

# Activity to explore p-hacking

XX YOUR NAMES HERE

2025-05-23

## Hack your way to scientific glory

During the course, we have learned about some of the perils of *dubious practices* such as overfitting models, making errors in extrapolation, and ignoring or misinterpreting null results. While there is not a single way to conduct an analysis, researchers need to be transparent in their decision-making process and acknowledge the limitations behind their research.

Before today's class, you were asked to read the Five Thirty Eight Article "Science Isn't Broken It's just a hell of a lot harder than we give it credit for" (<https://fivethirtyeight.com/features/science-isnt-broken>).

Today you will be working in pairs to explore the potential problems that arise when a non-ethical process is used to determine what to publish.

You will be working through the interactive visualization "Hack Your Way to Scientific Glory" provided by Five Thirty Eight (<https://projects.fivethirtyeight.com/p-hacking>), playing the role of a social scientist "with a hunch" about the association between who is in office (Republicans vs. Democrats) and the state of the economy.

An alternative location for the page can be found at <https://web.archive.org/web/20250117051842/http://projects.fivethirtyeight.com/p-hacking>.

Your mission is to come up with a set of analytic choices that lead to a "statistically significant" result (e.g,  $p < 0.05$ ).

## Step 0: appoint a leader and a scribe

- The leader will be the person whose birthday is coming up next. The leader is responsible for organizing work on the activity.
- The scribe will be the other person. The scribe is responsible for writing up the first draft of the submission and working with the leader to submit your report.

# Hack Your Way To Scientific Glory

You're a social scientist with a hunch: **The U.S. economy is affected by whether Republicans or Democrats are in office.** Try to show that a connection exists, using real data going back to 1948. For your results to be publishable in an academic journal, you'll need to prove that they are "statistically significant" by achieving a low enough p-value.

## 1 CHOOSE A POLITICAL PARTY

Republicans

Democrats

## 2 DEFINE TERMS

Which politicians do you want to include?

- ☐ Presidents
- ☐ Governors
- ☒ Senators
- ☒ Representatives

How do you want to measure economic performance?

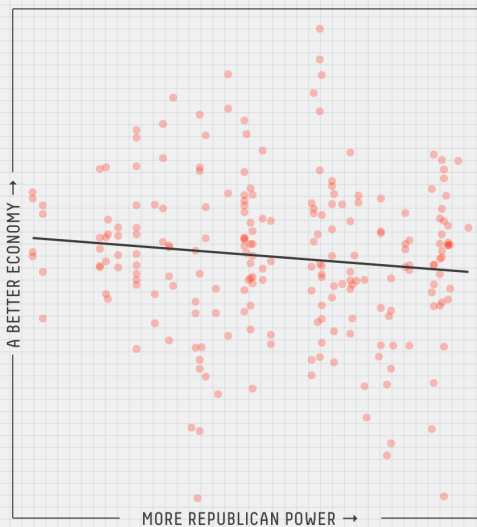
- ☒ Employment
- ☐ Inflation
- ☒ GDP
- ☒ Stock prices

Other options

- ☒ Factor in power  
Weight more powerful positions more heavily
- ☒ Exclude recessions  
Don't include economic recessions

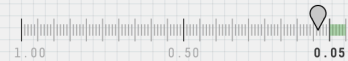
## 3 IS THERE A RELATIONSHIP?

Given how you've defined your terms, does the economy do better, worse or about the same when more Republicans are in power? Each dot below represents one month of data.



## 4 IS YOUR RESULT SIGNIFICANT?

If there were no connection between the economy and politics, what is the probability that you'd get results at least as strong as yours? That probability is your p-value, and by convention, you need a **p-value of 0.05 or less** to get published.



### Result: Almost

Your **0.09** p-value is close to the 0.05 threshold. Try tweaking your variables to see if you can push it over the line!

If you're interested in reading real (and more rigorous) studies on the connection between politics and the economy, see the work of Larry Bartels and Alan Blinder and Mark Watson.

Data from The @unitedstates Project, National Governors Association, Bureau of Labor Statistics, Federal Reserve Bank of St. Louis and Yahoo Finance.

Figure 1: Five Thirty Eight p-hacking explorer

**Step 1: Determine which party you will be exploring**

- If the leader's birthday is an ODD number, you will be fitting models for REPUBLICANS
- If the leader's birthday is an EVEN number, you will be fitting models for DEMOCRATS

**Step 2: Define terms**

Once you access the website, you'll need to make some decisions.

- Which politicians do you want to include?
- How do you want to measure economic performance?
- Other options

**Step 3: Is there a relationship?**

- What do you conclude about the outcome when more REPUBLICANS/DEMOCRATS are in office?

**Step 4: Is your result significant?****Step 5: Write up your results**

In a brief report (no more than one page, single spaced plus a figure) aimed at an audience of your classmates, summarize your model and your results. Include a screenshot of the scatterplot and the result. What do you conclude? Be sure to discuss the potential issues of p-hacking and the "Garden of Forking Paths"

**Step 6: Submit your report**