Computational Social Science

Final Projects & Shiny Applications

Dr. Thomas Davidson

Rutgers University

February 22, 2024

Plan

1. Course updates
2. Final projects
3. Developing Shiny applications in R

Course updates

- Homework 2 released
  - Github Classroom link in Canvas Module
  - Topics: tidyverse + ggplot2, APIs, and webscraping
  - Deadline: Next Wednesday, 2/28, 5pm ET

Final projects

Goals

- Build an interactive data visualization app, focused on a topic of social scientific interest
- Showcase skills and knowledge developed in this class
- Add to your portfolio or resume

Final projects

Five steps

1. Data collection
2. Data cleaning
3. Data analysis
4. Data visualization
5. Building and deploying an app

Final projects

Data collection

- Collect a dataset related to social science
- Data sources:
  - APIs
  - Webscraping
  - Existing datasets (e.g. General Social Survey, American National Election Study)

Final projects

Data cleaning

- Process dataset to extract relevant data
  - String pre-processing (more next week)
  - Parsing HTML
  - Merging datasets
  - Selecting, grouping, mutating, etc.

Final projects

Data analysis

- Conduct an analysis of the dataset. Possibilities include
  - Data summaries
  - Descriptive statistics (mean, median, mode, etc.)
  - Statistical tests (correlation, t-test, chi-squared, etc.)
  - Statistical modeling (regression, machine learning, topic modeling, etc.)

Final projects

Data visualization

- Construct visualizations of the dataset
  - Relationships between variables
  - Trends over time
  - Animations
  - Maps
- *All apps must have an interactive component*

Final projects

Building and deploying an app

- Use R Shiny to build an interactive app
- Deploy the app on the web using `Shinyapps.io`

Final projects

Final deliverables

1. A deployed Shiny app for interactive data visualization
2. An organized Github repository for the project
3. Documentation and a short written description of the app

Final projects

Milestones

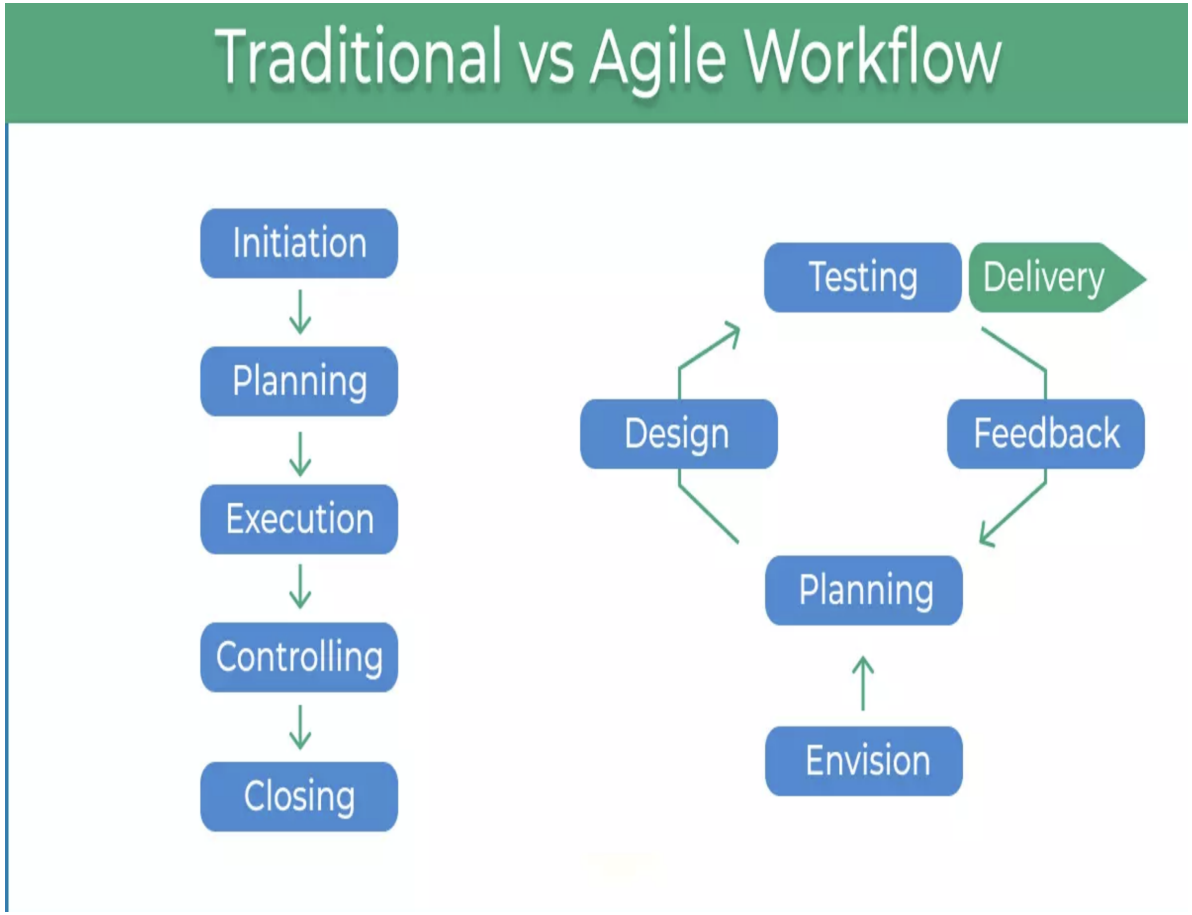
1. Identify a suitable topic
   - *All projects are subject to approval*
2. Initial proposal (due 3/6)
   - A short description of the planned project including plan for data collection, cleaning, analysis, and visualization
3. Prototype (due 4/10)
   - A basic working prototype of your app
4. Presentation (4/29)
   - Short demo and discussion in class
5. App and deliverables submitted (5/6 at 5pm)

Final projects

Feedback

- There will be several opportunities for formal and informal feedback
  - Office hours
  - Discussions with classmates
  - Feedback on proposal and prototype
  - Presentation

Final projects



Final projects

Teamwork

- You can work on the project individually or as part of a team (maximum 4)
- Make a decision when you submit your proposal
- Teams will be require to submit an additional document describing contributions of each member
  - Ensure fair distribution of work
  - Git commit history to highlight contributions

Developing Shiny applications in R

What is Shiny?

- Shiny is a package you can use to build interactive web pages directly from R
  - It does not require any experience with HTML, CSS, or Javascript, etc.
- Apps can be hosted on standalone webpages, enabling anyone to access them
- Many extensions available, making Shiny a powerful tool for data visualization and construction of dashboards.

Developing Shiny applications in R

Example 1: Explore your weather

- A simple app to show data on eruptions of Old Faithful, a geyer in Yellowstone NP

Developing Shiny applications in R

Example 2: Bloodbanks in India

- This app maps bloodbanks in India using data from the Open Government Platform
- Interactive, clickable map
- Tabs to show raw data and provide further details

Developing Shiny applications in R

More examples

- There are many more examples in the Gallery on the Shiny website.

Developing Shiny applications in R

Extensions

- There are many extensions of the Shiny framework
  - D3 for data-driven visualization. See 'r2d3' website
  - Leaflet for interactive mapping
  - Plotly for interactive plots
- This Github page lists a ton of Shiny resources.

Developing Shiny applications in R

Further readings

- RStudio's Shiny cheatsheet is a really useful quick reference.
- *Mastering Shiny* by Hadley Wickham is available for free online.
- RStudio has a tutorial.
- Lots of videos on YouTube

Developing Shiny applications in R

Deployment

- Apps can be deployed for free using the Shiny Apps website.* You just need to do the following:
  - Sign up (you can use Github account)
  - Link to RStudio using credentials
  - Package app in appropriate format
  - Deploy
- Free accounts are limited to 5 apps and 25 active hours a month. You will need to ensure you do not exceed these limits.

Developing Shiny applications in R

Information on example files

1. `app.R` is an example of simple Shiny app to interactively plot correlations, along with other descriptive statistics
2. `app2.R` modifies the previous app to use the `shinydashboards` framework
3. `/app` is a directory containing two files, `ui.R` and `server.R` (the components of `app.R`)

- `test.R` uses the files in `/app` to launch a local version of the app
- `app_deploy_template.R` can be run to deploy the app to shinyapps.io (or this can be done via point-and-click interface)
  - You must add credentials and other information to this template before running it.

Next week

- Introduction to Natural Language Processing