**Daniel Vassiliou**

**Statistics**

(Stone 8)

Website location - [https://data.fivethirtyeight.com/#](https://data.fivethirtyeight.com/)

Github Data - <https://github.com/fivethirtyeight/data/tree/master/polls>

Data used – Current House polls.

The poll file named current house polls contains data since the most recent election. Data was published to be seen on FiveThirtyEight and I was able to view the “latest polls” on GitHub. The author of this poll is (Arron Bycoffe, Holly Fuong, and Rad Cliffem. But there have been 57 contributors to the publishing of this information. Polling averages run from 1968-2020 but do not provide information of presidential approval prior to the Biden presidency. This only tracked post-2020 trump favorability. The favorability polls prior to this time are not available.

(Stone 8/18)

* pollster\_id – identification number in regard to the pollster
* pollster – The research and or group information was deprived from.
* sponsor\_ids – Identification number in regard to the sponsor from the information provided.
* sponsors – The sponsors/institutions that provided the information to the (current) house polls study.
* display\_name - Name displayed from sponsor.
* pollster\_rating\_id – identification number regarding rating of research of sponsor/institution
* pollster\_rating\_name – name of sponsor/institution being rated.
* fte\_grade – Grade of information provided.
* methodology - how information was deprived.
* transparency\_score – score of how easily information was deprived.
* State – State sponsor was from
* start\_date – Date started with information provided
* end\_date – Date stated where information was ceased to be provided
* sponsor\_candidate\_id – ID related to the candidate.
* sponsor\_candidate – Candidate related with sponsor/institution
* sponsor\_candiddate\_party – Deams what part sponsor belongs to
* question\_id – ID related to the canidate/questions
* sample\_size – sample size related to sample question
* population – population related to either (Iv or rv)
* subpopulation – no information provided
* population\_full - population related to either (Iv or rv)
* tracking – no information provided
* created\_at – what time information was created at in regard to the spreadsheet
* notes – information in addition to individuals talked with orginal sponsors