The U.S. Election throughout History - Process Book

Duanduan Liu
Kanrong Yu
Weina Chen

Overview and Motivation

As the eye-catching 2016 US presidential eventually settled, there had been so many politically motivated visualizations like those provided by the famous FiveThirtyEight and NYTimes, and the public expectation towards the final result was somehow affected or even misled by them. Therefore, what prompts us to do this project is to create an unbiased data-driven visualization, which won't mislead people's interpretation from viz. Moreover, from this year's presidential result, we should admit election results are generally hard to predict despite many different polls made on large group of samples. However, we consider it still interesting and meaningful to look at the whole idea of the evolution of US presidential election results and investigate if there is correlation between the results and some other factors.

To be more specific, we aim to obtain a general picture of the US election through demographic of election in the following aspects:

- Discover presidential shift from 1920 to 2012.
- Detect how the states stacked up in every election and how they have swung during the time frame. (In progress)
- Expose the relative relations with the preference of newspaper endorsements e.g. NYT, the external environment like recession or recovered economy, and the trend of growth of the executive branch e.g. Federal Budget Receipts & Outlays, and GDP.

Related Work

- There Are Many Ways to Map Election Results. We've Tried Most of Them. - by NYTimes
  http://www.nytimes.com/interactive/2016/11/01/upshot/many-ways-to-map-election-results.html?_r=0
- LIVE RESULTS AND MAPS Election Results 2016
  http://graphics.wsj.com/elections/2016/results/
- The various Live US election results we watched in class

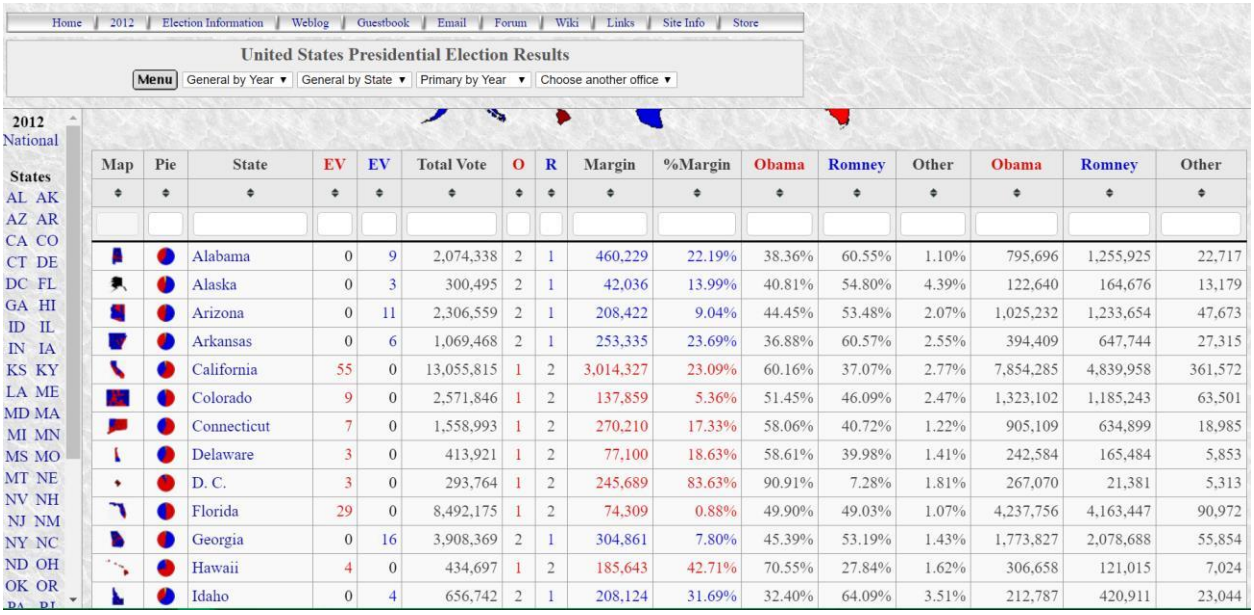
Data

1. Data Source

The main reference is from the Library of Congress, a post named "U.S. Election Statistics: A

Resource Guide" (https://www.loc.gov/rr/program/bib/elections/statistics.html). This resource guide compiles a list of online and print resources that contain U.S. election statistics for both federal and state elections. To be more specific, we mainly use three online resources listed in this guide:

- Dave Leip's Atlas of U.S. Presidential Elections, which is our main data source that includes the voting statistics for every election by state level.
http://uselectionatlas.org/



Home | 2012 | Election Information | Weblog | Guestbook | Email | Forum | Wiki | Links | Site Info | Store

United States Presidential Election Results

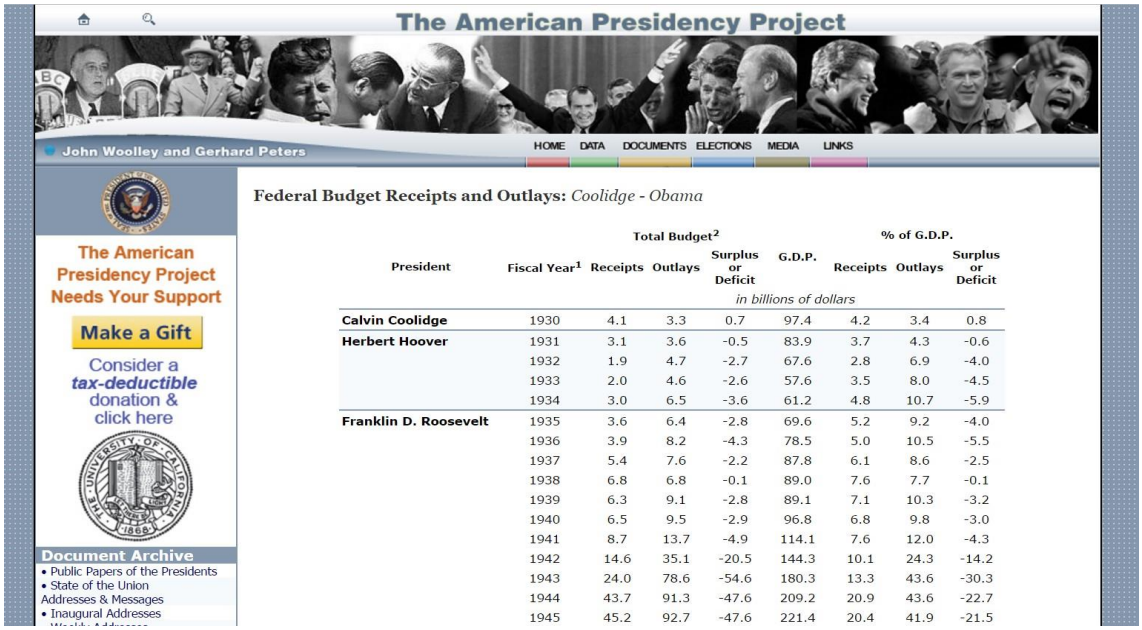
Menu | General by Year | General by State | Primary by Year | Choose another office

2012 National

States: AL AK AZ AR CA CO CT DE DC FL GA HI ID IL IN IA KS KY LA ME MD MA MI MN MS MO MT NE NV NH NJ NM NY NC ND OH OK OR

Map	Pie	State	EV	EV	Total Vote	O	R	Margin	%Margin	Obama	Romney	Other	Obama	Romney	Other
		Alabama	0	9	2,074,338	2	1	460,229	22.19%	38.36%	60.55%	1.10%	795,696	1,255,925	22,717
		Alaska	0	3	300,495	2	1	42,036	13.99%	40.81%	54.80%	4.39%	122,640	164,676	13,179
		Arizona	0	11	2,306,559	2	1	208,422	9.04%	44.45%	53.48%	2.07%	1,025,232	1,233,654	47,673
		Arkansas	0	6	1,069,468	2	1	253,335	23.69%	36.88%	60.57%	2.55%	394,409	647,744	27,315
		California	55	0	13,055,815	1	2	3,014,327	23.09%	60.16%	37.07%	2.77%	7,854,285	4,839,958	361,572
		Colorado	9	0	2,571,846	1	2	137,859	5.36%	51.45%	46.09%	2.47%	1,323,102	1,185,243	63,501
		Connecticut	7	0	1,558,993	1	2	270,210	17.33%	58.06%	40.72%	1.22%	905,109	634,899	18,985
		Delaware	3	0	413,921	1	2	77,100	18.63%	58.61%	39.98%	1.41%	242,584	165,484	5,853
		D. C.	3	0	293,764	1	2	245,689	83.63%	90.91%	7.28%	1.81%	267,070	21,381	5,313
		Florida	29	0	8,492,175	1	2	74,309	0.88%	49.90%	49.03%	1.07%	4,237,756	4,163,447	90,972
		Georgia	0	16	3,908,369	2	1	304,861	7.80%	45.39%	53.19%	1.43%	1,773,827	2,078,688	55,854
		Hawaii	4	0	434,697	1	2	185,643	42.71%	70.55%	27.84%	1.62%	306,658	121,015	7,024
		Idaho	0	4	656,742	2	1	208,124	31.69%	32.40%	64.09%	3.51%	212,787	420,911	23,044

- American Presidency Project: Presidential Election Data, from which we grabbed the financial data, GDP http://www.presidency.ucsb.edu/elections.php



The American Presidency Project

John Woolley and Gerhard Peters

HOME DATA DOCUMENTS ELECTIONS MEDIA LINKS

The American Presidency Project Needs Your Support

Make a Gift

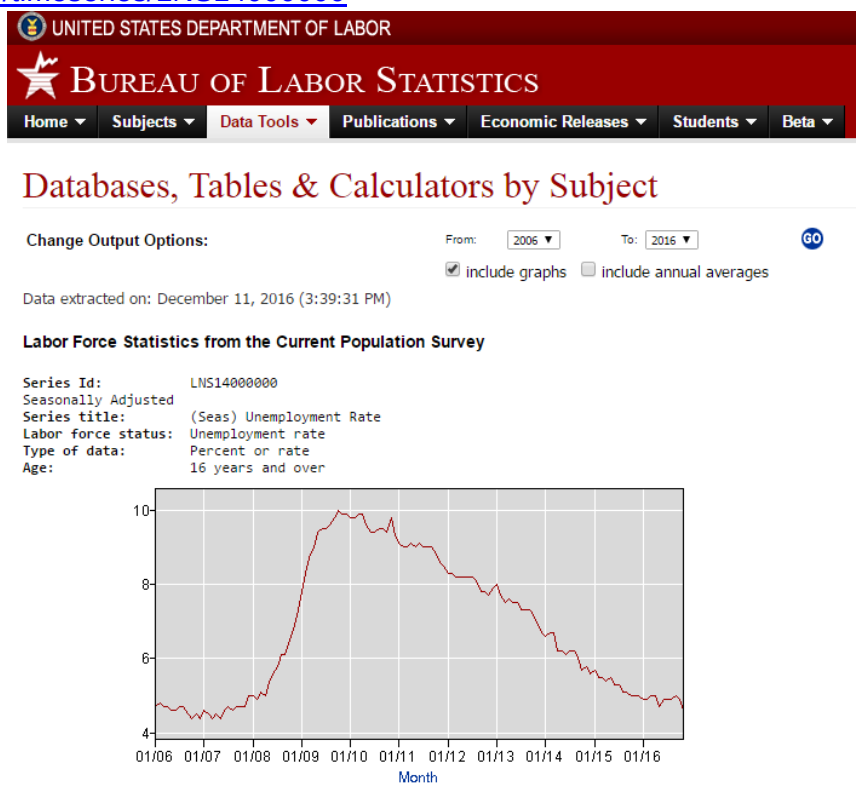
Consider a *tax-deductible* donation & click here

Document Archive
- Public Papers of the Presidents
- State of the Union Addresses & Messages
- Inaugural Addresses

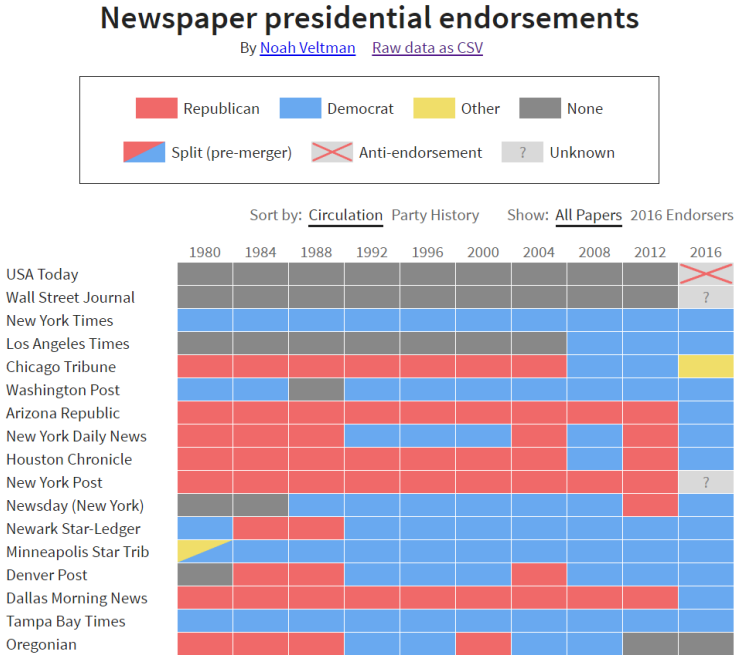
Federal Budget Receipts and Outlays: *Coolidge - Obama*

President	Fiscal Year[1]	Total Budget[2] Receipts	Outlays	Surplus or Deficit	G.D.P.	% of G.D.P. Receipts	Outlays	Surplus or Deficit
		in billions of dollars						
Calvin Coolidge	1930	4.1	3.3	0.7	97.4	4.2	3.4	0.8
Herbert Hoover	1931	3.1	3.6	-0.5	83.9	3.7	4.3	-0.6
	1932	1.9	4.7	-2.7	67.6	2.8	6.9	-4.0
	1933	2.0	4.6	-2.6	57.6	3.5	8.0	-4.5
	1934	3.0	6.5	-3.6	61.2	4.8	10.7	-5.9
Franklin D. Roosevelt	1935	3.6	6.4	-2.8	69.6	5.2	9.2	-4.0
	1936	3.9	8.2	-4.3	78.5	5.0	10.5	-5.5
	1937	5.4	7.6	-2.2	87.8	6.1	8.6	-2.5
	1938	6.8	6.8	-0.1	89.0	7.6	7.7	-0.1
	1939	6.3	9.1	-2.8	89.1	7.1	10.3	-3.2
	1940	6.5	9.5	-2.9	96.8	6.8	9.8	-3.0
	1941	8.7	13.7	-4.9	114.1	7.6	12.0	-4.3
	1942	14.6	35.1	-20.5	144.3	10.1	24.3	-14.2
	1943	24.0	78.6	-54.6	180.3	13.3	43.6	-30.3
	1944	43.7	91.3	-47.6	209.2	20.9	43.6	-22.7
	1945	45.2	92.7	-47.6	221.4	20.4	41.9	-21.5

- Labor Force Statistics from the Current Population Survey, from which we obtained the unemployment rate data
http://data.bls.gov/timeseries/LNS14000000



- Newspaper presidential endorsements, where we excerpted the endorsement of top 10 US newspapers (by circulation)
http://noahveltman.com/endorsements/



2. Data Processing

For Data Cleaning,

1) To keep the column consistency in “each year.csv”, we uses Python to manualLy add EV_I, I_Percen, I_Actual with “0” value when there’s only two parties (Democratic and Republican) at that year. By doing this, we got same number of columns in each file and could easily import relevant columns from the csv files using JavaScript.
2) Because of the admission of Alaska and Hawaii as states in 1959, the presidential election data before 1960 didn’t have their information and similar situation goes with Washington DC, we added new rows with “0” values for these states in order to keep the row consistency for every year data file.
3) Also, we deleted % for the D_Percen, R_Percen and I_Percen columns and delete the thousand separators.

To avoid too much calculation in later JavaScript code, we did the processing work in Python in advance. Including:

1) Add a RD_Diff column as R_Percen - D_Percen
2) Put the state abbreviation name in State_Abbr
3) Based on RD_Diff, attach relevant color HEX towards it (in ten levels, from deep red to deep blue
4) Besides the HEX, color rank is assigned (for sake of the sorting of state blocks in the stacked bar chart)
5) Recalculate Other_Percen as 1 - D_Percen - R_Percen - I_Percen as in the raw data, and these 4 columns add up to more than 1.

```
# Sort by index
df.sort_index(inplace=True)

# Delete the final Row
df = df[df['State'] != "Total"]

# Add RD_Diff
df = add_RD_Diff(df)

# Add State_Abbr
df = add_State_Abbr(df)

# Add Color
df = add_color(df)

# Re caculate Other_Percen
df = tranform_Other_Percen(df, year)

# Add Color Rank
df = add_color_rank(df)

# Save df to CSV
folder_name = "Data/Cleaning/CSV/" + str(year) + ".csv"
df.to_csv(folder_name, index=False)
```

6) To make the state electoral shift chart, df_now["Shift"] = (-df_now["D_Percen"] +

df_last["D_Percen"]) + (df_now["R_Percen"] - df_last["R_Percen"]) is added.

Here's the final look of each year.csv.

State	EV_D	EV_R	Total_Vote	D_Percen	R_Percen	Other_Percer	D_Actual	R_Actual	Other_Actual	I_Actual	EV_I	I_Percen	RD_Diff	State_Abbr	Color	colorRank	Last_Color	Shift
Alabama	0	9	2074338	38.36	60.55	1.09	795696	1255925	22717	0	0	0	22.19	AL	#fcae91	2	#fcae91	0.61
Alaska	0	3	300495	40.81	54.8	4.39	122640	164676	13179	0	0	0	13.99	AK	#fee5d9	1	#fcae91	-7.54
Arizona	0	11	2306559	44.45	53.48	2.07	1025232	1233654	47673	0	0	0	9.03	AZ	#fee5d9	1	#fee5d9	0.55
Arkansas	0	6	1069468	36.88	60.57	2.55	394409	647744	27315	0	0	0	23.69	AR	#fcae91	2	#fee5d9	3.83
California	55	0	13055815	60.16	37.07	2.77	7854285	4839958	361572	0	0	0	-23.09	CA	#6baed6	-3	#6baed6	0.93
Colorado	9	0	2571846	51.45	46.09	2.46	1323102	1185243	63501	0	0	0	-5.36	CO	#eff3ff	-1	#eff3ff	3.59
Connecticut	7	0	1558993	58.06	40.72	1.22	905109	634899	18985	0	0	0	-17.34	CT	#bdd7e7	-2	#6baed6	5.03
Delaware	3	0	413921	58.61	39.98	1.41	242584	165484	5853	0	0	0	-18.63	DE	#bdd7e7	-2	#6baed6	6.35
D. C.	3	0	293764	90.91	7.28	1.81	267070	21381	5313	0	0	0	-83.63	DC	#08519c	-5	#08519c	2.3

Apart from 29 election year state level data files, we also generated a total file, with state level electoral statistics, media endorsement, financial data and background information.

Year	Color	D_Percen	R_Percen	EV_Total	USA Today	Wall Street J	New York Tir	Los Angeles	Chicago Trib	Washington	Arizona Repu	New York Da	Houston Chr	New York Po	D_EV	R_EV	I_EV	Keyword
1976	#08519c	50.08	48.01	537			#08519c								297	240	0	Watergate scand
1980	#a50f15	41.01	50.75	538	#888888	#888888	#08519c	#888888	#a50f15	#08519c	#a50f15	#a50f15	#a50f15	#a50f15	49	489	0	Iran hostage crisi
1984	#a50f15	40.56	58.77	538	#888888	#888888	#08519c	#888888	#a50f15	#08519c	#a50f15	#a50f15	#a50f15	#a50f15	13	525	0	Strong economic
1988	#a50f15	45.65	53.37	537	#888888	#888888	#08519c	#888888	#a50f15	#888888	#a50f15	#a50f15	#a50f15	#a50f15	111	426	0	Good economy, S
1992	#08519c	43.01	37.45	538	#888888	#888888	#08519c	#888888	#a50f15	#08519c	#a50f15	#08519c	#a50f15	#a50f15	370	168	0	Economy in reces
1996	#08519c	49.23	40.72	538	#888888	#888888	#08519c	#888888	#a50f15	#08519c	#a50f15	#08519c	#a50f15	#a50f15	379	159	0	Recovered econo
2000	#a50f15	48.38	47.87	537	#888888	#888888	#08519c	#888888	#a50f15	#08519c	#a50f15	#08519c	#a50f15	#a50f15	266	271	0	Controversy over
2004	#a50f15	48.26	50.73	537	#888888	#888888	#08519c	#888888	#a50f15	#08519c	#a50f15	#a50f15	#a50f15	#a50f15	251	286	0	911 in 2001, War
2008	#08519c	52.86	45.6	538	#888888	#888888	#08519c	#08519c	#08519c	#08519c	#a50f15	#08519c	#08519c	#a50f15	365	173	0	2008 economic cı
2012	#08519c	51.01	47.15	538	#888888	#888888	#08519c	#08519c	#08519c	#08519c	#a50f15	#a50f15	#a50f15	#a50f15	332	206	0	Economic recove

Exploratory Data Analysis

After we prepared and processed the raw data, we leveraged Tableau dashboard to come up with a general idea of how our data looks like and where each part should be arranged.

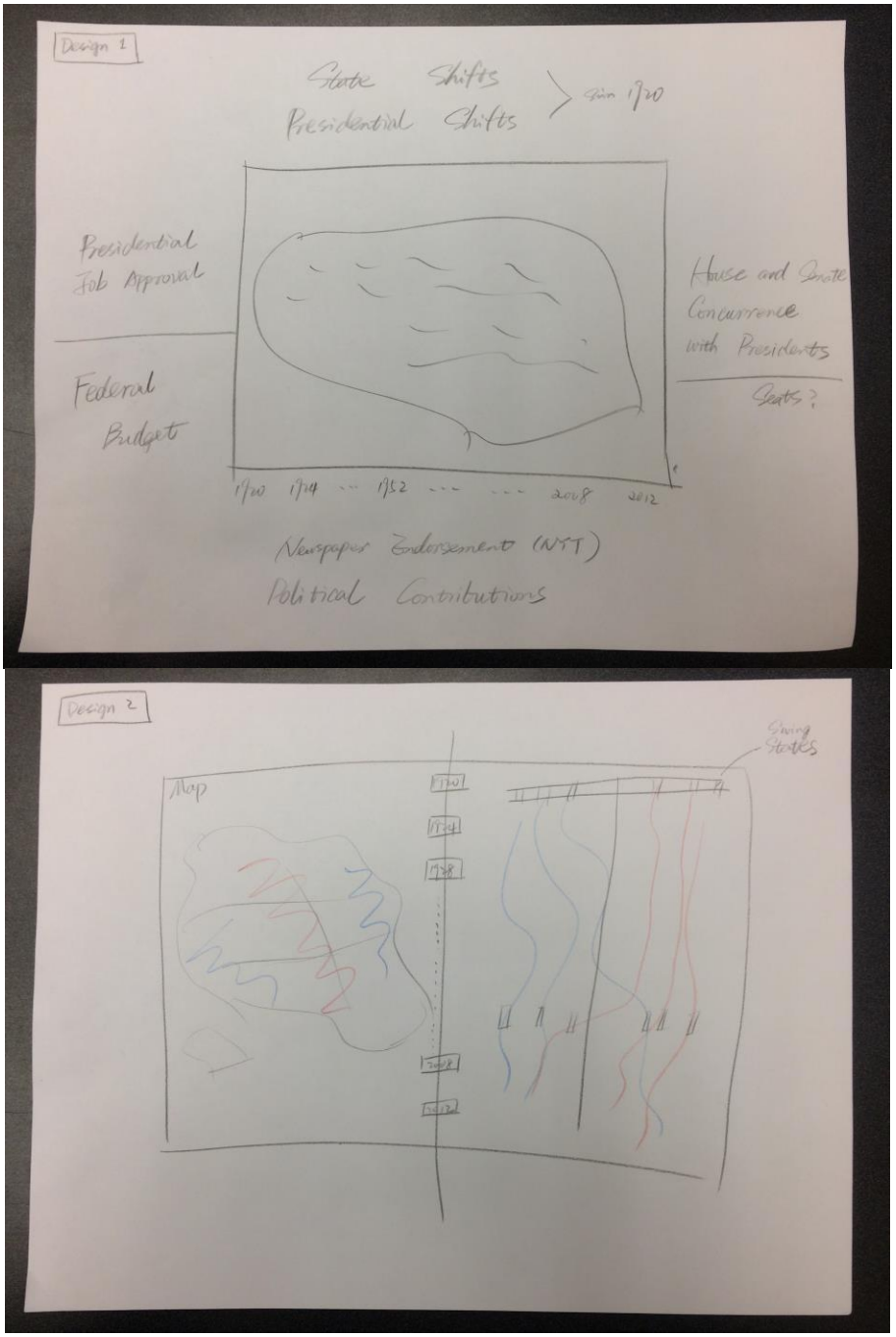
Here are some insights we obtained from Tableau dashboard, which helped us in our future viz design with D3.js.

- Titles are needed for some charts, to help the audience better understand what is displayed here.
- When interacting, tooltips can be a good choice; for mouse-click that cause changes in coordinated views, there should be a clear emphasis (possible ways can be change of color, size, etc.) on the according data and automatically reset afterwards (Tableau does poor in that).
- Scrolling down to check the viz in the bottom can be annoying, therefore we would try to set the layout in one screen, adaptive to any resolution.

Design Evolution

1. Early Stage

We sketched two different designs for historical data, as shown below. The first design focuses on the map display, and attempts to include various types of information in the vis; the second design focuses on the trend of every state, which may show some pattern of swing states.

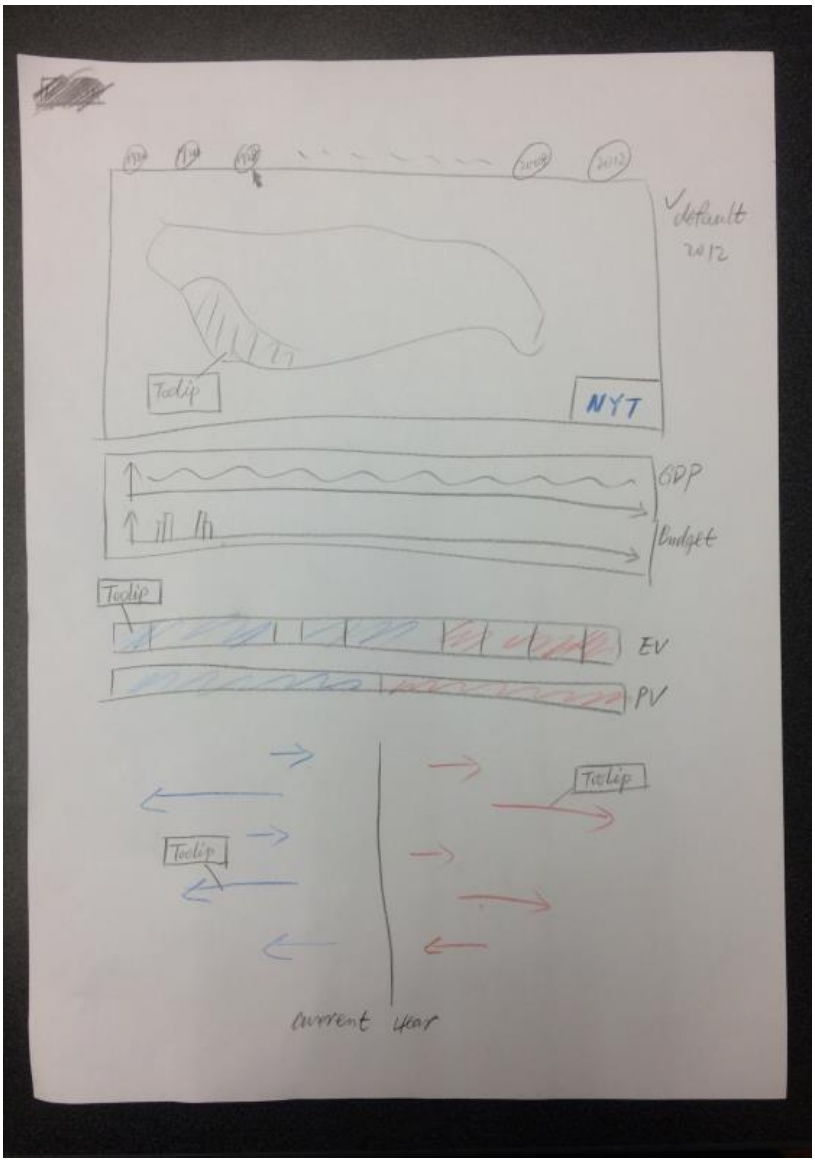


After discussion, we decided to use the first design, which is more consistent with our motivation to show as much information as possible in the vis.

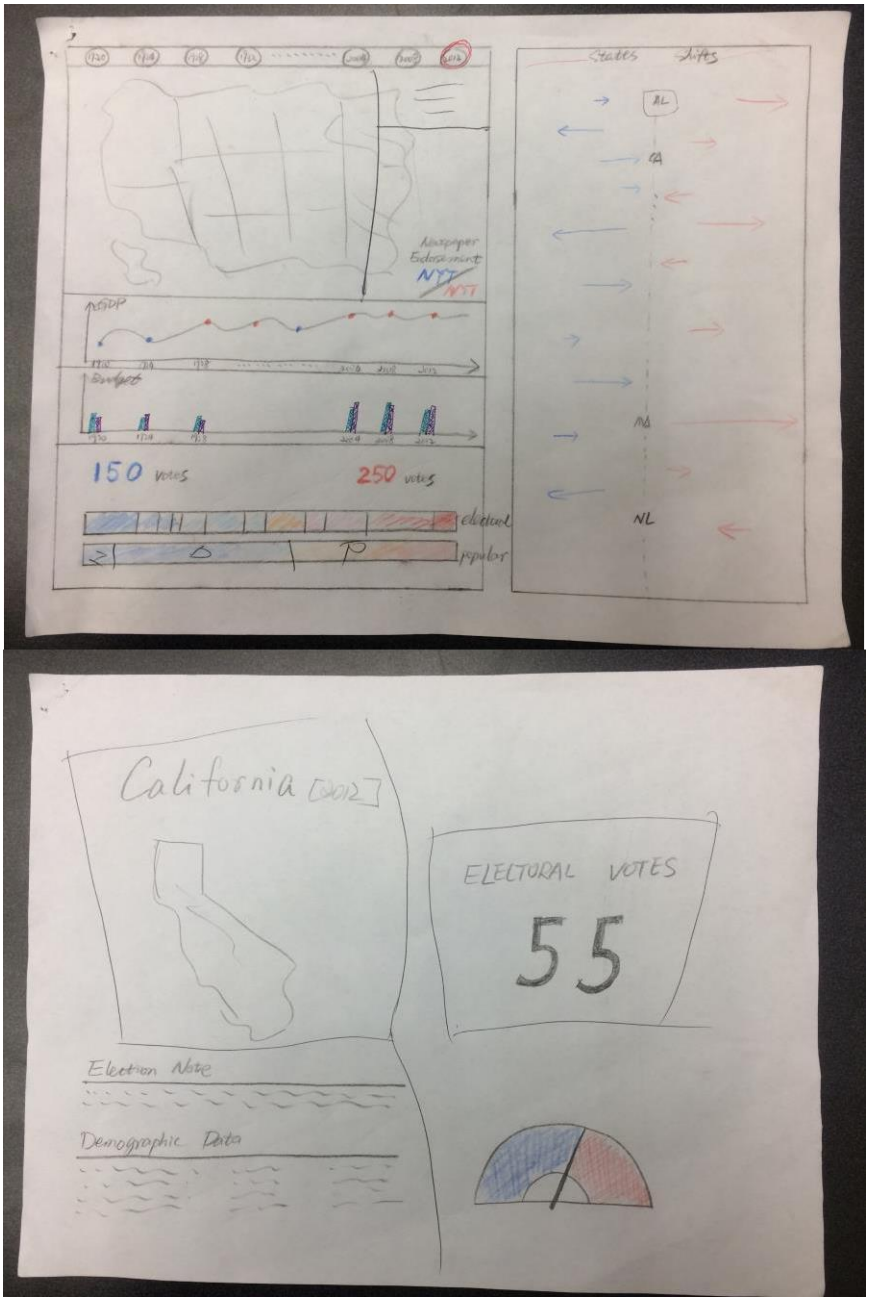
2. Vis Draft

Once we decided to use design 1, the next step is trying to generate a draft of our vis in more detail.

First, we created the whole viz in vertical-fashion:



But considering it will expand the whole page by using the scroll, it may not straightforward for audience to capture an entire idea at each year. Thus, we rearranged the page to a horizontal-fashion, and add a second page corresponding the first page.



Furthermore, we reconsidered the second page which is explicitly a zoom-in version of the first page, and concluded that the second page was totally redundant information. Therefore, the final vis draft is just the first page in horizontal type.

3. Operation Design

For timeline: “click” should trigger all the other vis elements.

For map: “hanover” should trigger tooltip of every state; “click” should trigger relevant highlights in other vis elements.

For stack bar: “hanover” should trigger tooltip of every state; “click” should trigger relevant highlights in map.

For shift slope chart: “hanover” should trigger tooltip of every state.

4. Vis Generating Process

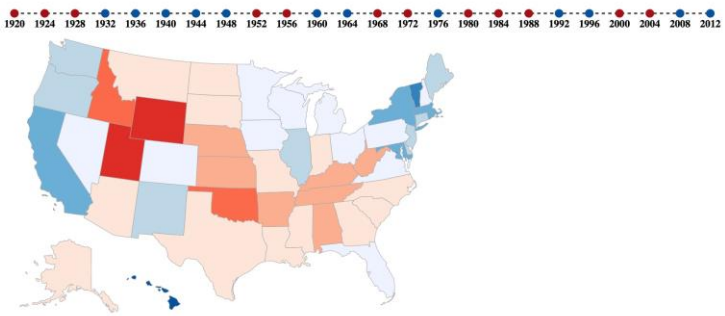
4.1 Layout Design

Each color block represents a separate div and is arranged as below.

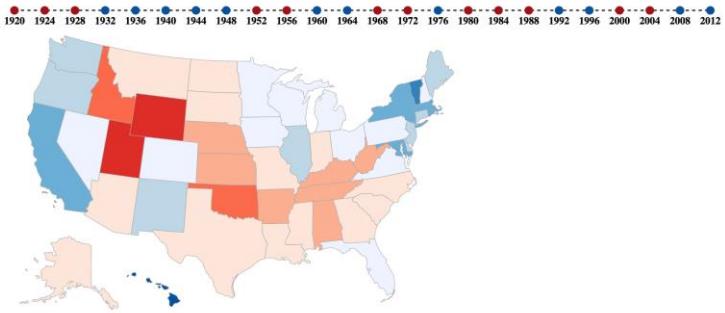
4.2 Timeline.

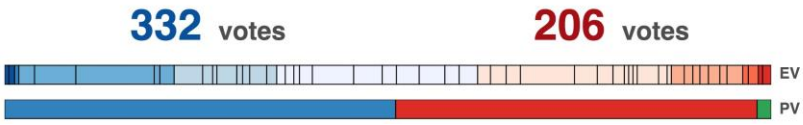
The timeline is adaptive to user's screen resolution, so are all of our other vis.

4.3 Map



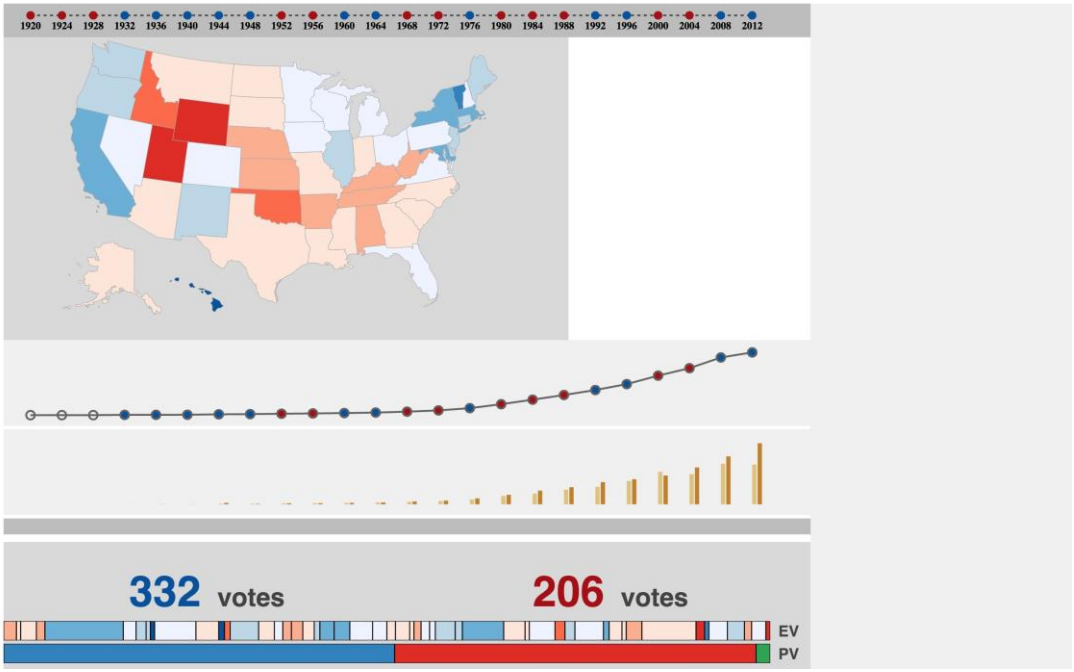
4.4 Information of Popular Vote and Electoral Vote



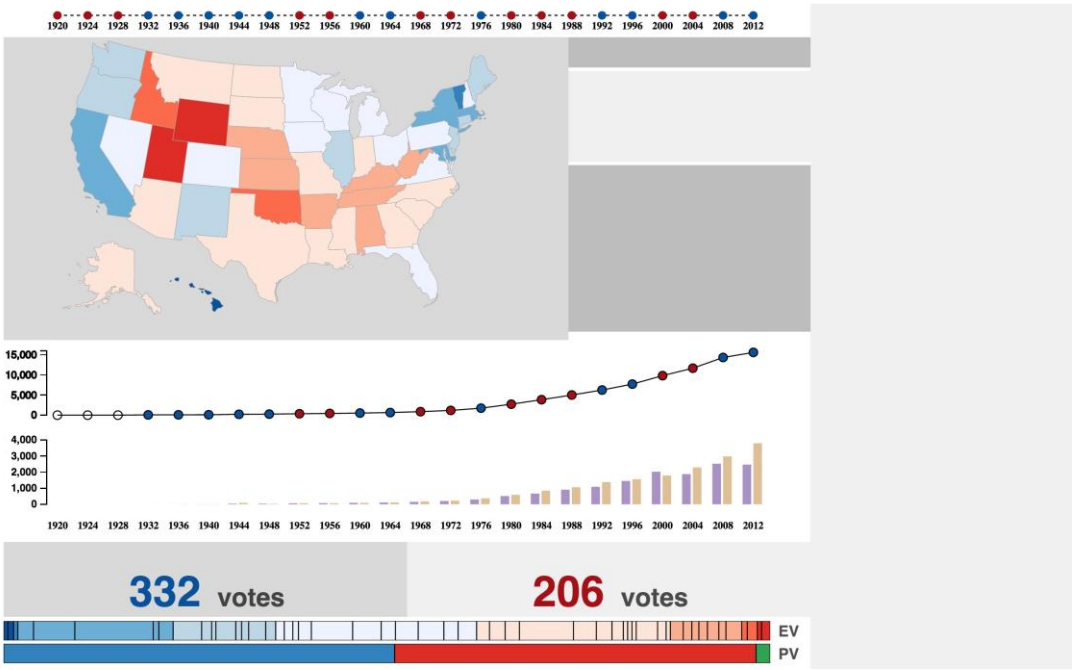


4.5 Financial Data

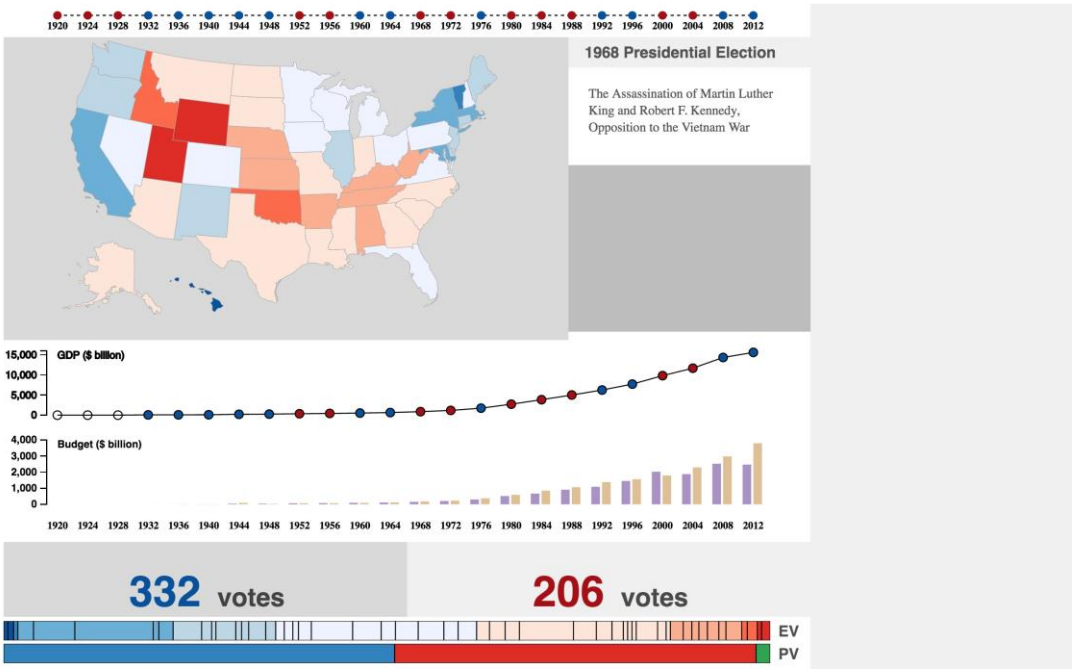
GDP and Federal budget trend charts are added here. While in the final version, we substitute the budget one with unemployment rate, which makes more sense and aligns with the key words we provide in the background info part.



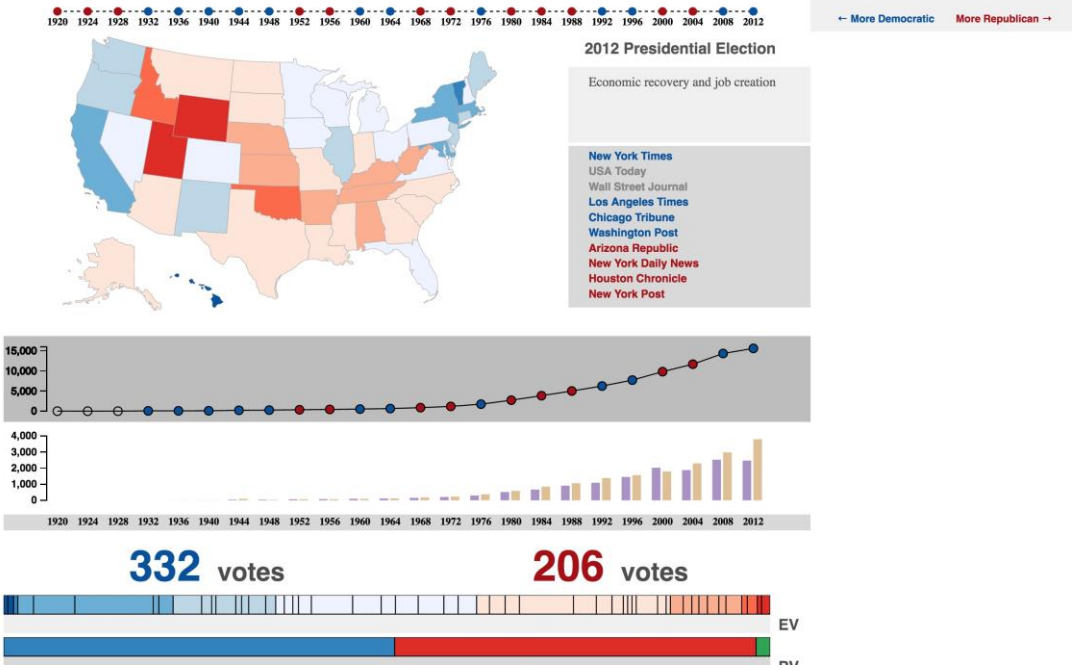
4.6 Redesign the Caption div as story-media
We divide it into 3 parts: storyTitle, story and media



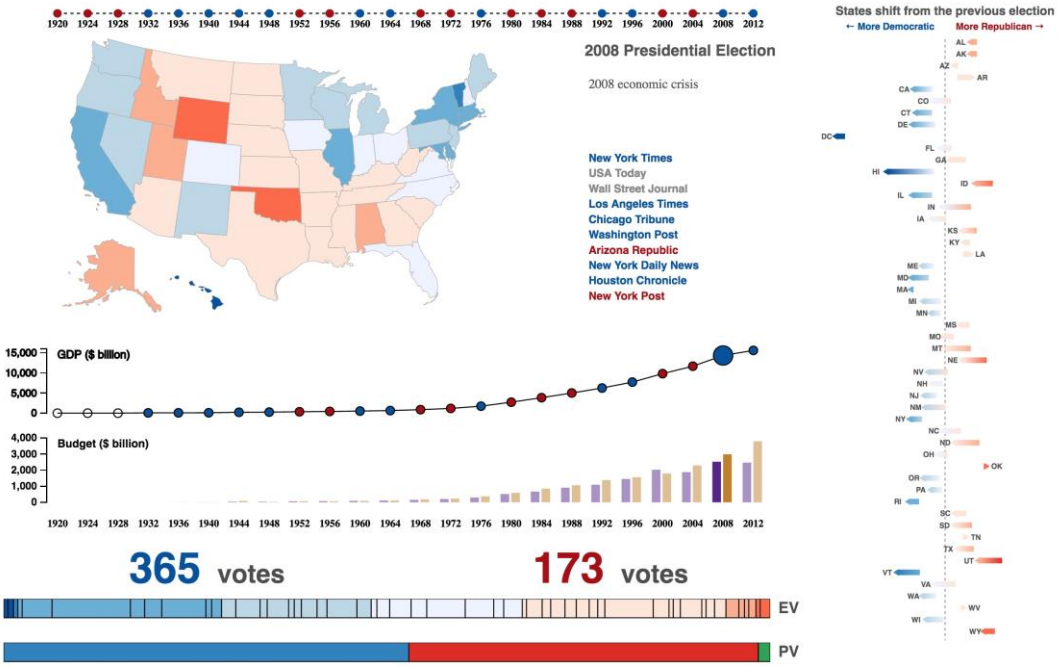
4.7 Story title and story with key words



4.8 Media Endorsement

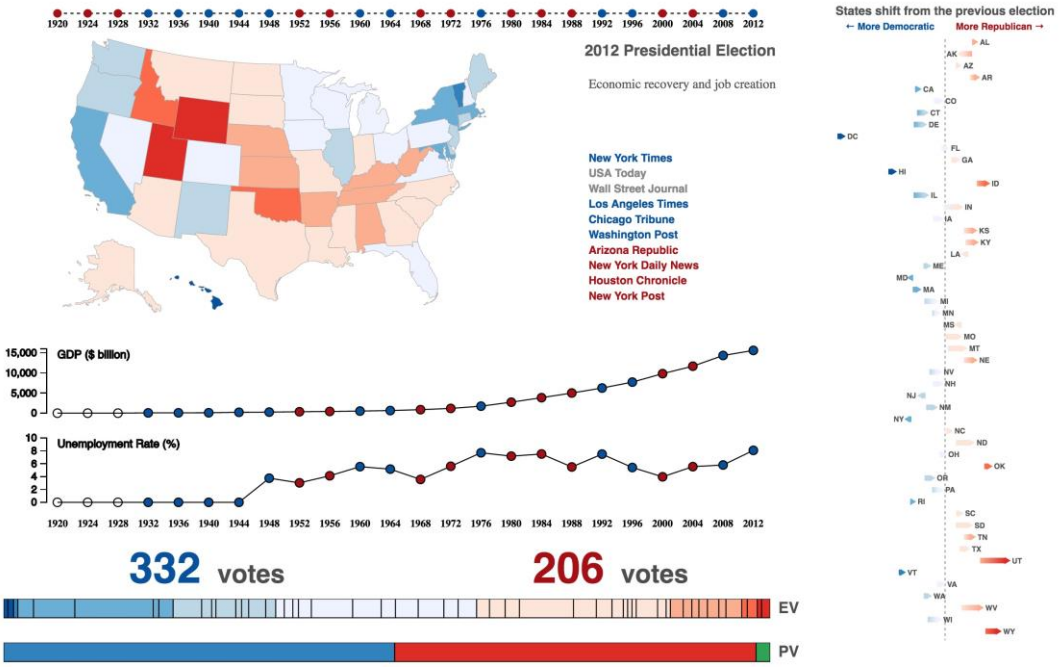


4.9 State Shift Chart



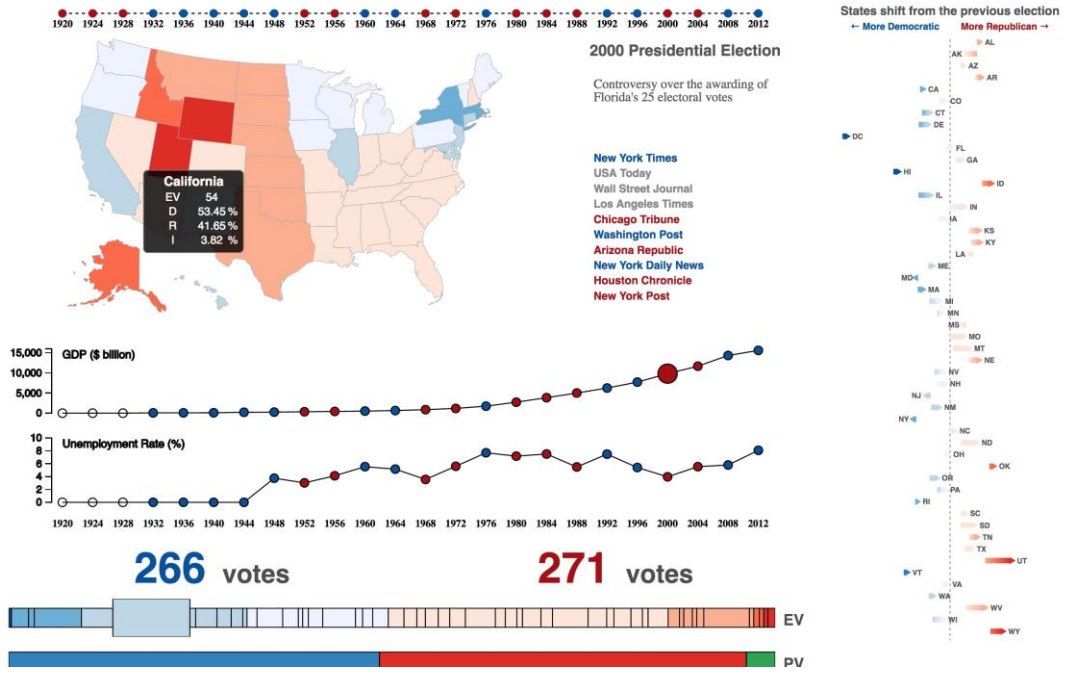
4.10 Unemployment Rate Trend

It replaces the formal Federal Budget one, as federal budget is always going higher and higher as time goes by but unemployment rate has closely relation with economy depression/ booming and is more related with every one's life.



4.11 From Static to Interactive

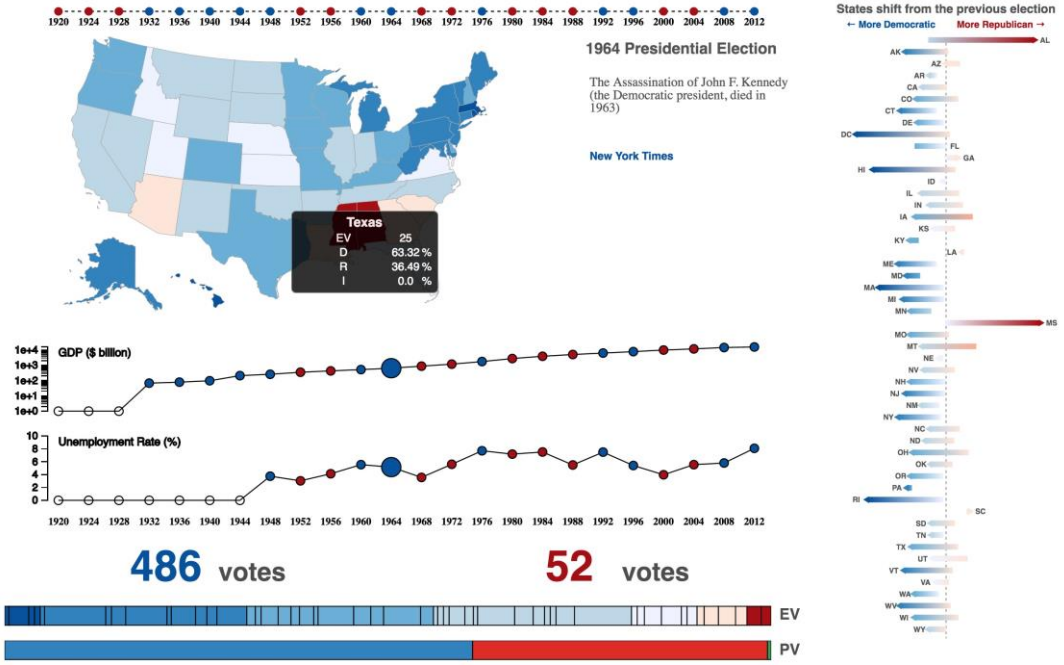
When clicking on a specific year, the according dot in the trend and block in the stack will pop bigger.

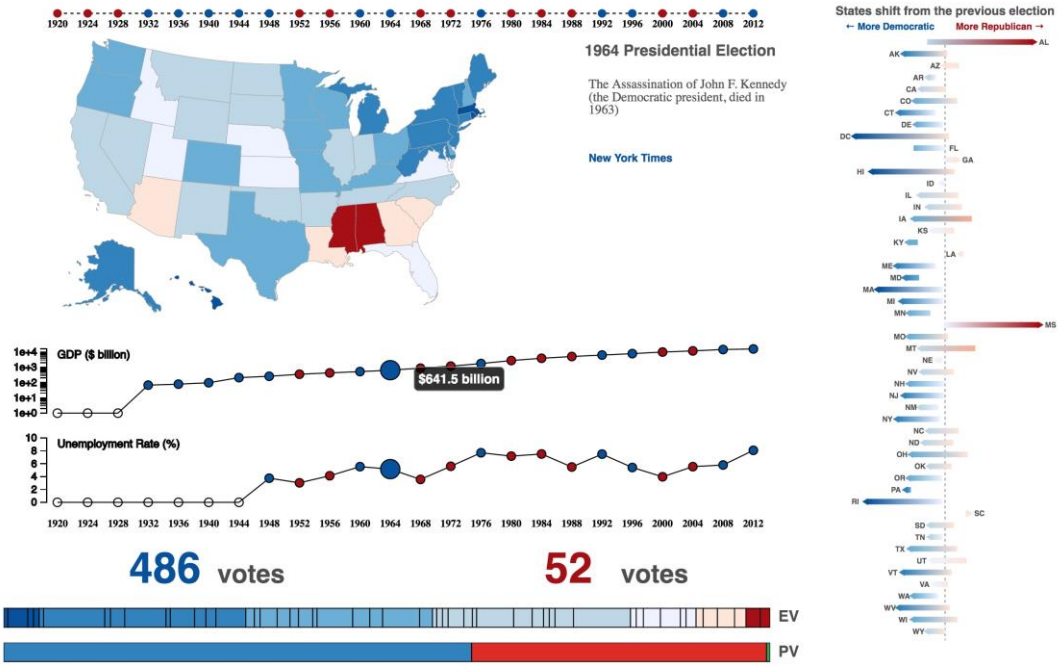


4.12 Tooltips

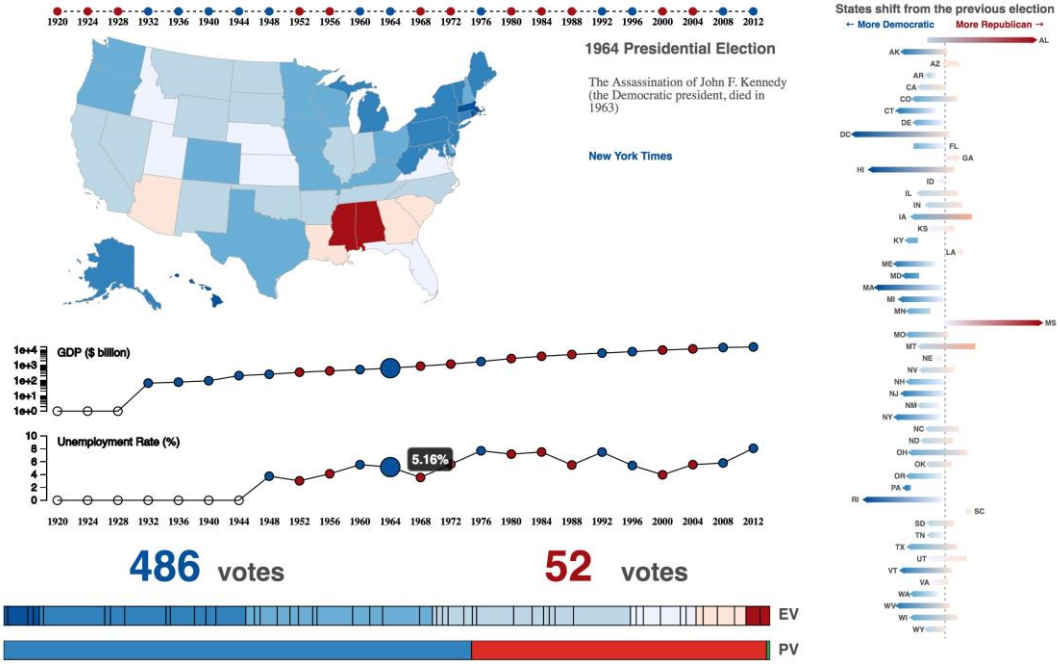
From map to trend chart as well as the shift chart, tooltip with data label or state name will help the audience understand the information more concisely.

Also, we fix the y-axis of GDP trend to make the subtle change more visible.

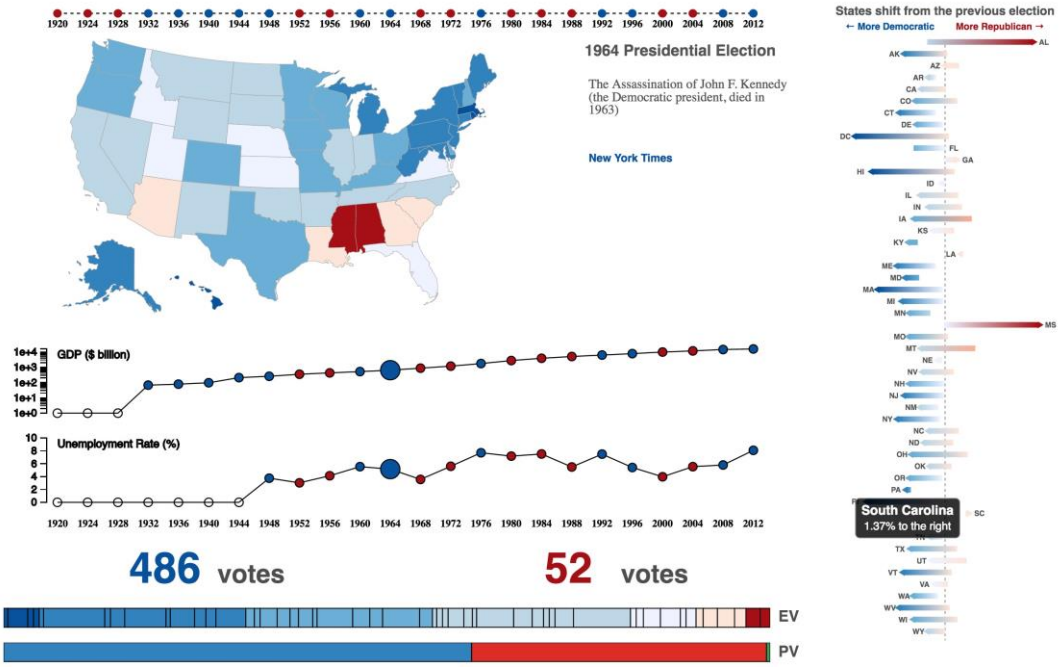



1920 1924 1928 1932 1936 1940 1944 1948 1952 1956 1960 1964 1968 1972 1976 1980 1984 1988 1992 1996 2000 2004 2008 2012
1964 Presidential Election
The Assassination of John F. Kennedy (the Democratic president, died in 1963)
New York Times
GDP ($ billion)
$641.5 billion
Unemployment Rate (%)
486 votes
52 votes
EV
PV

States shift from the previous election
← More Democratic
More Republican →

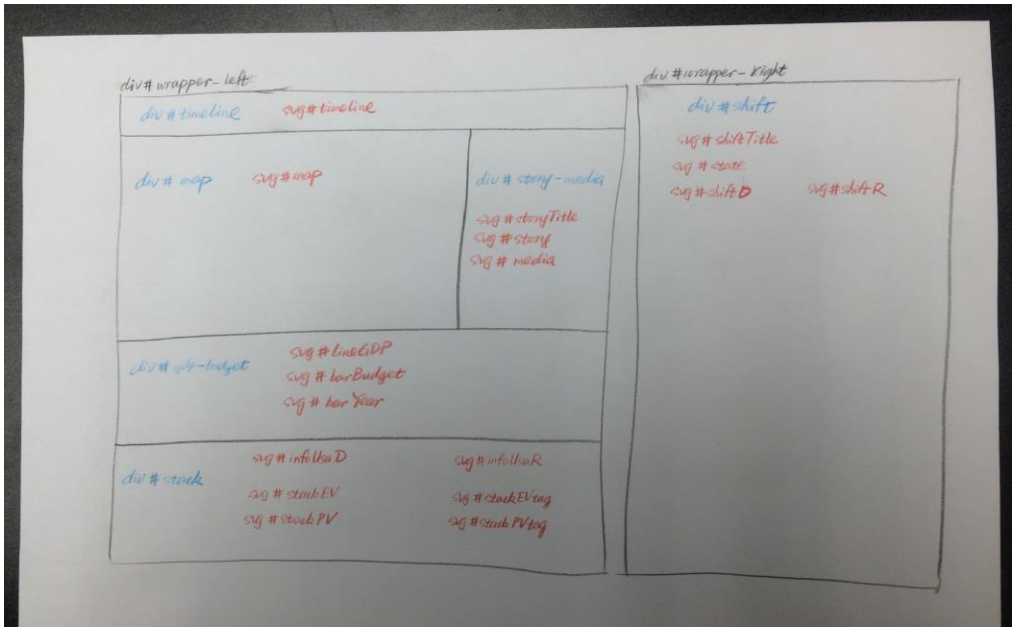

1920 1924 1928 1932 1936 1940 1944 1948 1952 1956 1960 1964 1968 1972 1976 1980 1984 1988 1992 1996 2000 2004 2008 2012
1964 Presidential Election
The Assassination of John F. Kennedy (the Democratic president, died in 1963)
New York Times
GDP ($ billion)
Unemployment Rate (%)
5.16%
486 votes
52 votes
EV
PV

States shift from the previous election
← More Democratic
More Republican →



Implementation

1. Vis Elements Structure



For div#wrapper-left, it includes:

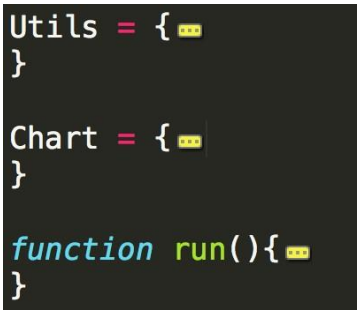
- div#timeline
  - svg#timeline
- div#map
  - svg#map
- div#story-media
  - svg#storyTitle
  - svg#story
  - svg#media
- div#gdp-budget
  - svg#lineGDP
  - svg#barBudget
  - svg#barYear
- div#stack
  - svg#infoD
  - svg#infoR
  - svg#stackEV
  - svg#stackPV
  - svg#stackEVtag
  - svg#stackPVtag

For div#wrapper-right, it includes:

- div#shift
  - svg#shiftTitle
  - svg#shiftState

2. Code Architecture

Generally speaking, Utils includes the interaction code and Chart includes all the static charts; with run function, they could be connected and executed together.



1) Utils

When a specific year on the timeline is clicked, the information in other charts like map, vote text, and stacked bars would be updated.

```
Utils = {

    updateMap: function(year){
    },

    updateStoryTitle: function(year){
    },

    updateStory: function(year){
    },

    updateStackEV: function(year){
    },

    updateStackPV: function(year){
    },

    updateUSADR: function(year){
    },

    updateGDPBudget: function(year){
    },

    updateStackEVForMap: function(id){
    },

    /*
    updateShift: function(data, category){}*/

    update(d){
    },

    updateForMap(d){
    }
}
```

2) Chart

So far, 11 Vis elements are included in Chart, other one like the state shift map is expected to be done in the next stage.

```
Chart = {

    timeline: function(id){
    },

    map: function(id, year){
    },

    storyTitle: function(id, year) {
    },

    story: function(id, year) {
    },

    lineGDP: function(id){
    },

    barBudget: function(id){
    },

    barYear: function(id) {
    },

    infoUSAD: function(id, year){
    },

    infoUSAR: function(id, year){
    },

    stackEV: function(id, year){
    },
```

Evaluation

We've heard much from the media of Election 2016, but what does the data say, about the US elections over the history? Join us to explore the data stories behind all things election!

- Firstly, with the red or blue dots on the timeline, we could quickly and easily figure out the winning party of each election.
- Then, we could go a step further to look at the hue US map, which gives us a straightforward feeling of how "red" or "blue" a state is. Besides, we could find a sorted detail about every state's electoral votes in the stack bar below, along with the information of popular votes in that election year.
- Later, we could get some key events occurred during that election period, along with media endorsement information.
- Moreover, we could also get an overview of change in GDP and unemployment rate over time.
- For the interest in state shift scenario, we can find how every state shifts from the previous election.
- And speaking of more detail, we could see the relevant tooltips when mouse over the state in map or in the shift chart, or mouse over every point in the GDP and unemployment rate charts.

From this visualization, we found some interesting insights. For example:

- At year 1924, 1948 and 1968, there are some states vote for a third party other than the Democratic and the Republican.
- At year 1932 and 1936, we can see almost the whole country vote for the Democratic, with the effect of the 1929 wall street crash and the great depression, Franklin Roosevelt promised reform in his policy called the new deal. And during the election of 1936, he was still working to push the provisions of his new deal economic policy through congress and courts.
- While at year 1972 almost the whole country vote for the Republican, emphasizing a good economy and his success in foreign affairs, the Republican won the election in a massive landslide.
- But at year 2000, there is a very close tie between two parties. It was the closet election since 1876.
- Now let's focus on the states shift chart, we found some patterns of election during certain period. As designed in the chart, the length of bar denoted to what extent that state's political opinion changes, the direction of bar denoted whether the state is more democratic or more republican. For example, from year 1964 to 1980, we can observe clearly a flip pattern election by election

For future work, as we've seen several good storytelling examples like the NYTime's How Trump Can Influence Climate Change, http://www.nytimes.com/interactive/2016/12/08/us/trump-climate-change.html?smid=tw-nytimes&smtyp=cur&_r=1, which is more narrative with the caption showing along with the data, we could leverage this idea to let our vis more understandable.