

Assignment 2

Due 2/20/2022 10pm

Due February 20th 10pm

Submit by uploading to GitHub classes submission link: <https://classroom.github.com/a/uem2UTaU>

Assignment 2 Goals:

By the end of assignment 2 you will have the ability to build a Shiny application that can:

- Load in an external .csv
- Save results within your app
- Use ActionButtons to trigger events
- Fit several statistical models within Shiny

George will present a demonstration on Monday along with some a lecture on the new UI and steps to saving output within the app.

You should have enough information to work on the data uploading step. Wickham Chapter 9 is a good resource for data uploads.

Assignment 2 Directions

There are two versions of this assignment. Version 2 is required for all students. Version 1 is optional but we will provide feedback to you if you submit it.

Version 1

Version 1 is an application that will compare 2 separate regressions fit with the same dataset. The application will require the following steps:

1. Build the ability to load in an external data set (stored as a .csv)
2. After your dataset has been uploaded, develop a UI element (dropdowns, checkboxes or a drag/drop) that will allow you select a single outcome variable and select 2 sets of predictors (one set for each of the two models) predictor variables to be used to fit two separate regression analyses.
3. After fitting the two models, show AIC for each of the models. If the two models are nested (one model is the same as the other but includes an additional variable(s))

In R you can use the `anova()` function to perform this test:

```
# example of a nested model
library(ggplot2)
```

```
## Warning in register(): Can't find generic 'scale_type' in package ggplot2 to
## register S3 method.
```

```
m1 <- lm(price ~ carat, diamonds)
m2 <- lm(price ~ carat + cut + color, diamonds)

anova(m1, m2)
```

```
## Analysis of Variance Table
##
## Model 1: price ~ carat
## Model 2: price ~ carat + cut + color
##   Res.Df      RSS Df Sum of Sq    F    Pr(>F)
## 1   53938 1.2935e+11
## 2   53928 1.1061e+11 10 1.8731e+10 913.21 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# an example of non nested models
m1 <- lm(price ~ carat, diamonds)
m2 <- lm(price ~ cut, diamonds)

# you can use the AIC command to get AIC
AIC(m1)
```

```
## [1] 945466.5
```

Version 2

Version 2 will be required it too is a regression app but rather than fitting 2 “side by side models” this app will have the ability to fit models one at a time and record the results.

Your application will require you to:

1. Build the ability to load in an external data set (stored as a .csv)
2. Select an outcome and predictor variable for **one** model
- 3 fit this model and save the results
4. Build a dropdown were you can access the results of all saved models
5. results should include AIC and general the general regression summary

Bonus This is completely optional but if you would like to make a more advanced app add the option to perform leave one out cross validation.

Data

If you have a particular dataset you are working with, feel free to use it! I have also included the `WaffleDivorce` dataset as a .csv. The outcome for this dataset the divorce rate. The data set consists of 50 rows, wach row represents a state.