Reproducible Research in R, Lesson 4 - Completed Notebook

Christy Garcia, Ph.D. and Christopher Prener, Ph.D.

(October 17, 2018)

Table of Contents

## Introduction

This notebook extends our use of knitr and other reproducible research techniques discussed in the first three lessons.

## Dependencies

This notebook requires the following packages:

# tidyverse packages  
library(ggplot2) # data plotting  
library(magrittr) # pipe operator  
library(readr) # read csv files  
  
# other packages  
library(here) # file path management

## here() starts at /Users/chris/GitHub/DSS/research-04

library(janitor) # frequency tables  
library(knitr) # support for document knitting

## Load Data and Modify Output

This notebook requires two data files from the data/ folder, the mpg.csv data and the starwars.csv data.

auto <- read\_csv(here("data", "mpg.csv"))

star <- read\_csv(here("data", "starwars.csv"))

*The code chunks both include results='hide' syntax, that prevents their output from appearing in your knit document. The code, however, will appear!*

If there are specific messages, like read\_csv() produces, these will be hidden as well.

## Inline Information

We can embed statistics, like the average highway fuel efficiency (23.4401709). We can round this to two digits so we get a shorter output using the round() function - 23.44.

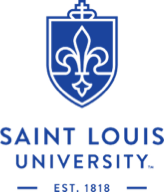
## Modifying Output with Images

There are two ways to include images. First, we can use pre-made images in our documents. For example, we could embed the SLU DSS logo. By including echo=FALSE, we supress the code from being included in our output document!



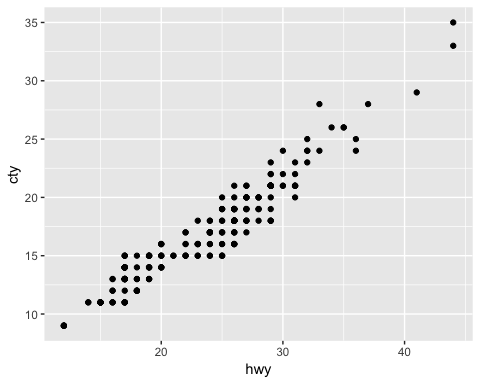
DSS Logo

We could also add the SLU logo itself.

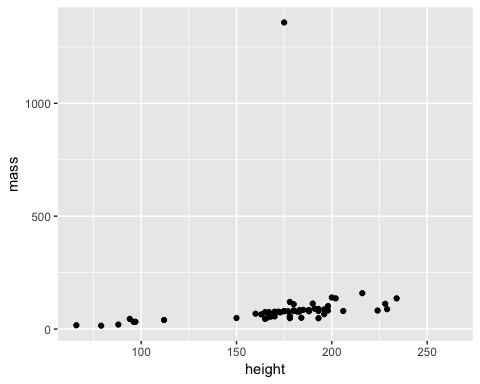


SLU Logo

Or we could add a plot:



Or another plot:



This last plot adds a warning to our document, which we can supress with warning=FALSE.

## Adding Tables

If we want to create a simple frequency table, we can do that with the janitor package:

## class n percent  
## 2seater 5 0.02136752  
## compact 47 0.20085470  
## midsize 41 0.17521368  
## minivan 11 0.04700855  
## pickup 33 0.14102564  
## subcompact 35 0.14957265  
## suv 62 0.26495726

By using the kable function, we can improve the display of this table:

|  |  |  |
| --- | --- | --- |
| class | n | percent |
| 2seater | 5 | 0.0213675 |
| compact | 47 | 0.2008547 |
| midsize | 41 | 0.1752137 |
| minivan | 11 | 0.0470085 |
| pickup | 33 | 0.1410256 |
| subcompact | 35 | 0.1495726 |
| suv | 62 | 0.2649573 |

This can also be used for arbitrary information:

|  |  |
| --- | --- |
| name | program |
| Chris | Sociology |
| Christy | Spanish |

## Other formats

In order to switch to other formats, you need to edit the YAML header of your .Rmd file. You can choose from a variety of [document](https://bookdown.org/yihui/rmarkdown/documents.html) and [presentation](https://bookdown.org/yihui/rmarkdown/presentations.html) formats.

We’ll quickly adjust this notebook to knit as a word document instead of html notebooks and .md output files.

Change the output: line of your YAML to look like this:

output: word\_document

When you knit, the word document will be created instead of the previous output types.

## Adding a Table of Contents

If we want to add a table of contents to this document, we can edit the YAML like so:

output:  
 word\_document:  
 toc: true  
 toc\_depth: 2

## Citing References

To cite references, you need four components. First, you’ll need a .bib file [or another bibligoraphy format](https://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html) saved in your notebook/ directory. We’ve provided a sample .bib file, which uses LaTeX bibligoraphy styles that can be obtained from Google Scholar to populate our bibliography.

Second, you need to have edited the notebook header to include a bibligography reference:

bibliography: bibliography.bib

Third, you need to include an in-text reference. For example, we could discuss the book *R for Data Science* (Wickham and Grolemund 2016). We use the brackets and @ to indicate that we want to make an in-text citation. We can also talk about an author’s work, like Xie’s recent work (2018). Or we can talk about an author’s work on a specific page, like Wickham and Grolemund (2016, 62).

Fourth, we need to add a final heading at the bottom of the file titled ## References or ## Works Cited.

## Works Cited

Wickham, Hadley, and Garrett Grolemund. 2016. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. " O’Reilly Media, Inc.".

Xie, Yihui, JJ Allaire, and Garrett Grolemund. 2018. *R Markdown: The Definitive Guide*. CRC Press.