

✓ A Prediction Analysis on the New 2024 Election

09/15/2024

With the 2024 election approaching, both the U.S. and the global community are eagerly anticipating who will become the next President of the United States. I plan to use data to forecast the outcome of this election.

Currently, the race is between two prominent candidates from the major political parties: Donald Trump, the 45th U.S. President and Republican candidate, and Kamala Harris, the current Vice President and Democratic candidate. My goal is to analyze available data to predict which of these candidates is likely to win.

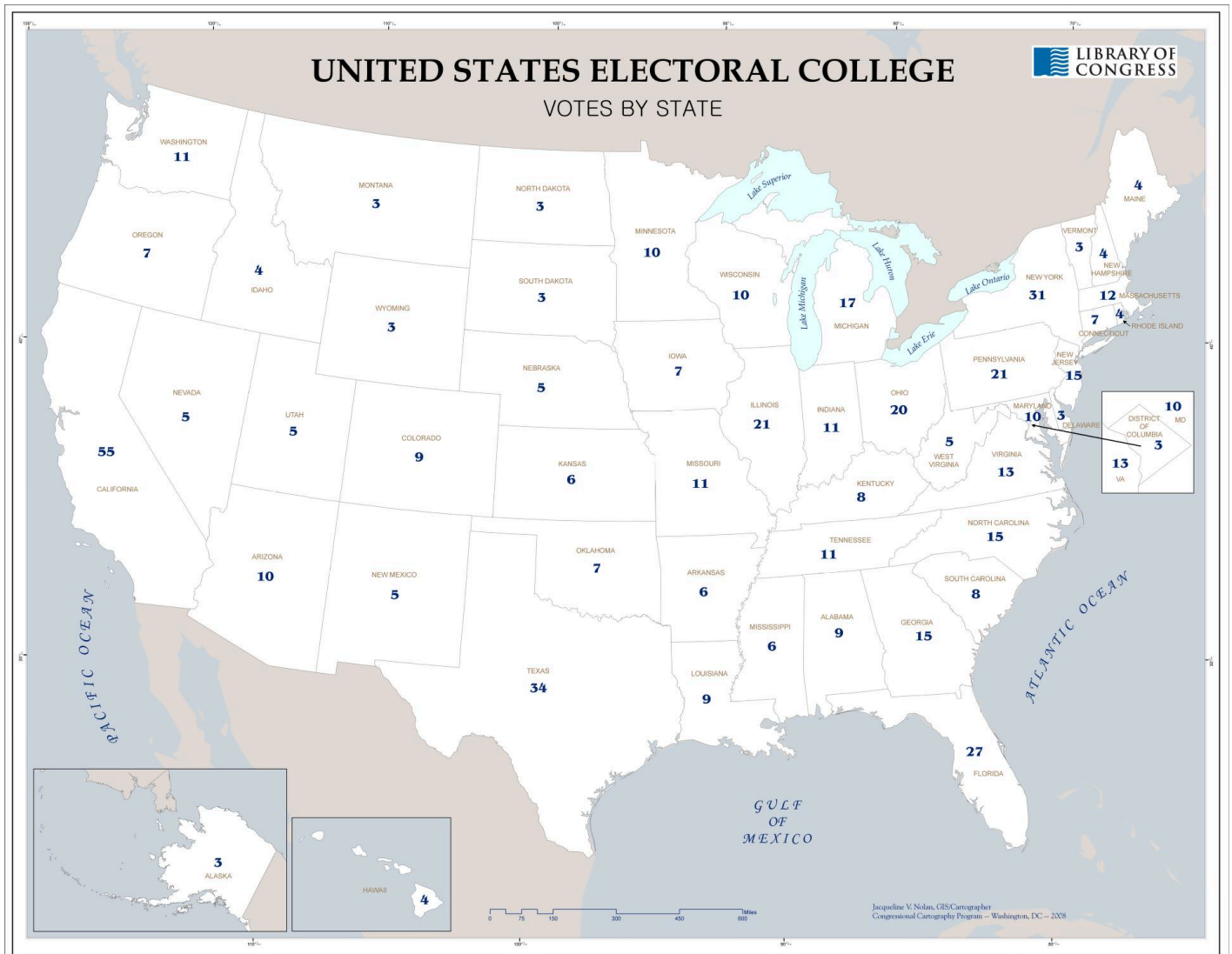
✓ Electoral College

In the U.S., the Electoral College is responsible for electing the President and Vice President. This process involves selecting electors who then cast votes for the two offices. After the electors cast their votes, Congress is responsible for counting them. Unlike other elections where candidates are elected directly by popular vote, the President and Vice President are elected through this Electoral College system.

Each state has a number of electors equal to its total number of Congressional representatives, which includes both House and Senate members. Washington, D.C., also has three electors, bringing the total to 538 electors. Political parties in each state select their own slate of potential electors, and the method for choosing electors varies by state.

How does the Electoral College process work?

1. After we cast your ballot for president, our vote goes to a statewide tally. In 48 states and Washinton, D.C., the winner will get all the electoral votes for that state. Other states like Maine and Nebraska assign their electors using a proportional system.
2. A presidential candidate needs the vote of at least 270 electors which is more than half of all electors to win the presidential election.
3. In most cases, a projected winner is announced on election night in November after you vote. However, the actual Electoral College vote takes place in mid-December when the electors meet in their states

Figure 1. Map of the United States Electoral College Votes By State

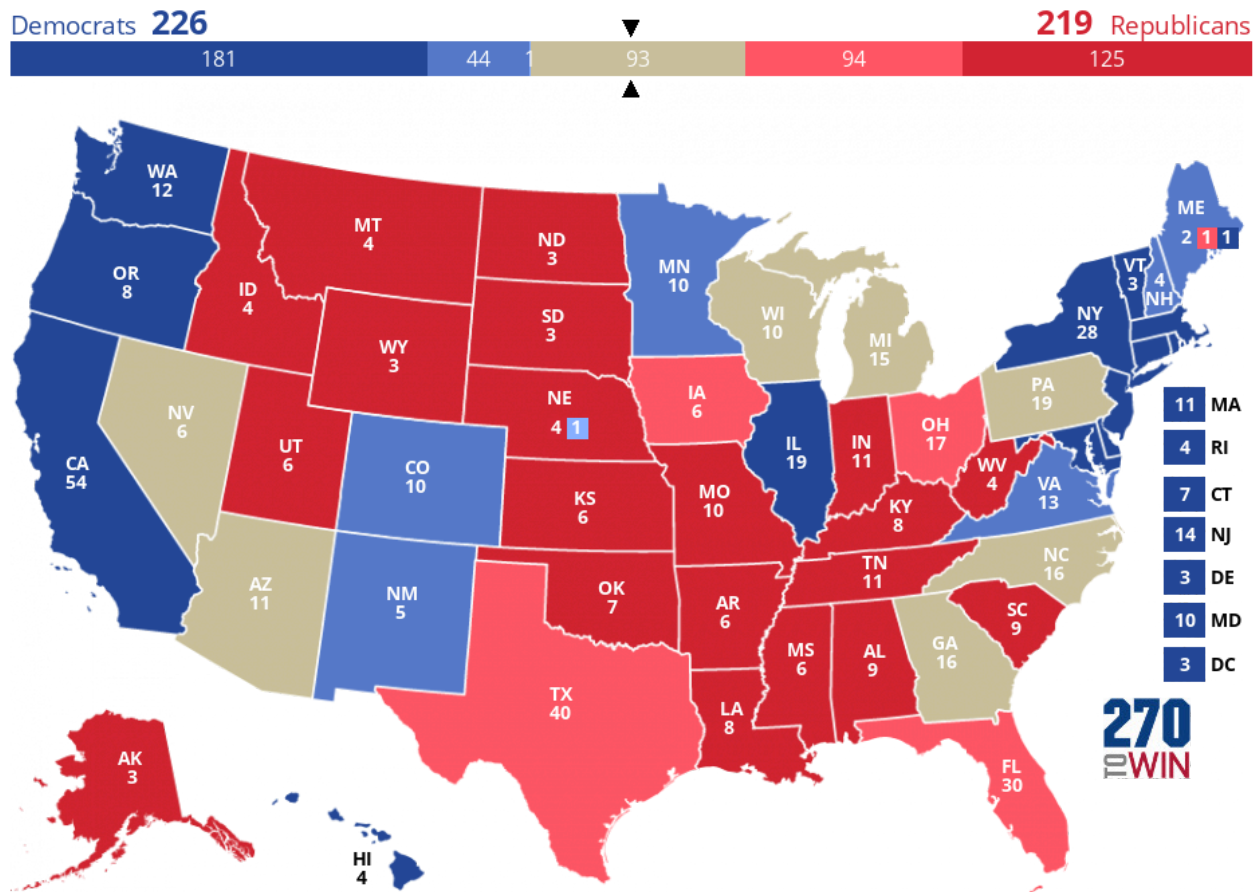
This figure depicts how many Electoral College votes each states has. Each state has a different number of Electoral College votes.

✓ Interpretation

The Electoral College will essential determine the future presidential candidate and in order for a presidential candidate to win, they will need to get atleast 270 Electoral College votes.

We will now look at some current polls to see how the president candidates are doing. We will see which Electoral State votes they gotten so far and also see the swing states(Electorial State where it's a close tie).

Figure 2. Map of the United States Electoral College Votes where the States are more Harris leaning or Trump leading or is a swing state.



This figure depicts how many Electoral College votes Harris and Trump has gotten. Currently we see that Harris has gotten 223 votes and Trump has gotten 188 votes, with 127 votes still available from swing states.

Note: The votes will constantly change.

Methodology

Based on my interpretation, the outcome of the election could still go either way due to the way the Electoral College functions. Note that presidential candidates do not need to win all the swing states to reach 270 electoral votes. Additionally, the speed of vote counting matters because once a candidate reaches or surpasses 270 electoral votes through various combinations, the remaining votes in other states may no longer impact the outcome.

My methodology involves tracking the number of Electoral College votes already leaning toward either Harris or Trump and focusing on the swing states, as they will ultimately determine the election.


✓ Gathering Data

Source: <https://www.statista.com/statistics/1035442/electoral-votes-republican-democratic-parties-since-1828/>

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import math
from google.colab import drive

drive.mount("/content/gdrive", force_remount=True)

df = pd.read_csv('/content/gdrive/MyDrive/pastElection.csv') #Gets data from CSV fil
df
```

 Mounted at /content/gdrive

| | Characteristic | Democrat | Republican |
|----|----------------|----------|------------|
| 0 | 2020 | 56.88% | 43.12% |
| 1 | 2016 | 42.19% | 56.69% |
| 2 | 2012 | 61.71% | 38.29% |
| 3 | 2008 | 67.84% | 32.16% |
| 4 | 2004 | 46.66% | 53.16% |
| 5 | 2000 | 49.44% | 50.37% |
| 6 | 1996 | 70.45% | 29.55% |
| 7 | 1992 | 68.77% | 31.23% |
| 8 | 1988 | 20.63% | 79.18% |
| 9 | 1984 | 2.42% | 97.58% |
| 10 | 1980 | 9.11% | 90.89% |
| 11 | 1976 | 55.20% | 44.61% |
| 12 | 1972 | 3.16% | 96.65% |
| 13 | 1968 | 35.50% | 55.95% |
| 14 | 1964 | 90.33% | 9.67% |
| 15 | 1960 | 56.32% | 40.78% |
| 16 | 1956 | 13.75% | 84.18% |
| 17 | 1952 | 16.76% | 83.24% |
| 18 | 1948 | 57.06% | 35.59% |
| 19 | 1944 | 81.36% | 18.64% |
| 20 | 1940 | 84.56% | 15.44% |
| 21 | 1936 | 98.49% | 1.51% |
| 22 | 1932 | 88.89% | 11.11% |
| 23 | 1928 | 16.38% | 83.62% |
| 24 | 1924 | 25.61% | 71.94% |
| 25 | 1920 | 23.92% | 76.08% |
| 26 | 1916 | 52.16% | 47.83% |
| 27 | 1912* | 81.92% | 1.51% |
| 28 | 1908 | 33.54% | 66.46% |
| 29 | 1904 | 24.68% | 65.22% |

| | | | |
|-----------|----------|--------|--------|
| 29 | 1904 | 34.68% | 65.32% |
| 30 | 1900 | 34.68% | 65.32% |
| 31 | 1896 | 39.37% | 60.63% |
| 32 | 1892 | 62.39% | 32.66% |
| 33 | 1888 | 41.90% | 58.10% |
| 34 | 1884 | 54.64% | 45.39% |
| 35 | 1880 | 42.01% | 57.99% |
| 36 | 1876 | 49.86% | 50.14% |
| 37 | 1872** | 17.61% | 81.25% |
| 38 | 1868 | 27.21% | 72.79% |
| 39 | 1864*** | 9.01% | 90.99% |
| 40 | 1860**** | 27.72% | 59.40% |

```
df['Democrat'] = df['Democrat'].str.replace('%', '').astype(float)
df['Republican'] = df['Republican'].str.replace('%', '').astype(float)
df
```



| | Characteristic | Democrat | Republican |
|----|----------------|----------|------------|
| 0 | 2020 | 56.88 | 43.12 |
| 1 | 2016 | 42.19 | 56.69 |
| 2 | 2012 | 61.71 | 38.29 |
| 3 | 2008 | 67.84 | 32.16 |
| 4 | 2004 | 46.66 | 53.16 |
| 5 | 2000 | 49.44 | 50.37 |
| 6 | 1996 | 70.45 | 29.55 |
| 7 | 1992 | 68.77 | 31.23 |
| 8 | 1988 | 20.63 | 79.18 |
| 9 | 1984 | 2.42 | 97.58 |
| 10 | 1980 | 9.11 | 90.89 |
| 11 | 1976 | 55.20 | 44.61 |
| 12 | 1972 | 3.16 | 96.65 |
| 13 | 1968 | 35.50 | 55.95 |
| 14 | 1964 | 90.33 | 9.67 |
| 15 | 1960 | 56.32 | 40.78 |
| 16 | 1956 | 13.75 | 84.18 |
| 17 | 1952 | 16.76 | 83.24 |
| 18 | 1948 | 57.06 | 35.59 |
| 19 | 1944 | 81.36 | 18.64 |
| 20 | 1940 | 84.56 | 15.44 |
| 21 | 1936 | 98.49 | 1.51 |
| 22 | 1932 | 88.89 | 11.11 |
| 23 | 1928 | 16.38 | 83.62 |
| 24 | 1924 | 25.61 | 71.94 |
| 25 | 1920 | 23.92 | 76.08 |
| 26 | 1916 | 52.16 | 47.83 |
| 27 | 1912* | 81.92 | 1.51 |
| 28 | 1908 | 33.54 | 66.46 |
| 29 | 1904 | 34.68 | 65.32 |

| | | | |
|-----------|----------|-------|-------|
| 30 | 1900 | 34.68 | 65.32 |
| 31 | 1896 | 39.37 | 60.63 |
| 32 | 1892 | 62.39 | 32.66 |
| 33 | 1888 | 41.90 | 58.10 |
| 34 | 1884 | 54.64 | 45.39 |
| 35 | 1880 | 42.01 | 57.99 |
| 36 | 1876 | 49.86 | 50.14 |
| 37 | 1872** | 17.61 | 81.25 |
| 38 | 1868 | 27.21 | 72.79 |
| 39 | 1864*** | 9.01 | 90.99 |
| 40 | 1860**** | 27.72 | 59.40 |

```
df['Characteristic'] = df['Characteristic'].str.replace('*', '').astype(int)
df
```




| | Characteristic | Democrat | Republican |
|----|----------------|----------|------------|
| 0 | 2020 | 56.88 | 43.12 |
| 1 | 2016 | 42.19 | 56.69 |
| 2 | 2012 | 61.71 | 38.29 |
| 3 | 2008 | 67.84 | 32.16 |
| 4 | 2004 | 46.66 | 53.16 |
| 5 | 2000 | 49.44 | 50.37 |
| 6 | 1996 | 70.45 | 29.55 |
| 7 | 1992 | 68.77 | 31.23 |
| 8 | 1988 | 20.63 | 79.18 |
| 9 | 1984 | 2.42 | 97.58 |
| 10 | 1980 | 9.11 | 90.89 |
| 11 | 1976 | 55.20 | 44.61 |
| 12 | 1972 | 3.16 | 96.65 |
| 13 | 1968 | 35.50 | 55.95 |
| 14 | 1964 | 90.33 | 9.67 |
| 15 | 1960 | 56.32 | 40.78 |
| 16 | 1956 | 13.75 | 84.18 |
| 17 | 1952 | 16.76 | 83.24 |
| 18 | 1948 | 57.06 | 35.59 |
| 19 | 1944 | 81.36 | 18.64 |
| 20 | 1940 | 84.56 | 15.44 |
| 21 | 1936 | 98.49 | 1.51 |
| 22 | 1932 | 88.89 | 11.11 |
| 23 | 1928 | 16.38 | 83.62 |
| 24 | 1924 | 25.61 | 71.94 |
| 25 | 1920 | 23.92 | 76.08 |
| 26 | 1916 | 52.16 | 47.83 |
| 27 | 1912 | 81.92 | 1.51 |
| 28 | 1908 | 33.54 | 66.46 |
| 29 | 1904 | 34.68 | 65.32 |

| | | | |
|-----------|------|-------|-------|
| 30 | 1900 | 34.68 | 65.32 |
| 31 | 1896 | 39.37 | 60.63 |
| 32 | 1892 | 62.39 | 32.66 |
| 33 | 1888 | 41.90 | 58.10 |
| 34 | 1884 | 54.64 | 45.39 |
| 35 | 1880 | 42.01 | 57.99 |
| 36 | 1876 | 49.86 | 50.14 |
| 37 | 1872 | 17.61 | 81.25 |
| 38 | 1868 | 27.21 | 72.79 |
| 39 | 1864 | 9.01 | 90.99 |
| 40 | 1860 | 27.72 | 59.40 |

```
years = df['Characteristic']
democrat_votes = df['Democrat']
republican_votes = df['Republican']
```

```
# Number of bars
n = len(years)
```

```
# Position of bars on x-axis
bar_width = 0.4
index = np.arange(n)
```

```
plt.figure(figsize=(30, 20))
```

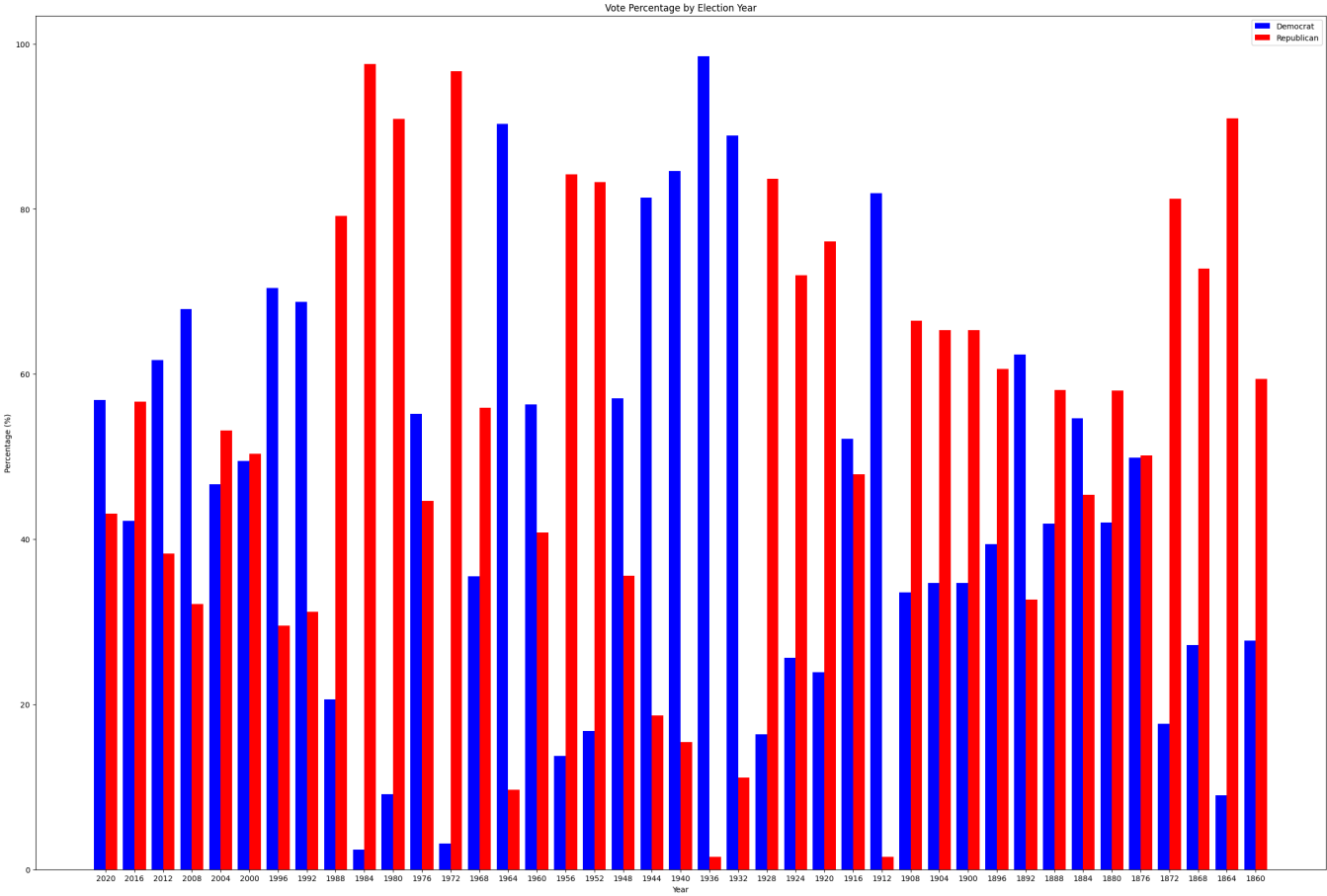
```
# Plot bars for Democrats
plt.bar(index - bar_width / 2, democrat_votes, bar_width, label='Democrat', color='b')
```

```
# Plot bars for Republicans
plt.bar(index + bar_width / 2, republican_votes, bar_width, label='Republican', color='r')
```

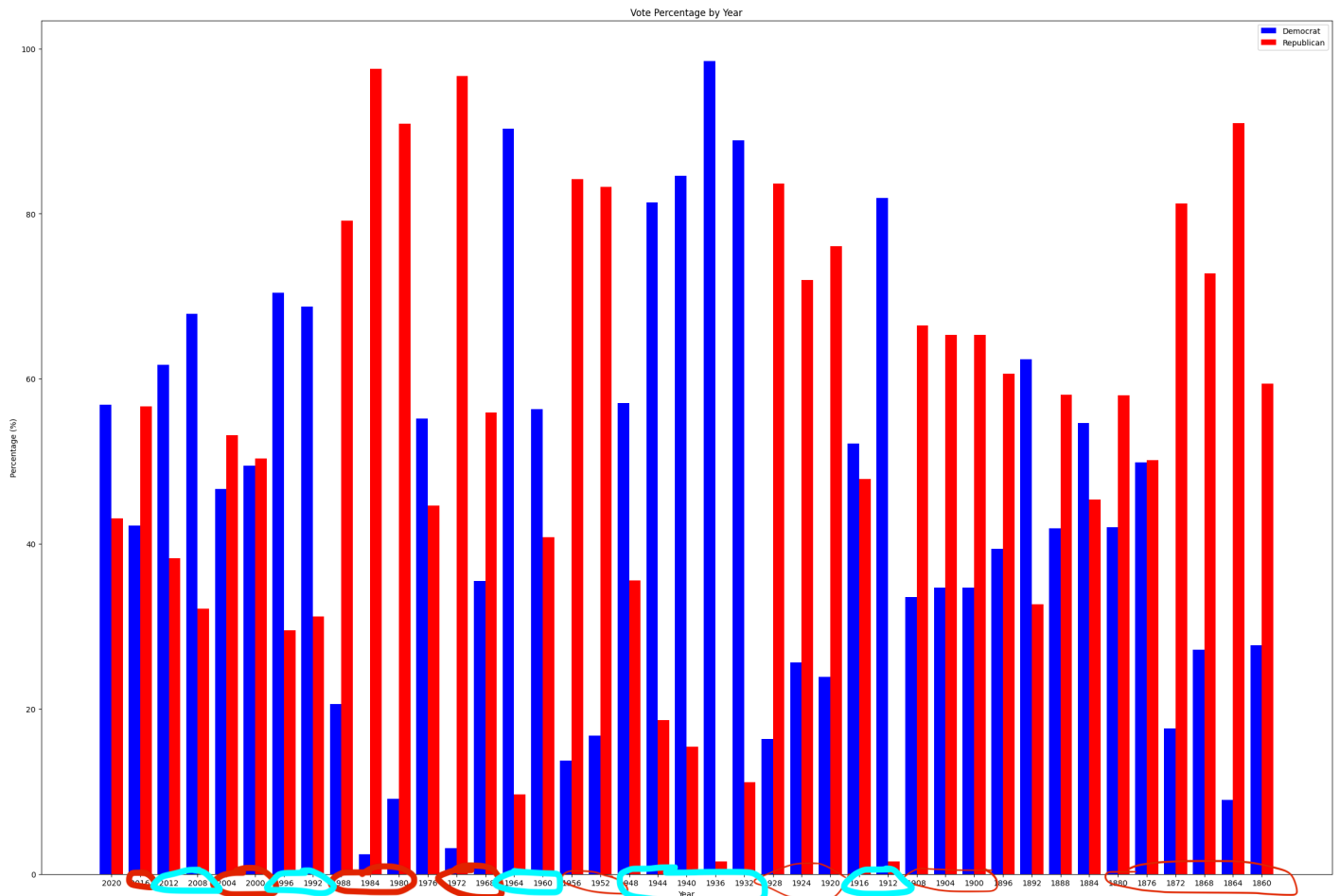
```
# Add labels and title
plt.xlabel('Year')
plt.ylabel('Percentage (%)')
plt.title('Vote Percentage by Election Year')
plt.xticks(index, years) # Set the position and labels for the x-ticks
```

```
# Add a legend
plt.legend()
```

```
# Show the plot
plt.show()
```



Throughout U.S. history, there have been many instances of both Democrats and Republicans winning. There have been periods when Democrats won consistently, as well as periods when Republicans did.



Based on this graph, we can see that Democratic and Republican presidents have frequently alternated over time. In the past, there were more Republican presidents, whereas more recently, there have been more Democratic presidents. Additionally, notice how the percentage of votes is less one-sided now. Historically, a president from one party would receive the majority of the Electoral College votes, but now the candidates from both parties tend to receive similar amounts of votes, making the competition more interesting and competitive.

Given this information, I believe that voters are becoming more polarized and more united simultaneously.

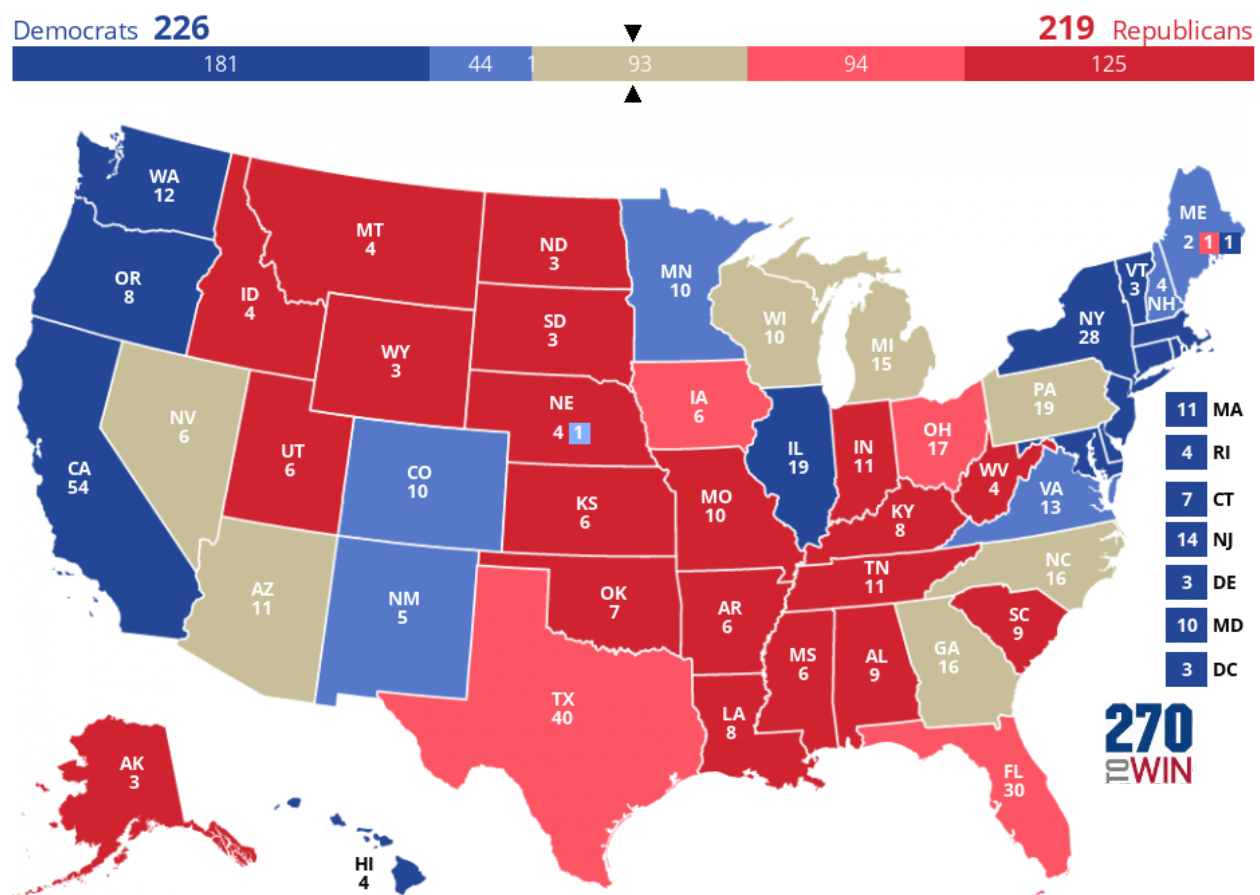
Based on this graph, I also predict that future elections will continue to have closer results, with Democrats and Republicans receiving similar amounts of Electoral College votes. Future

presidential elections will likely see a switch between Democratic and Republican candidates approximately every 1 to 2 elections.

With U.S. citizens becoming more polarized yet simultaneously more united, it suggests that diverse values and beliefs are no longer adequately represented by just two major political parties. This could indicate that it might be time to introduce more political parties to better reflect the true values of the American people.

Based on this graph, I cannot say with absolute certainty whether Harris or Trump would win.

Now refer back to **figure 2**, notice that there are a total of 7 states that is currently colored gray. This means that there the swing states and notice how Harris nor Trump has gotten 270 votes yet in the polls as of now.



This figure depicts how many Electoral College votes Harris and Trump has gotten. Currently we see that Harris has gotten 226 votes and Trump has gotten 219 votes, with 127 votes still available from swing states.

Note: The votes will constantly change.

```
df2 = pd.read_csv('/content/gdrive/MyDrive/swingStates.csv') #Gets data from CSV fil
df2
```



| | Characteristic | Harris | Trump |
|----|----------------|--------|--------|
| 0 | Arizona | 46.80% | 48.40% |
| 1 | Georgia | 48% | 48.30% |
| 2 | Pennsylvania | 47.60% | 47.60% |
| 3 | Nevada | 48% | 47.40% |
| 4 | North Carolina | 47.80% | 48.90% |
| 5 | Michigan | 48.30% | 47.10% |
| 6 | Wisconsin | 49.10% | 47.30% |
| 7 | NaN | NaN | NaN |
| 8 | NaN | NaN | NaN |
| 9 | NaN | NaN | NaN |
| 10 | NaN | NaN | NaN |
| 11 | NaN | NaN | NaN |
| 12 | NaN | NaN | NaN |
| 13 | NaN | NaN | NaN |
| 14 | NaN | NaN | NaN |
| 15 | NaN | NaN | NaN |
| 16 | NaN | NaN | NaN |
| 17 | NaN | NaN | NaN |
| 18 | NaN | NaN | NaN |
| 19 | NaN | NaN | NaN |
| 20 | NaN | NaN | NaN |
| 21 | NaN | NaN | NaN |
| 22 | NaN | NaN | NaN |
| 23 | NaN | NaN | NaN |
| 24 | NaN | NaN | NaN |
| 25 | NaN | NaN | NaN |
| 26 | NaN | NaN | NaN |
| 27 | NaN | NaN | NaN |
| 28 | NaN | NaN | NaN |
| 29 | NaN | NaN | NaN |


| | | | |
|----|-----|-----|-----|
| 30 | NaN | NaN | NaN |
| 31 | NaN | NaN | NaN |
| 32 | NaN | NaN | NaN |
| 33 | NaN | NaN | NaN |
| 34 | NaN | NaN | NaN |
| 35 | NaN | NaN | NaN |
| 36 | NaN | NaN | NaN |
| 37 | NaN | NaN | NaN |
| 38 | NaN | NaN | NaN |
| 39 | NaN | NaN | NaN |
| 40 | NaN | NaN | NaN |

Source: <https://www.statista.com/statistics/1428865/general-election-swing-state-polling-biden-trump-us/>

```
print(df2.shape) #Checks the number of rows and columns in dataframe. The dataframe
df2= df2.dropna() #Removes the rows that contains NULL values.
print(df2.shape)
```

```
⇒ (41, 3)
   (7, 3)
```

```
df2['Harris'] = df2['Harris'].str.replace('%', '').astype(float)
df2['Trump'] = df2['Trump'].str.replace('%', '').astype(float)
df2
```




```
<ipython-input-7-a7283e59fe80>:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <https://pandas.pydata.org/pandas-docs/stab>
df2['Harris'] = df2['Harris'].str.replace('%', '').astype(float)
<ipython-input-7-a7283e59fe80>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: <https://pandas.pydata.org/pandas-docs/stab>
df2['Trump'] = df2['Trump'].str.replace('%', '').astype(float)

| | Characteristic | Harris | Trump |
|---|----------------|--------|-------|
| 0 | Arizona | 46.8 | 48.4 |
| 1 | Georgia | 48.0 | 48.3 |
| 2 | Pennsylvania | 47.6 | 47.6 |
| 3 | Nevada | 48.0 | 47.4 |
| 4 | North Carolina | 47.8 | 48.9 |
| 5 | Michigan | 48.3 | 47.1 |
| 6 | Wisconsin | 49.1 | 47.3 |

df2



| | Characteristic | Harris | Trump |
|---|----------------|--------|-------|
| 0 | Arizona | 46.8 | 48.4 |
| 1 | Georgia | 48.0 | 48.3 |
| 2 | Pennsylvania | 47.6 | 47.6 |
| 3 | Nevada | 48.0 | 47.4 |
| 4 | North Carolina | 47.8 | 48.9 |
| 5 | Michigan | 48.3 | 47.1 |
| 6 | Wisconsin | 49.1 | 47.3 |

```
# Define new data for additional columns  
votes = [11, 16, 19, 6, 16, 15, 10]  
  
new_data = {'Votes': votes}  
df2 = df2.assign(**new_data)
```

df2