back to back from 1945-1969, these fours all clustered together showing minimal variance during the spending of their administrations