



**GESIS** Leibniz Institute  
for the Social Sciences

# Reproducible Research in Action: A Hands-On Binder Tutorial

*Leibniz*  
Leibniz  
Association

R2CASS@ICWSM, June '25

Lorraine Saju (GESIS, Köln)



# From Theory to Practice: Bringing Reproducibility to Life

- **You've Heard About:**
  - Open Science & computational reproducibility
  - Computational Environments (Jupyter, Docker, repo2docker)
  - Binder Hub
- **Today's Goal:**
  - **Launch & Interact** with fully reproducible R & Python analyses via Binder.
  - Understand the **ease and power** of sharing executable research.

# Launch Your Reproducible Environment!

## Python Demo



<https://tinyurl.com/8xv63xw7>

## R demo



<https://tinyurl.com/845y7a3s>

# Binder: Your Reproducible Research Bridge

ICWSM-Reproducibility-Python-... Public Pin Watch 0

main Go to file + <> Code

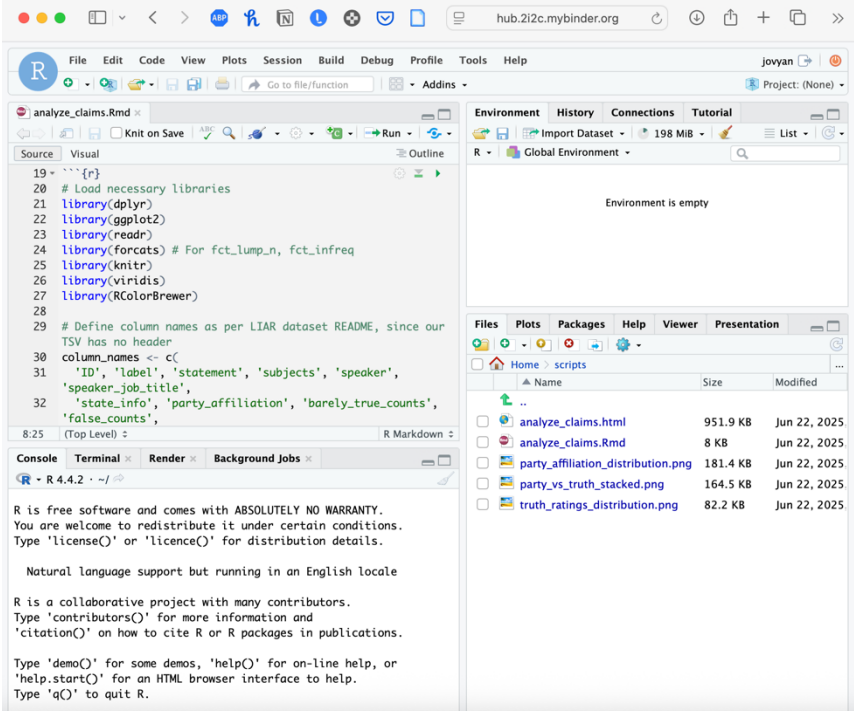
File	Description	Time
binder	Initial functional setup: Add Python...	yesterday
data	Initial functional setup: Add Python...	yesterday
scripts	Added stacked bar plot to visuali...	15 hours ago
.gitignore	Initial functional setup: Add Python...	yesterday
LICENSE	Initial commit	yesterday
README.md	Update Readme	yesterday

Your Code + Config Files (e.g., install.R, requirements.txt)



Turn a Git repo into a collection of interactive notebooks

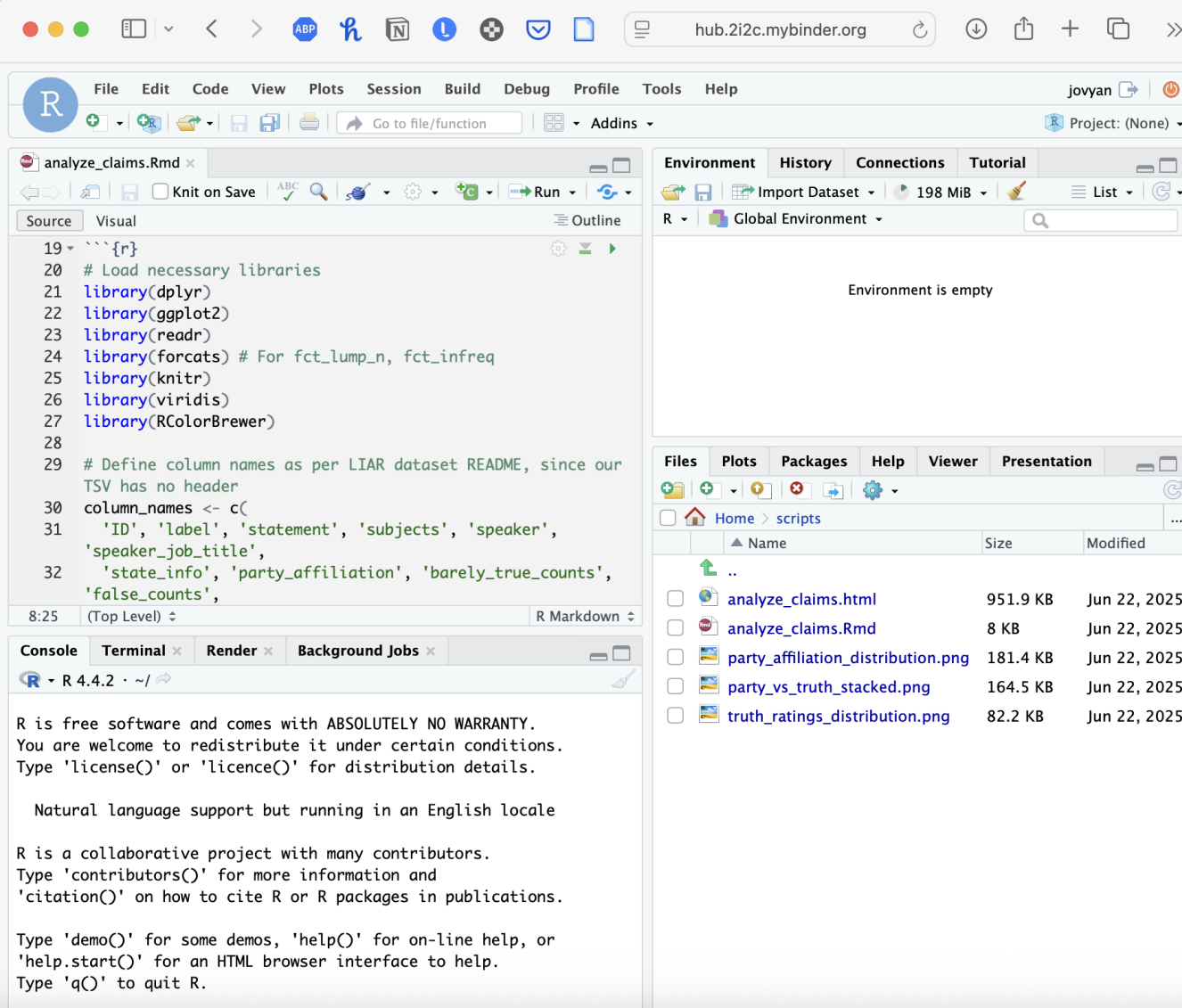
**Binder Hub (builds Docker image with repo2docker)**



**Live, Interactive Environment (RStudio / JupyterLab)**

**Result:** Zero setup, shareable, executable research for anyone, anywhere.

# R Demo: Inside the remote Rstudio instance



The screenshot shows a web browser window at [hub.2i2c.mybinder.org](https://hub.2i2c.mybinder.org) displaying a remote RStudio session. The interface includes a menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help), a toolbar, and a main workspace divided into several panes.

**Source Pane:** Displays the R script `analyze_claims.Rmd` with the following code:

```
19 {r}
20 # Load necessary libraries
21 library(dplyr)
22 library(ggplot2)
23 library(readr)
24 library(forcats) # For fct_lump_n, fct_infreq
25 library(knitr)
26 library(viridis)
27 library(RColorBrewer)
28
29 # Define column names as per LIAR dataset README, since our
   TSV has no header
30 column_names <- c(
31   'ID', 'label', 'statement', 'subjects', 'speaker',
   'speaker_job_title',
32   'state_info', 'party_affiliation', 'barely_true_counts',
   'false_counts',
33 )
```

**Environment Pane:** Shows the Global Environment, which is currently empty.

**Files Pane:** Displays a file explorer view of the `scripts` directory. The files listed are:

Name	Size	Modified
analyze_claims.html	951.9 KB	Jun 22, 2025
analyze_claims.Rmd	8 KB	Jun 22, 2025
party_affiliation_distribution.png	181.4 KB	Jun 22, 2025
party_vs_truth_stacked.png	164.5 KB	Jun 22, 2025
truth_ratings_distribution.png	82.2 KB	Jun 22, 2025

**Console Pane:** Shows the R version (R 4.4.2) and the following output:

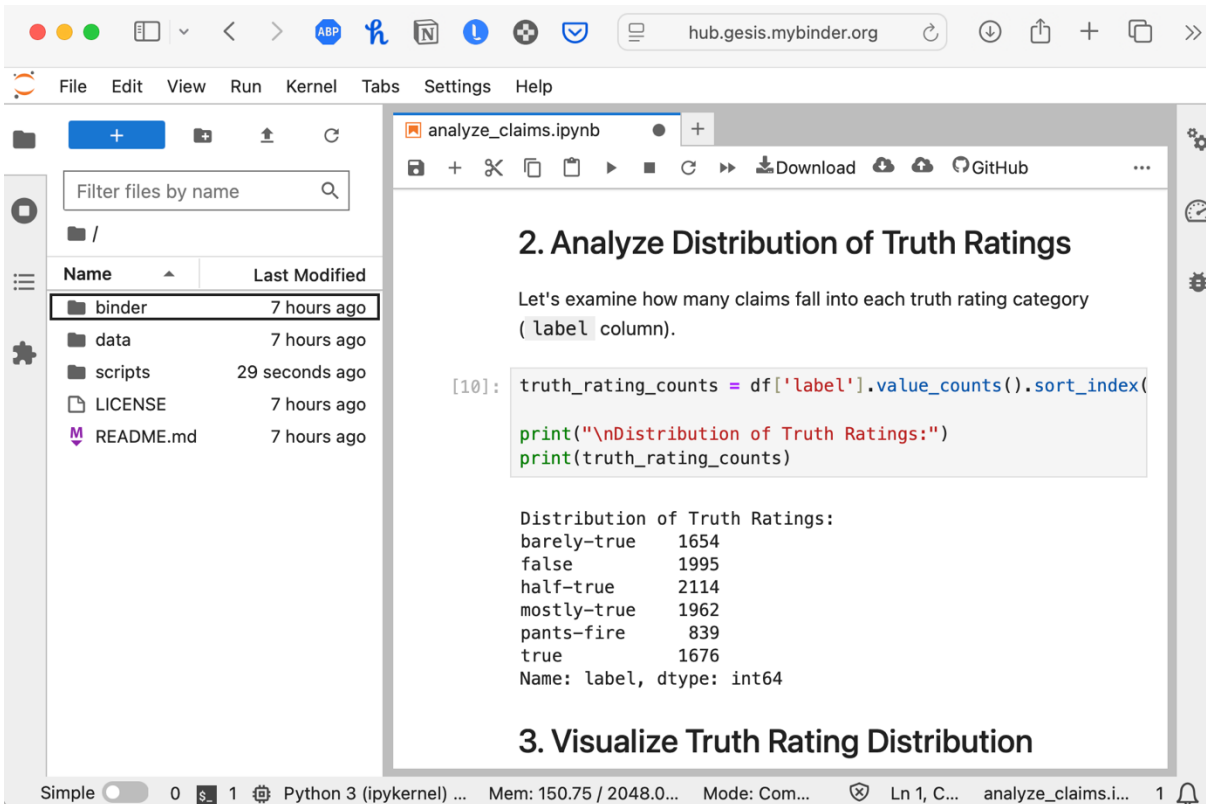
```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

# Python Demo: Unpacking Politician Claims in JupyterLab



hub.gesis.mybinder.org

File Edit View Run Kernel Tabs Settings Help

Filter files by name

Name Last Modified

- binder 7 hours ago
- data 7 hours ago
- scripts 29 seconds ago
- LICENSE 7 hours ago
- README.md 7 hours ago

## 2. Analyze Distribution of Truth Ratings

Let's examine how many claims fall into each truth rating category (label column).

```
[10]: truth_rating_counts = df['label'].value_counts().sort_index()
print("\nDistribution of Truth Ratings:")
print(truth_rating_counts)
```

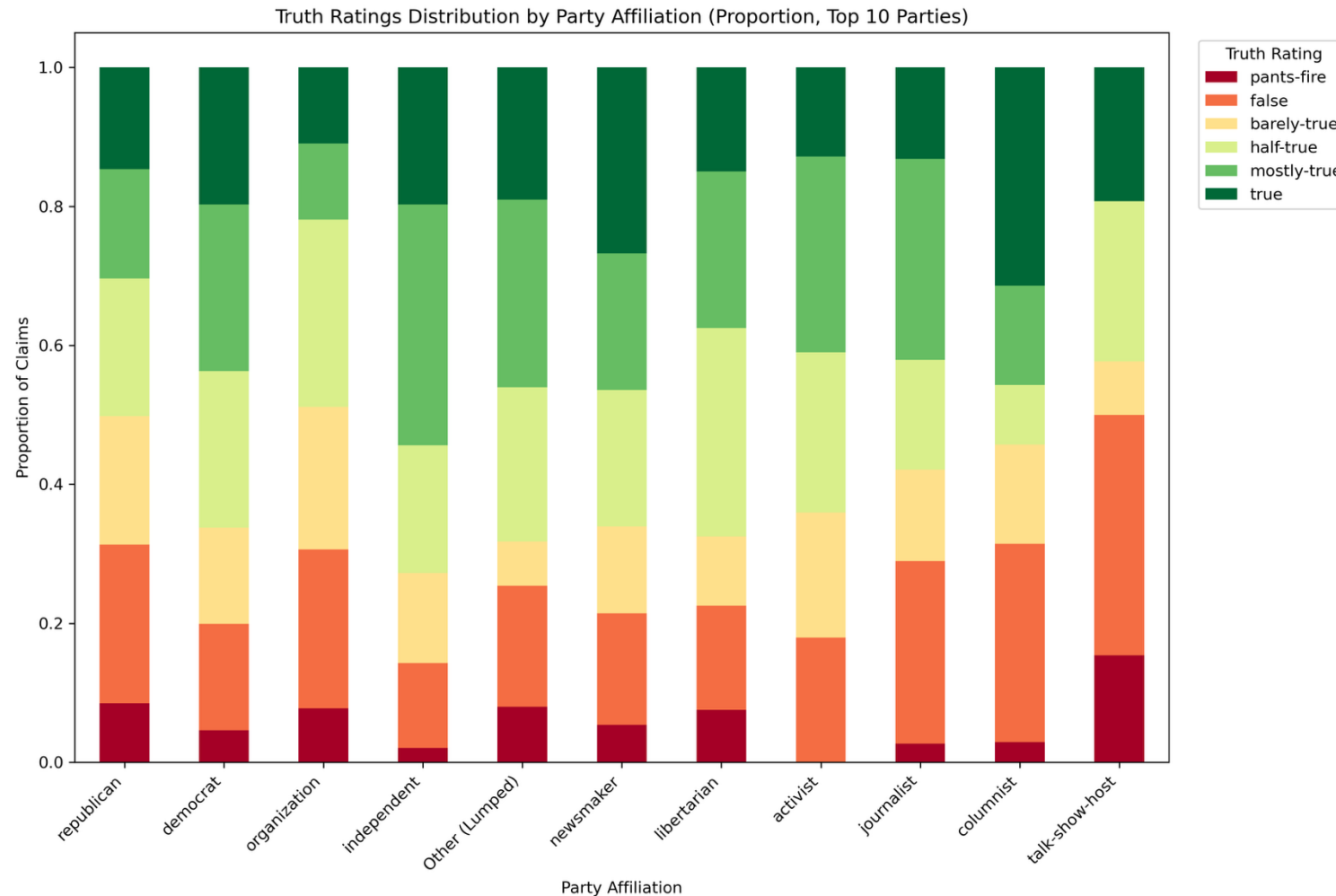
Distribution of Truth Ratings:

barely-true	1654
false	1995
half-true	2114
mostly-true	1962
pants-fire	839
true	1676
Name: label, dtype: int64	

## 3. Visualize Truth Rating Distribution

Simple 0 1 Python 3 (ipykernel) ... Mem: 150.75 / 2048.0... Mode: Com... Ln 1, C... analyze\_claims.i... 1

# Python Demo: Politician Claims Analysis Walkthrough



# Key Takeaways: Reproducibility at Your Fingertips

- **Binder's Power:**
  - **Zero Setup:** No local installations needed.
  - **Interactive:** Run, modify, explore code live.
  - **Reproducible:** Ensures consistent results across environments.
  - **Shareable:** A single link to your executable research.
- **Impact for Research:**
  - Increases transparency & trust
  - Accelerates collaboration
  - Lowers barriers to engagement



# Your Next Steps: Binderize Your Own Project!

Steps	Python	R
1. Organise your code: Clear folders, meaningful names, use of relative paths.	-	-
2. Define environment: Document language version List external libraries	Python version -> runtime.txt Dependencies: requirements.txt (e.g., pandas==2.0.3) (or environment.yml for Conda)	sessionInfo() -> runtime.txt Dependencies: install.R (e.g, inst all.packages("tidyverse"))
3. Push to <b>GitHub</b> and make your repo public		
4. Go to <b>mybinder.org</b> and paste your repository URL		
5. Copy your binder badge		

# Further Resources

- **Explore Our Demos:**
  - <https://tinyurl.com/8xv63xw7>
  - <https://tinyurl.com/845y7a3s>
- **Learn More:**
  - [mybinder.org](https://mybinder.org) (Official Binder site)
  - **repo2docker Documentation:**
    - General: [repo2docker.readthedocs.io/en/latest/](https://repo2docker.readthedocs.io/en/latest/)
    - Configuration Files: [repo2docker.readthedocs.io/en/latest/config\\_files.html](https://repo2docker.readthedocs.io/en/latest/config_files.html)
  - [The Turing Way \(Guide to Reproducible Research\)](#)
  - BERD Reproducible Research Course: [berd-nfdi.github.io/BERD-reproducible-research-course/](https://berd-nfdi.github.io/BERD-reproducible-research-course/)

# Thank You For Your Attention



## Questions?