



# Voting Project

By Carlos Saloni and Jaromir Parker



# HISTORY/MOTIVATION

For our project we chose voting, because of the recent political climate. We hope that when people see our project it reminds them of the importance of voting and making their voice heard.



# CHALLENGES

One of the hardest things about our project was troubleshooting our code. The amount of times we had to redo code and find substitutions was at times overwhelming.

# Organization of Data

US Elections Dataset | Kaggle

notebook3ff180c0d9 | Kaggle

https://www.kaggle.com/carlosparker/notebook3ff180c0d9/edit

notebook3ff180c0d9 Draft saved

File Edit View Run Add-ons Help

Run All

Draft Session (0m)

	fips	votes	fraction_votes
count	2.451100e+04	24611.000000	24611.000000
mean	2.667152e+07	2306.252773	0.304524
std	4.200978e+07	9861.183572	0.231401
min	1.001000e-03	0.000000	0.000000
25%	2.109100e+04	68.000000	0.094000
50%	4.208100e+04	358.000000	0.273000
75%	9.090012e+07	1375.000000	0.479000
max	9.560004e+07	590502.000000	1.000000

In[201]:

```
Electionresults_data
```

Out[201]:

	state	state_abbreviation	county	fips	party	candidate	votes	fraction_votes
0	Vermont	VT	Sutton	95000197.0	Republican	John Kasich	20	0.227
1	Vermont	VT	Tunbridge	95000204.0	Republican	John Kasich	36	0.319
2	Vermont	VT	Weathersfield	95000220.0	Republican	Ted Cruz	46	0.111
3	Vermont	VT	Weston	95000228.0	Republican	John Kasich	32	0.471
4	Vermont	VT	Whiting	95000233.0	Republican	Ben Carson	0	0.000
...	...	...	...	...	...	...	...	...
24606	Wisconsin	WI	Racine	55101.0	Republican	Donald Trump	11756	0.321
24607	Wisconsin	WI	Rusk	55107.0	Republican	Ted Cruz	1339	0.410
24608	Wisconsin	WI	Sawyer	55113.0	Republican	Donald Trump	2034	0.539
24609	Wyoming	WY	Weston	56045.0	Democrat	Bernie Sanders	1	0.500
24610	Wyoming	WY	Hot Springs-Washakie	95600030.0	Republican	Donald Trump	0	0.000

24611 rows x 8 columns

+ Code + Markdown

In[202]:

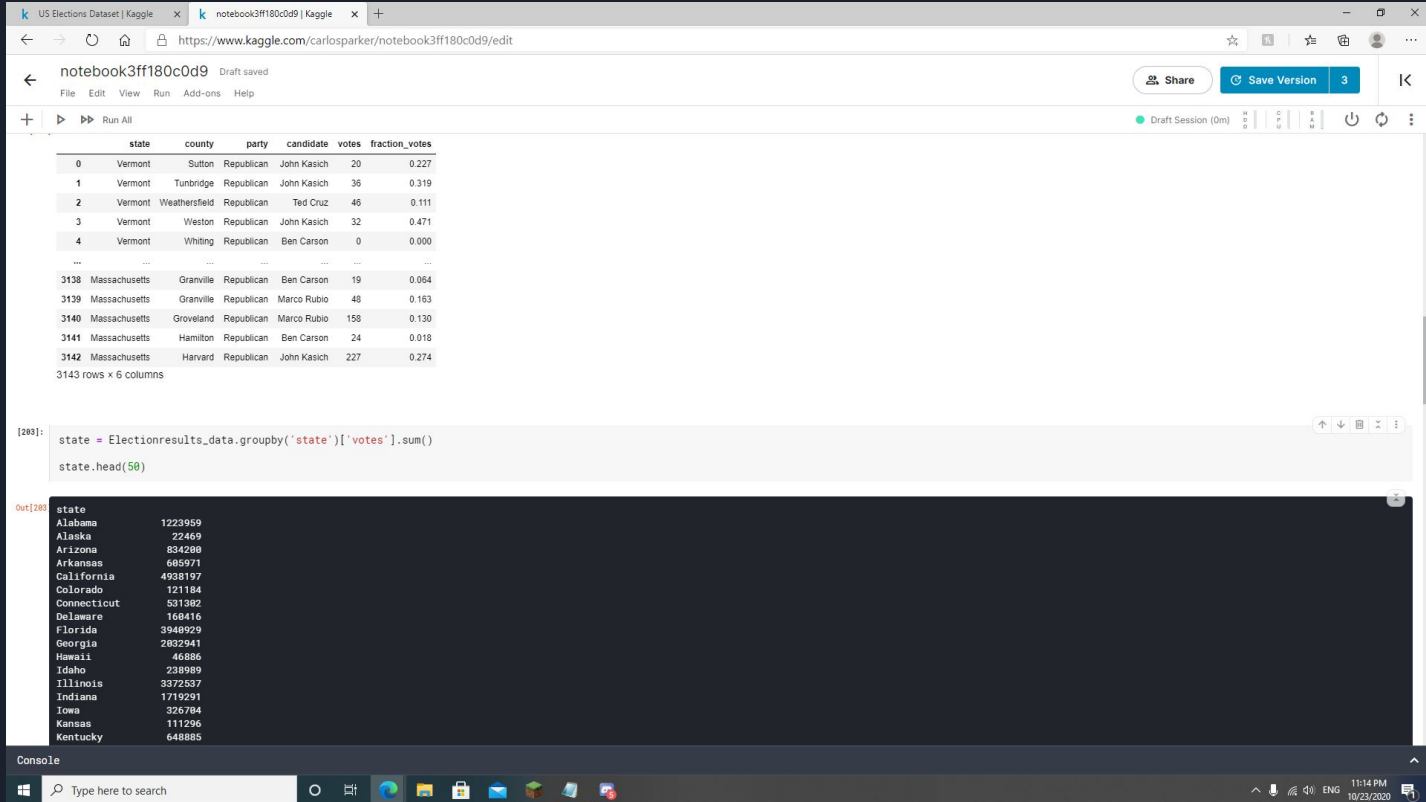
```
Electionresults_data.drop(columns=['fips', 'state_abbreviation'], inplace=True)  
Electionresults_data.head(3143)
```

Console

Type here to search

11:14 PM  
10/23/2020

# Organization of data



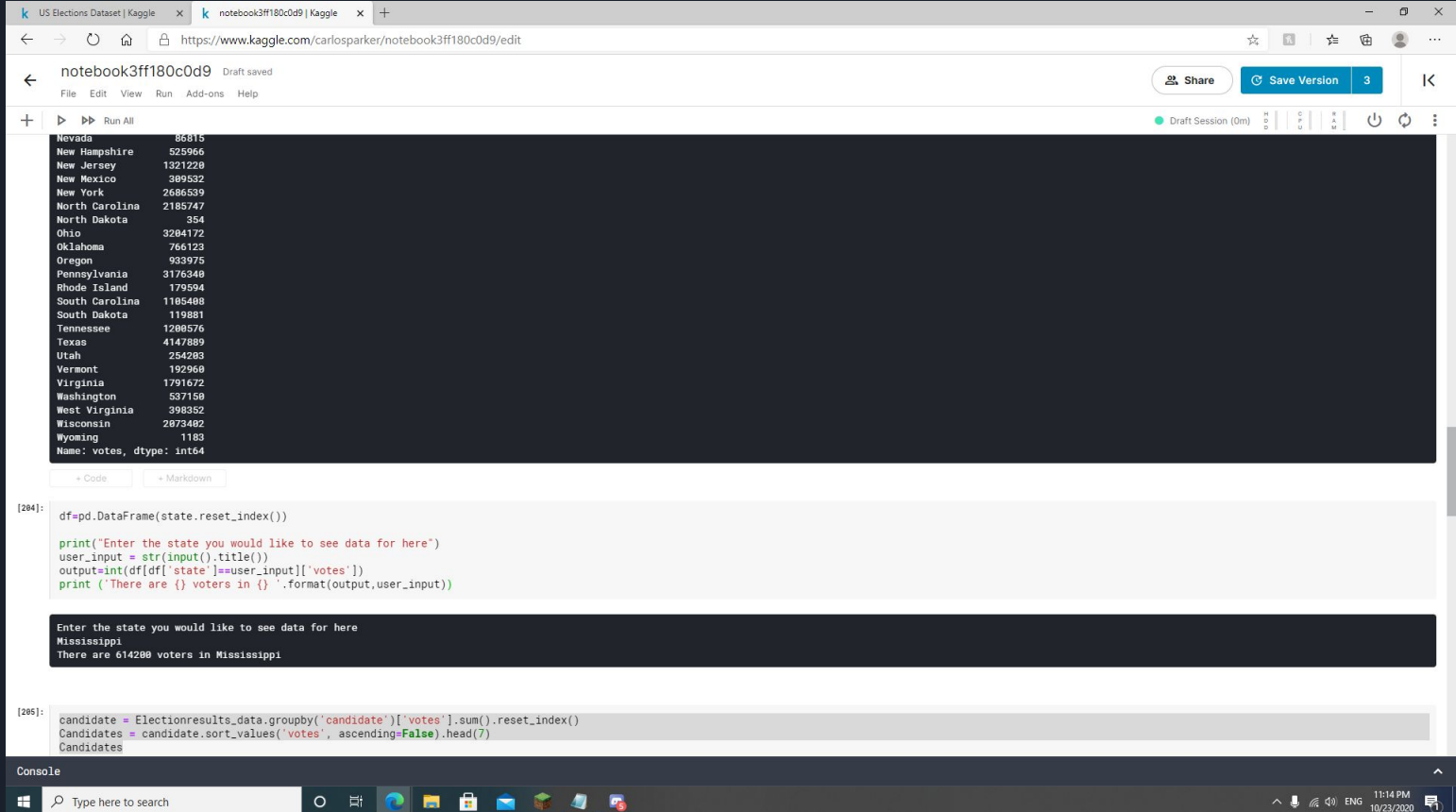
The screenshot shows a Kaggle notebook interface. At the top, there are browser tabs for 'US Elections Dataset | Kaggle' and 'notebook3ff180c0d9 | Kaggle'. The notebook title is 'notebook3ff180c0d9' and it is in 'Draft saved' mode. The notebook contains a data table with 6 columns: 'state', 'county', 'party', 'candidate', 'votes', and 'fraction\_votes'. Below the table, a code cell is executed, showing the output of a pandas groupby operation.

	state	county	party	candidate	votes	fraction_votes
0	Vermont	Sutton	Republican	John Kasich	20	0.227
1	Vermont	Tunbridge	Republican	John Kasich	36	0.319
2	Vermont	Weathersfield	Republican	Ted Cruz	46	0.111
3	Vermont	Weston	Republican	John Kasich	32	0.471
4	Vermont	Whiting	Republican	Ben Carson	0	0.000
...	...	...	...	...	...	...
3138	Massachusetts	Granville	Republican	Ben Carson	19	0.064
3139	Massachusetts	Granville	Republican	Marco Rubio	48	0.163
3140	Massachusetts	Groveland	Republican	Marco Rubio	158	0.130
3141	Massachusetts	Hamilton	Republican	Ben Carson	24	0.018
3142	Massachusetts	Harvard	Republican	John Kasich	227	0.274

```
[289]: state = Electionresults_data.groupby('state')['votes'].sum()
state.head(50)
```

```
Out[289]: state
Alabama      1223959
Alaska       22469
Arizona      834288
Arkansas     685971
California   4938197
Colorado     121184
Connecticut  531382
Delaware     168416
Florida      3948929
Georgia     2832941
Hawaii       46886
Idaho        238989
Illinois     3372537
Indiana      1719291
Iowa         326784
Kansas       111296
Kentucky     648885
```

# Organization of Data



The screenshot shows a Kaggle notebook interface. At the top, there are browser tabs for 'US Elections Dataset | Kaggle' and 'notebook3ff180c0d9 | Kaggle'. The address bar shows the URL 'https://www.kaggle.com/carlosparker/notebook3ff180c0d9/edit'. The notebook title is 'notebook3ff180c0d9' with a 'Draft saved' status. The interface includes a 'Share' button, a 'Save Version' button (set to 3), and a 'Run All' button. The main content area displays a list of US states and their corresponding voter counts:

Nevada	86815
New Hampshire	525966
New Jersey	1321228
New Mexico	389532
New York	2686539
North Carolina	2185747
North Dakota	354
Ohio	3284172
Oklahoma	766123
Oregon	933975
Pennsylvania	3176348
Rhode Island	179594
South Carolina	1185488
South Dakota	119881
Tennessee	1280576
Texas	4147889
Utah	254288
Vermont	192868
Virginia	1791672
Washington	537158
West Virginia	398352
Wisconsin	2873482
Wyoming	1183

Below the list, there is a code cell with the following code:

```
[284]: df=pd.DataFrame(state.reset_index())

print("Enter the state you would like to see data for here")
user_input = str(input().title())
output=int(df[df['state']==user_input]['votes'])
print ('There are {} voters in {}'.format(output,user_input))
```

The output of the code cell is:

```
Enter the state you would like to see data for here
Mississippi
There are 614288 voters in Mississippi
```

Below the output, there is another code cell with the following code:

```
[285]: candidate = Electionresults_data.groupby("candidate")["votes"].sum().reset_index()
Candidates = candidate.sort_values("votes", ascending=False).head(7)
Candidates
```

At the bottom of the notebook, there is a 'Console' section with a search bar and a Windows taskbar. The taskbar shows the time as 11:14 PM on 10/23/2020.

# Visualization of Data

US Elections Datasets | Kaggle x notebook3ff180c0d9 | Kaggle x

https://www.kaggle.com/carlosparker/notebook3ff180c0d9/edit

notebook3ff180c0d9 Draft saved

File Edit View Run Add-ons Help

Share Save Version 3

Run All

Draft Session (0m)

	candidate	votes
5	Hillary Clinton	15692452
4	Donald Trump	13302541
1	Bernie Sanders	11959102
14	Ted Cruz	7603006
7	John Kasich	4159949
8	Marco Rubio	3321076
0	Ben Carson	564553

```
import plotly.express as px

fig = px.pie(Candidates, values = Candidates['votes'], names = Candidates['candidate'])
fig.update_layout(title = 'Vote percentages by Candidate over entire nation')
fig.show()
```

Vote percentages by Candidate over entire nation

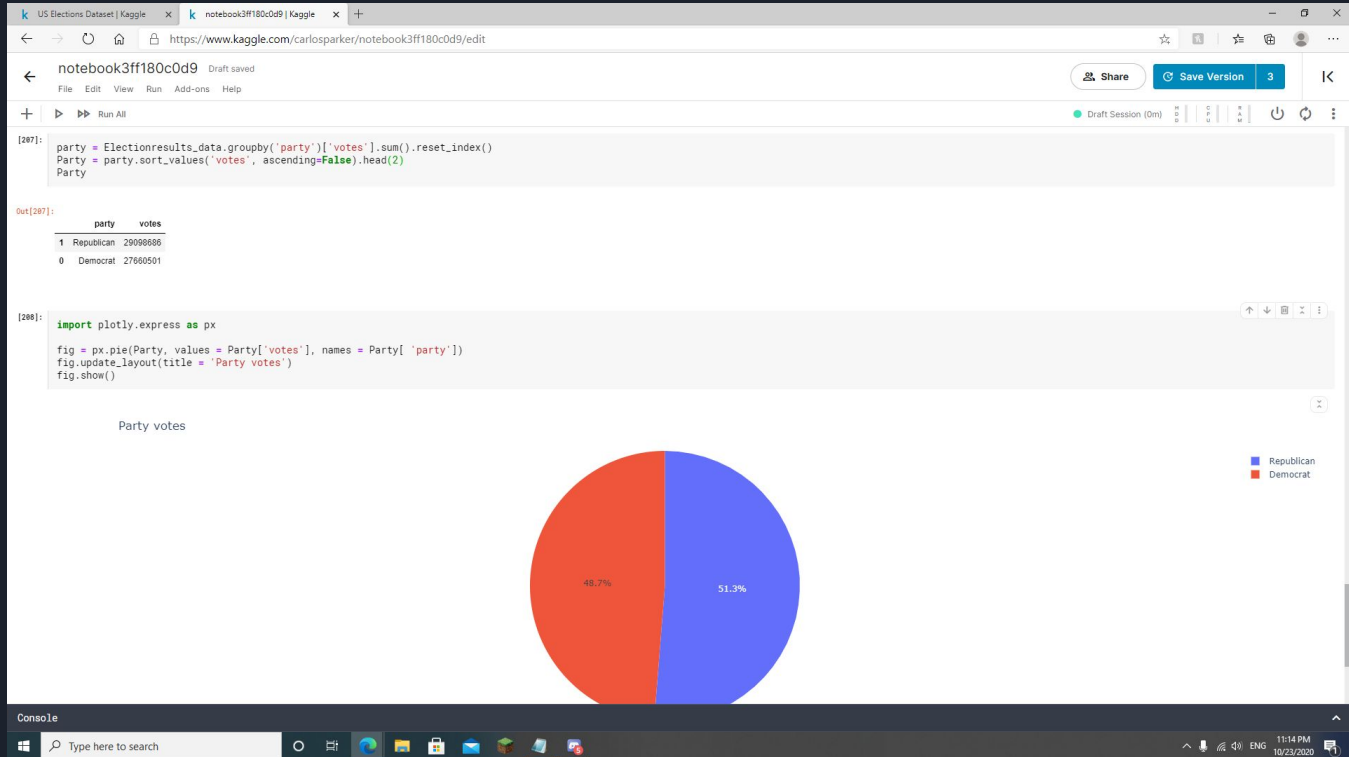
Candidate	Percentage
Hillary Clinton	27.7%
Donald Trump	23.5%
Bernie Sanders	21.1%
Ted Cruz	7.35%
John Kasich	13.4%
Marco Rubio	5.87%
Ben Carson	0.997%

Console

Type here to search

11:14 PM 10/23/2020

# Visualization of Data (continued)



The screenshot shows a Kaggle notebook interface with the following content:


```
[287]: party = ElectionResults_data.groupby('party')['votes'].sum().reset_index()
Party = party.sort_values('votes', ascending=False).head(2)
Party
```

Out[287]:

	party	votes
1	Republican	29098686
0	Democrat	27660501

```
[288]: import plotly.express as px
fig = px.pie(Party, values = Party['votes'], names = Party['party'])
fig.update_layout(title = 'Party votes')
fig.show()
```

Party votes



A pie chart titled "Party votes" showing the distribution of votes between two parties. The Republican party (blue) holds 51.3% of the votes, and the Democrat party (red) holds 48.7%. A legend on the right identifies the colors: blue for Republican and red for Democrat.

Console

Type here to search

11:14 PM 10/23/2020