Data Science: Productivity Tools
Murders Project

Benito J Palacios

2023-11-12

Murders

It is our first project named *Preparing for a Data Science Project*. We want to create two subdirectories to hold the raw data and intermediate data. We'll call these:

- data
- rda (R Data)

We need two scripts to load data:

- download-data.R: it's a script to download the file into the data directory for remote server.
- wrangle-data.R: it's a script to read this data and prepare a table that we can use for further analysis. This file is saved in rda directory. Although not the case here, this approach is often practical because generating the data object we use for final analysis and plots can be complex and time consuming. So we run this process one time and save the file.

Now we're ready to write the analysis file. We're going to call this one ***analysis.R***. The content should be something like this.

```
library( tidyverse )
load( "rda/murders.rda" )

murders %>% mutate( abb - reorder( abb, rate ) ) %>%
    ggplot( aes( abb, rate ) ) +
    geom\_bar( width=0.5, stat="identity", color="black" )+
    coord\_flip()
```

If you run this analysis, you'll see that it generates a file. Now, we want to save the generated file for use in a report or a presentation. We can do this using *ggsave*, it's part of the *ggplot package*. But where do we put the graph?. We will save plots to a directory called ***figs***. So we have to create this directory. Now we can add the next line to the code below:

*ggsave("figs/barplot.png")