CS&SS 321 - Intro to R Markdown

Andrew Yuanzhou Wang

The YAML Header

YAML means "Yet Another Markup Language". You can use this section to customize or add metadata, such as: - `date` - `abstract`

You can also add formatting under `output: mainfont`: Set the main font. `fontsize`: 11pt: Adjust the base font size. `header-includes`: Include custom LaTeX packages or commands (e.g. for advanced header/footer customization).

The setup chunk

You should reserve this for global chunk options, loading libraries/packages, and if needed clearing the environment.

Chunk options can be global (applies for your entire document) or local (applies only to a specific chunk). To set global options you can modify the \texttt{opts_chunk$set()} parameters. - `echo = TRUE` shows your code chunks in the final document. - `cache = TRUE` tells your machine to save code output as you work, which can be useful if you are slowly working through a large markdown file with lots of code.

You can load libraries in the setup or elsewhere, but it's generally best to load everything at the top. - Install packages with `install.packages("packagename")`. These usually only need to be installed once until R gets an update and you have to reinstall all of them again. - Load libraries with `library(libraryname)`. You must install the package containing the library before you can do this.

It is vital that you set your directory correctly. All data should be saved to this directory or the same directory as your script or markdown file. To do so, you can use `setwd([directorylocation])`in this chunk.

Code Chunks

You can embed code like below. You can customize a few options in the header:

- `include = FALSE` hides code and results from the finished file.
- `echo = FALSE` hides code but not results.
- `message = FALSE` hides messages generated by code.
- `warning = FALSE` hides warnings.

Defaults for the above are all `TRUE`.

Activity: Make a cat by customizing chunk options.

```r
cat("  /\\_/\\  (\n ( ^.^ ) _)")
message("   \\\"/  (\n ( | | )")
##    \"/  (
## ( | | )
warning("   /  (\n | )")
cat("(__d b__)")
```

Other formatting

Header level is indicated by the number of hashtags.

`fig.[someoption]` allows you to format figures, and `results = '[someoption]'` allows you to format results. Obviously replace "`someoption`" with the relevant option.

To add math notation, embed Latex math with "...". Some math environments, like `align` do not need the dollar signs and work on their own.

You can change font options for some specific words to `monospace`, **bold**, **different bold**, or *italic*.

Trying it out

Let's make sure that we have everything set up correctly. Please download the file here and put it in the correct directory.

Now let's make a plot. To do so, we need to learn two things.

- `variablename <- value` will assign `value` to `variablename`
- `read.csv(filename.csv)` will return the contents of `filename.csv` as a dataframe.
- `plot(x,y)` will create a scatterplot.

Don't worry about what a dataframe is yet.

```r
df <- read.csv("uscrowdsynth.csv")
plot(df$income, df$monthtotal)
```
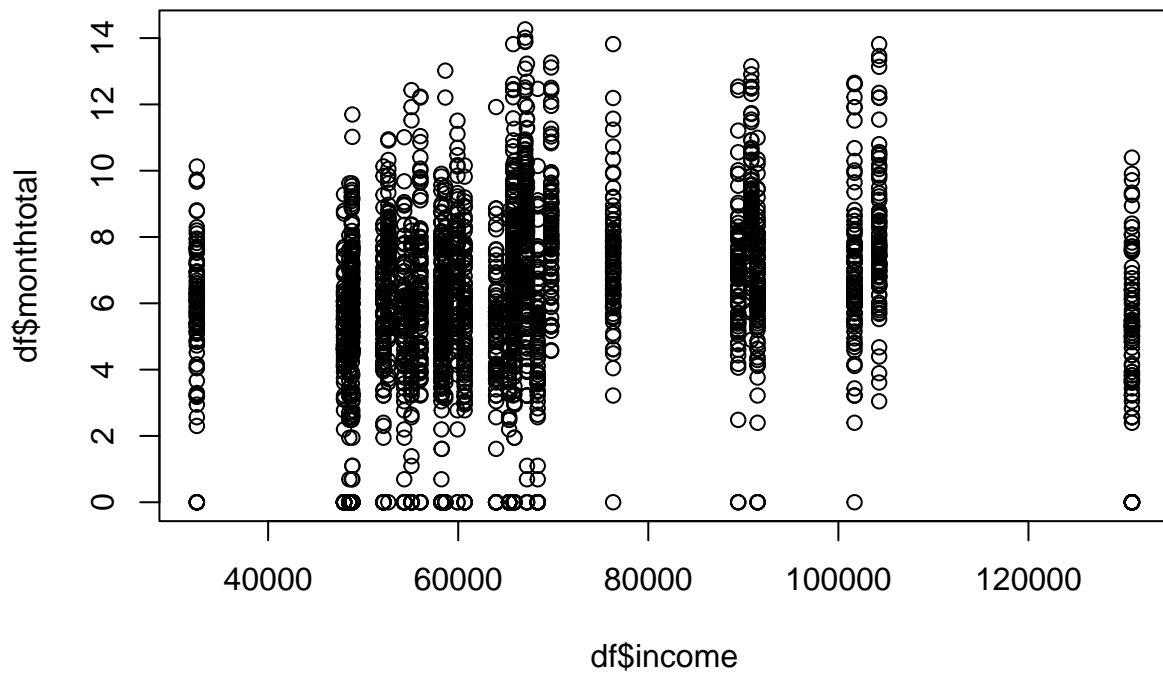
Now click "Knit" to turn it into a PDF. It should work.