

APPROVAL RATING DASHBOARD FOR U.S. PRESIDENTS

Project Proposal

Mia Silva Louise Oh Lidio Jaimes
Oliver Brown

2025-11-13

Table of contents

Research Question	2
Existing Work	2
Work Outline	3
The Data	4
Programming Paradigms	4
Software and Packages	4
Our Product	5
Timeline	6
Task Distribution	6

Github Repo Link

Project 4 GitHub Repo

Research Question

To what extent can a predictive model, using key economic indicators (like inflation and unemployment) and major political event data, accurately forecast near-term U.S. approval ratings? Furthermore, how do simulated changes in these economic indicators proportionally impact future approval scenarios?

This research question seeks to quantify the economy's role in a leader's public perception, while also examining the time lag before economic shifts actually impact approval ratings. Furthermore by building a simulative model, it aims to create a practical, forward-looking tool for testing "what-if" scenarios, moving beyond simple historical analysis to show how hypothetical future economic conditions could shape public opinion.

Existing Work

Past work on presidential approval is in two main areas. The first is public-facing data dashboards. The most famous was Nate Silver's (and later owned by ABC) FiveThirtyEight, which was a statistical product in itself. Its "work" was running a proprietary weighted-averaging model on polls, giving more credibility to pollsters with better historical accuracy. Other dashboards, like The Economist's approval tracker, RealClearPolitics, and Gallup's Presidential Job Approval Center, serve as aggregators or historical archives. These present demographic breakdowns or simple rolling averages to show how support differs by group and to track the most current trends.

The second area is the academic statistical modeling that provides the blueprint for these predictions. This research, from pioneers like Edward Tufte and Ray Fair, established that approval is not random. It is overwhelmingly driven by the economy (like inflation and GDP growth). These models evolved to include other key, non-economic variables. For instance, John Mueller's work identified the "rally 'round the flag" effect (a temporary spike in approval after a crisis), while Alan

Abramowitz's "Time for a Change" model quantified the predictable "erosion of support" that a president experiences over time.

Work Outline

- Data Gathering

This phase involves fetching necessary time-series data, including monthly economic indicators like CPI and unemployment from the BLS API, and scraping poll-by-poll approval ratings from The American Presidency Project and the National Election Studies.

- Data Cleaning

We then use dplyr to process the raw, separate files. The primary challenge will be joining the existing approval data with the economic data in order to create a single, aligned time-series dataset.

- Model Training

Using the clean, aligned historical dataset, we will train a supervised machine learning model to find the mathematical relationship between the economic variables and approval ratings. This trained model object will be saved as an .rds file so the Shiny app can load it for making predictions.

- Building the Dashboard

This phase involves creating the R Shiny application, using a ui.R file for the layout and a server.R file for the logic. The server will load the saved .rds model and use reactive expressions to instantly update forecast charts based on user inputs.

- Adding Additional Features if Time Allows

If the core product is finished, this time would be used to enhance the model, such as by incorporating "major political events" data. We could also expand the dashboard's "Historical Data" page by adding more economic variables for the user to explore.

The Data

Economic data will primarily come from the Bureau of Labor Statistics API, which provides key predictive features like monthly inflation and unemployment. Historical and current presidential approval rates will be sourced from The American Presidency Project. This poll data will be aggregated into a monthly average to align with the economic data.

Programming Paradigms

This paradigm is used for all data cleaning and feature engineering, primarily through R’s dplyr package. This approach allows for a clean, readable, and reproducible data pipeline. We use a chain of functions to transform data, such as group_by() and summarise() to resample daily polls into monthly averages, left_join() to merge approval data with economic tables. Each step takes a data frame and returns a new data frame, which is “piped” to the next function.

- Supervised Machine Learning

This is the paradigm used for the “Prediction and Modelling” page. We treat the time-series forecast as a supervised learning problem. This involves training a predictive model on a “labeled” historical dataset. In this context, the features are the historical economic indicators (like inflation and unemployment). The label is the corresponding presidential approval rating for that month. The model learns the mathematical relationship between these economic features and the approval rating, allowing it to predict future ratings when given new economic data.

Software and Packages

- Tidyverse (data wrangling and visualization)
- Tidymodels (ML)

- RShiny
- ML engines (e.g. glmnet, ranger, kknn, kernlab)
- DT, knitr (for tables)

Our Product

Dashboard

- Home Page (Summary of Product/Tutorial)
- Historical Data

Here, the user can select a president of interest and view the approval ratings over the course of their time in office. Moreover, the user can also select variables of interest such as inflation, cost of healthcare, median salary, etc.

The dashboard will then display these in multiple. The top panel will show the president's approval rating, and separate panels will appear beneath it for each selected economic variable. All charts will share the same x-axis (time), allowing the user to easily compare trends vertically across the different metrics, each with its own clear and independent y-axis.

- Prediction and Modelling

This page will focus on applying a pre-trained statistical model to recent data to generate a near-term forecast for the current president. It will display a chart showing the model's prediction for the next 3-6 months, which is calculated by feeding the latest available data into the model.

The core of this page is the interactive simulation tool. Here, the user can adjust sliders corresponding to key economic variables (for example, "Set future inflation to...") to run their own "what-if" scenarios. As the sliders are moved, the model will instantly re-calculate and display the proportional impact of these hypothetical changes on the president's approval forecast. This section will also briefly show the model's key drivers, highlighting which features, such as unemployment or inflation, the model found to be the most significant predictors.

- Reference Page

Timeline

- November 19, Wednesday: finalize project plan
- November 21, Friday: draft dashboard template
- Week of Thanksgiving - everyone works on their portion of the dashboard
- December 1, Monday: final draft of dashboard
- December 8, Monday: finalize presentation pdf
- December 9, Tuesday: draft of project write up
- December 10, Wednesday: project presentation DUE
- December 16, Tuesday: final draft of project write up
- December 18, Thursday: project write up DUE

Task Distribution

- Home Page & Summary of Product (Everyone)
- UI (Mia)
- Prediction and Modelling (Oliver and Mia)
- Historical Data and Interactive Approval Rating Plots (Louise and Lidio)
- Reference Page (Everyone)