Assignment #1

Your Name

2024-11-22

Contents

1 Housekeeping

I typically create a *Housekeeping* section where I install packages, create functions, and load data.

1.1 Packages

```
## Load packages
library(broom)
library(conflicted)
library(kableExtra)
library(knitr)
library(tidyverse)

## Prefer over all other packages
conflict_prefer("filter", "dplyr")
```

1.2 Functions

```
## Create function that doubles the input digit
double_it <- function(number){
  double = number * 2
  return(double)
}
```

1.3 Data

```
## Load data
df <- read_csv("data/favorability_polls.csv")
```

2 Exploratory data analysis

Notice that you can create headers in Quarto documents. Quarto will use these headers to print the table of contents.

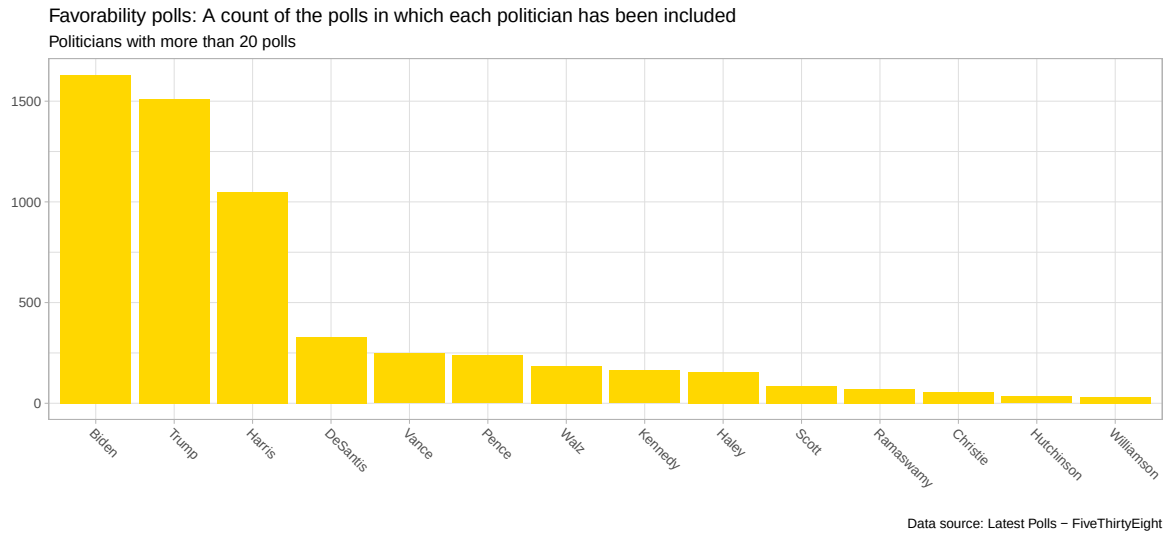
Another cool thing about Quarto is that you can manage your bibliography and references; for instance:

According to Wickham, Çetinkaya-Rundel, and Grolemund (2023), we conduct exploratory data analysis, among other things, to generate new questions about the data or refine the questions we already have.

To do so, you just need to copy the BibTeX reference from Google Scholar and paste it into the *references.bib* file located in the working directory. Check the citation syntax, here.

Quarto allows you to include code and print code output:

```
## Plot
df |>
  count(politician) |>
  arrange(desc(n)) |>
  filter(n > 20) |>
  mutate(
    politician = gsub(".* ", "", politician),
    politician = factor(politician, levels = politician)
  ) |>
  ggplot(aes(x = politician, y = n))  +
  geom_bar(stat = "identity", fill = "gold") +
  theme_light() +
  theme(axis.text.x = element_text(angle = -45, hjust = 0)) +
  labs(
    title = "Favorability polls: A count of the polls in which each
     ↪  politician has been included",
    subtitle = "Politicians with more than 20 polls",
    x = "",
    y = "",
    caption = "Data source: Latest Polls - FiveThirtyEight"
  )
```



You can set the figures' width and depth as general patterns in the YAML metadata. Likewise, you can set specific parameters by typing, for example, `#| fig-width: 11` at the first line of the chunk in question.

Moreover, you can insert LaTeX math:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

You can also insert inline LaTeX math: $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

> **!** Install extension
>
> If the iconic LaTeX string is not rendering in your computer, you must run the following command in the RStudio Terminal:
>
> *quarto add quarto-ext/fancy-text*

Remember, you can download this model Quarto document from GitHub. Please read my blog post about it.

As you will see, the model Quarto document has very extensive YAML metadata. Feel free to change parameters and customize the PDF file the way you find prettier.

By the way, here is some code to print neat regression tables using the `kableExtra` and `broom` packages:

```r
## Fit the model
model <- lm(mpg ~ cyl + hp, data = mtcars)

## Print regression table
kableExtra::kable(
  x = broom::tidy(model),
  caption = "Regression output",
  col.names = c("Term", "Estimate", "SE", "t-statistic",
   ↪  "p-value"),
  digits = c(0, 2, 2, 2, 4),
  format = "pipe"
)
```

Table 1: Regression output

Term	Estimate	SE	t-statistic	p-value
(Intercept)	36.91	2.19	16.85	0.0000
cyl	-2.26	0.58	-3.93	0.0005
hp	-0.02	0.02	-1.27	0.2125

Hit the *Render* button to turn the Quarto document into a PDF file. The new file will be created in the working directory.

References

Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2023. *R for Data Science*. " O'Reilly Media, Inc.".