

Project 3 (Final)

Creating a Shiny App

The goal of this project is to create a nice looking shiny app that can be used to explore data and model it.

Project Work

All project work should be done in a github repo. Ideally, you have connected RStudio with github and can work from the command line within RStudio. All major updates should be made through github so we can track your activity.

We will run your app from gitHub using RStudio so make sure it is set up correctly to do that (see lecture for reference to setting it up, try it yourself to make sure it works!).

Be sure to have a section on your readme page that includes all packages required for the app to work. You should have a line that would install all the packages used (so we can easily grab that and run it prior to running your app).

You should also include a line of code on the readme that we can copy and paste that will run your repo from RStudio. You should check that this works with an “empty” version of RStudio (that is, one that doesn’t have objects you’ve already created existing in your environment). You want to make sure anyone can run the app using the code!

Find a data set you are interested in

For this project I’m going to let you choose your own data set. You can either pull in a data set via a file or you could pull data from the web within your app (for instance baseball data ([link](#)) or twitter data ([link](#))).

Choose something you are interested in and have ideas for investigating!

App Requirements

- Your app should have multiple pages (tabs) to it. I don’t care if you use the built in tabs for shiny or a package like shinydashboard - use the method you prefer.
 - An information page that describes the data, the purpose of the app, and how to navigate it
 - A data exploration page where common numerical and graphical summaries can be created by the user. The user should be able to change variables and subset the data in some way.
 - A page with either clustering (include a dendrogram) or principal components analysis (include a biplot) - again where the user can specify aspects of the method.
 - A page for modeling - see below for details
 - A page that allows the user to scroll through the data, subset it, and save the data as a file (.csv is fine but whatever you’d like)
- You should have at least two different dynamic UI elements.
- You should have a button that allows the user to save a plot they are viewing to a file.
- You should utilize the ability to click on a plot or select a region in some way (see the `plotly` package if you’d like to use that functionality)
- You should include some type of math type (maybe an equation or just a special symbol you need to use `mathJax` for).
- You should include a link to something and some other formatted text.
- Modeling

- (At least) two supervised learning models (feel free to branch out to things we didn't discuss if you'd like)
- You should give the user some functionality for choosing model settings (variables used, number of trees, etc.) and for changing relevant output
- You should give the user a way to use the model for prediction (they should be able to select the values of the predictors).

Submission

You should simply post your github repo URL as your submission. You don't need to use github pages on this one unless you want to.

Notes

- There are some nice shiny additions such as a status bar that you can add to display to the user that your model is running.
- Remember that most people that have apps make their code freely available (often on gitHub). One thing you might do is search through apps to find someone that say has the option to download the data set. Check their code on gitHub to see how they did it and if it makes sense for you. The openness of the R community is really one of the great benefits of using R!
- This project is pretty open ended! Have fun with it and make something that you can show off to others - Good luck :)

Rubric for Grading (total = 100 points)

Item	Points	Notes
Working Tabs	5	Worth either 0 or 5
Information page	10	Worth either 0, 5, or 10
Data exploration page	20	Worth either 0, 5, ... or 20
Unsupervised learning page	10	Worth either 0, 5, or 10
Data page with subsetting	5	Worth either 0 or 5
Dynamic UI elements	10	Worth either 0, 5, or 10
Save plot & data	5	Worth either 0 or 5
Plot click or hover	5	Worth either 0 or 5
Math type/Link/formatted text	5	Worth either 0 or 5
Supervised Learning models & Prediction	25	Worth either 0, 5, ..., or 25

Notes on grading:

- For each item in the rubric, your grade will be lowered one level for each each error (syntax, logical, or other) in the code and for each required item that is missing or lacking a description.
- **If your work was not completed and documented using your github repo you will lose 50 points on the project.**
- You should use Good Programming Practices when coding (see wolfware). If you do not follow GPP you can lose up to 25 points on the project.
- You should use appropriate markdown options/formatting (you can lose up to 20 points) for not doing so