

Washington Post

After the initial data cleaning described in our Data Cleaning document, there were 554 fact-checks published by the Washington Post between January 1st 2016 and June 30th, 2021. The dataset only includes fact-checks in which the claimant was a political figure, not Donald Trump, and assigned a Democratic or Republican affiliation. This dataset is called Version 8. We then did some light data cleaning prior to our analysis to clarify claimant names and streamline aberrant textual ratings. That process is described below.

Claimant Name Cleaning

In a program called OpenRefine, we created a new column called ‘claimant_clean’ and manually sorted the names of all claimants in our dataset to determine when there were different spellings of the same name. This was done via the cluster function embedded in the application. When varying iterations were found of the same name, we selected a single name to refer to each person by. For example Charles Schumer was changed to Chuck Schumer. For the claimant column, this cleaning process resulted in a reduction of claimants from 188 individual names to 169. This did not reduce the number of claims, just streamlined claimant names so the same political figure was not listed under several different names. The original claimant names are stored under the variable “claimant”

Standardization of Textual Rating

Next we had to address the issue in textualRating. The Washington Post is not entirely consistent in its determinations of claim ratings. In theory, it has a standard ranking system ordered as follows: Geppetto Checkmark, One Pinocchio, Two Pinocchios, Three Pinocchios, Four Pinocchios. Of the 554 claims we pulled from the Washington Post, only 335 were originally ranked on the WaPo’s five scale ratings system. These definitions go as follows, sourced from <https://www.washingtonpost.com/politics/2019/01/07/about-fact-checker/>:

The Geppetto Checkmark: Statements and claims that contain “the truth, the whole truth, and nothing but the truth” will be recognized with our prized Geppetto checkmark. We tend to reserve this for claims that are unexpectedly true, so it is not awarded very often.

One Pinocchio: Some shading of facts. Selective telling of truth. Some omissions and exaggerations but no outright falsehoods. (You could view this as “mostly true.”)

Two Pinocchios: Significant omissions and/or exaggerations. Some factual error may be involved but not necessarily. A politician can create a false, misleading impression by playing with words and using legalistic language that means little to ordinary people. (Similar to “half true.”)

Three Pinocchios: Significant factual error and/or obvious contradictions. This gets into the realm of “mostly false.” But it could include statements which are technically correct (such as based on official government data) but are so taken out of context as to be very misleading.

Four Pinocchios: Whoppers.

Our original data had these textual ratings:

textualRating	n
Four Pinocchios	136
Three Pinocchios	100
Two Pinocchios	89
FALSE	46
Needs context	21
Not the whole story	19
Not the Whole Story	19
Lacks context	12
In dispute	7
Wrong	7
One Pinocchio	6
Cherry-picked number	5
Exaggerated	5
Mostly False	5
Spins the facts	5
Geppetto Checkmark	4
Lacks Context	4
Misleading	4
No evidence	4
Correct	3
Half true	3
Mis- leading	3
Needs Context	3
TRUE	3
Doubtful	2
Fuzzy math	2
Mostly false	2
Not the whole story.	2
Twists the facts	2
Accurate	1
Data is fishy	1
Depends on math	1
Depends on stats used	1
Disputed	1
Evidence gets stronger	1
Evidence is mixed	1
Facile comparison	1
Flip flop	1
Flip- Flop	1
Flubbed Talking Point	1
For the Record	1
Half Flop	1
Half True	1
Incorrect	1
Issue is more complex	1
It's a complex story	1
Lacks evidence	1
Mis- Spoke	1
Needs more context	1
Not based on science	1
Not clear yet	1
Not quite right	1

textualRating	n
Not the same	1
Quoted out of context	1
Spin	1
Spinning the facts	1
Staff error	1
Study in Dispute	1
Techni- cally right	1
Verdict Pending	1

Having the aim to analyze a body of claims all on the standardized rating scale, we began by reassigning certain ratings in OpenRefine based on the exact wording of the Washington Post’s own definitions.

7 **mostly false** ratings were reassigned **Three Pinocchios**.

4 **half true** ratings were reassigned to **Two Pinocchios**.

7 **true**, **correct**, and **accurate** ratings were reassigned the **Geppetto Checkmark**

353 claims were then rated on the standard ranking system.

201 claims remained tagged with anomaly ratings, meaning we could not reliably compare them to the rest of our data. To address this, we reached out to Glenn Kessler, the singular fact-checker at the Washington Post and brains behind the operation. He knows their scale inside and out. We sent over the remaining anomaly tags (sometimes grouping similar ones together) with more than 2 occurrences and received two instructions. First, he identified several tags that could be reassigned without further review. Second, he indicated several tags that he preferred to review and reassign himself. To not overburden Glenn during the manual review process, we removed all claims tagged with an anomaly rating that appeared only once or twice. For the first step, we turned to OpenRefine to make the changes he outlined. This was done to speed up the recoding process, though it could have been done in RStudio if necessary.

Reassignments are the following:

Not the whole story —> Two Pinocchio

Depends on math or on stats —> Two Pinocchio

Wrong or incorrect —> Four Pinocchio

Spins or twists the facts —> Three Pinocchio

Exaggerated —> Three Pinocchio

Cherry-picked number —> Two Pinocchio

False —> Four Pinocchio

This reassigned 115 claims onto the standardized rating scale, leaving us with 468 claims.

Then for the claims that he preferred to self identify we gave him a list of all the claims tagged with the following ratings:

Needs or lacks context

Disputed or in dispute or study in dispute

Misleading

No evidence or lacks evidence

This amounted to 62 total ratings, 58 of which he assigned new ratings to. The remaining four were also removed from the data. This step was done in RStudio.

In total, we were left with 526 claims of the original 554 now ranked by the standard Washington Post system, these are the ones we will use for continued analysis.

The Final Data

Utilizing this standardized dataset, the breakdown of claims by party and rating is found here.

Table 2: Claim Counts by Party

claimant_party	n
Democratic	279
Republican	247

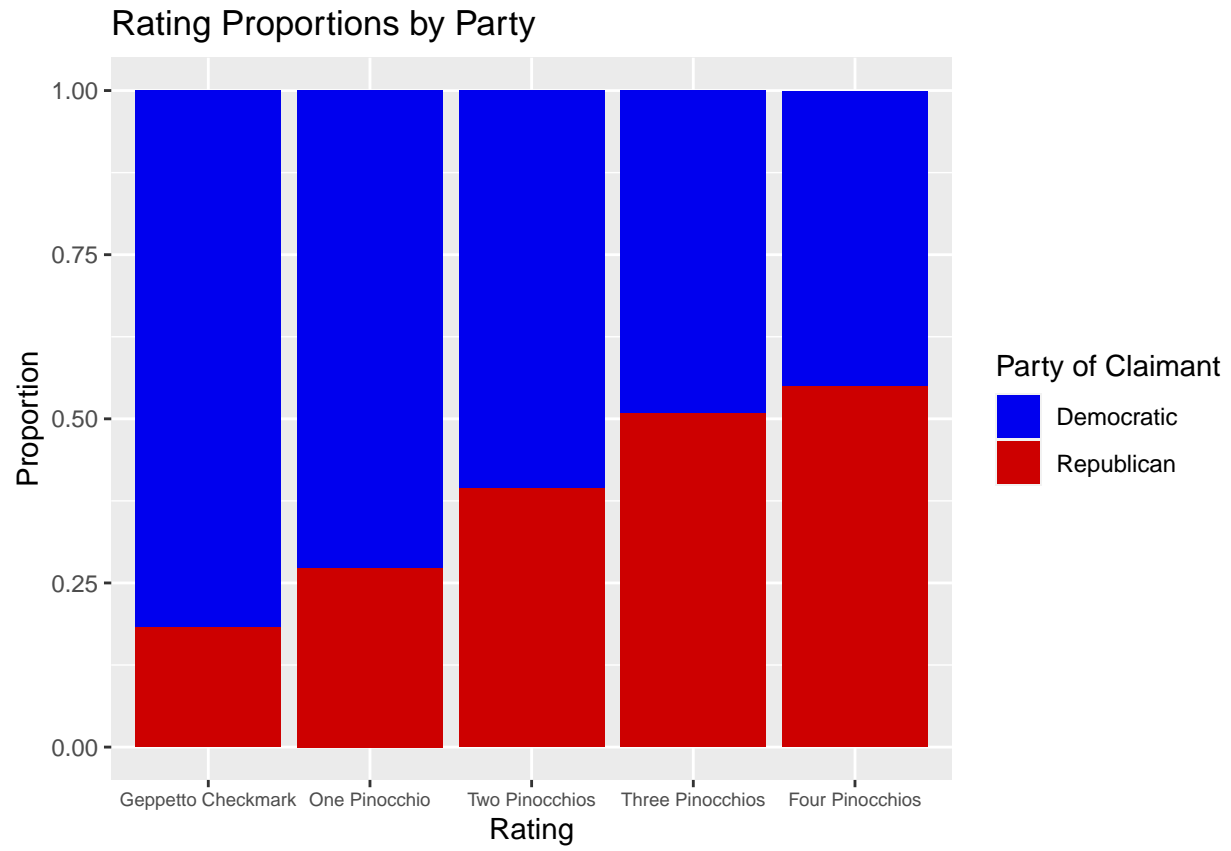
Table 3: Claim Counts by Rating

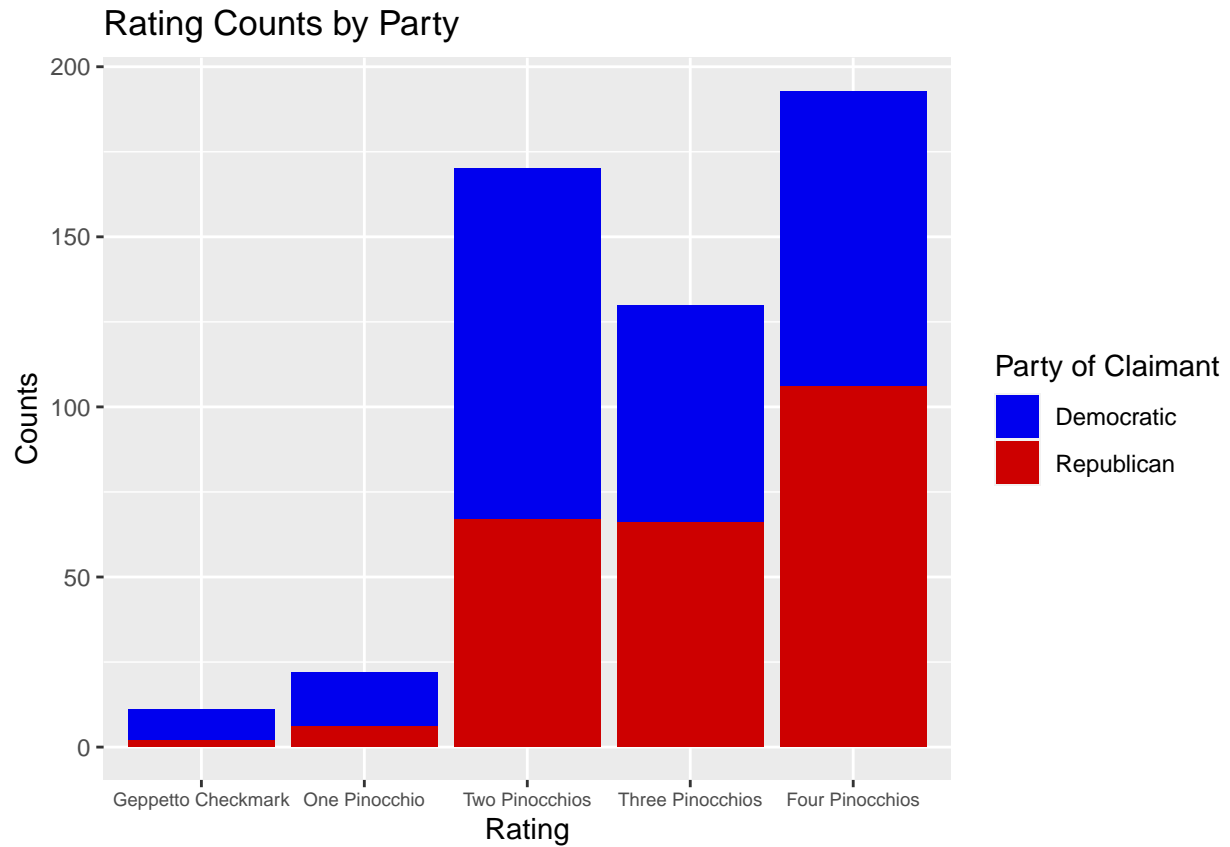
textualRating	n
Geppetto Checkmark	11
One Pinocchio	22
Two Pinocchios	170
Three Pinocchios	130
Four Pinocchios	193

Table 4: Claim Counts by Party and Rating

claimant_party	Geppetto Checkmark	One Pinocchio	Two Pinocchios	Three Pinocchios	Four Pinocchios
Democratic	9	16	103	64	87
Republican	2	6	67	66	106

Likely due to selection bias towards correcting falsehoods and/or an extremely high threshold of proof for truth, there are very few claims rated Geppetto Checkmark or One Pinocchio. As such, its hard to draw any significant conclusions from the limited data. The little information we have does appear to superficially indicate a higher ratio of Democratic truths to their Republican counterparts. There are dozens more datapoints under the Two, Three, and Four Pinocchios rating, both for Democratic and Republican politicians. By count, Democrat falsehoods exceed Republican's in the Two Pinocchios category, but they account for a lower proportion of claims rated Three and Four Pinocchios.





By proportion of total claims alone, and keeping in mind the limitations of any conclusion due to sample size, it appears that Democrats are fact-checked on more truthful claims, whereas Republicans account for a higher share of the egregious falsehoods checked by the Washington Post.

