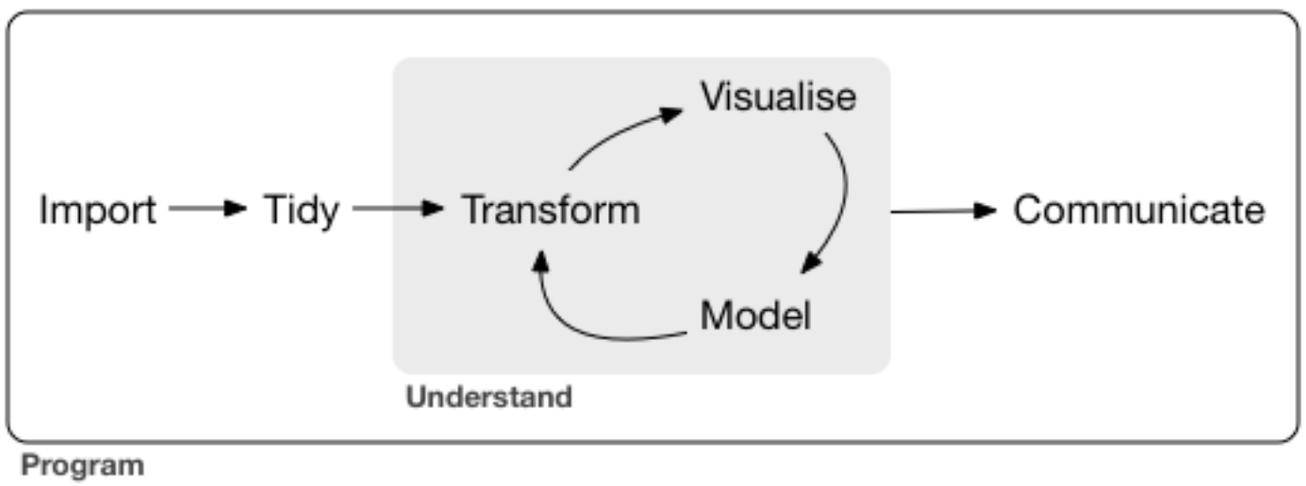
EDA Process and Discussions

EDA

Exploratory Data Analysis (EDA) is a term popularizes by the statistician John Tukey.

EDA generally represents the exploring part of the data analysis process where you are trying to “tease out” any sort of patterns in the data

The process can include more, but for us it basically means the plots and summaries we can create



EDA

An effective EDA should be purposeful and iterative

Purposeful: When you are trying to understand the data, you have to make intro plots to see what you are working with, but when you are presenting your work, the story of your EDA should tell a cohesive story. You should be building to something.

Iterative: This is connected to purposefulness. Your previous plots should inform the next ones. You may need to revisit previous plots based on new information learned from other parts of the EDA.

At this point we don't have great context for it, but the real purpose of EDA is to help inform you about modeling decisions (DS 2).

EDA

One of the best ways to understand a good EDA is to see it modeled:

Julia Silge: https://juliasilge.com/blog/drought-in-tx/

David Robinson: https://www.youtube.com/watch?v=Mkac8DHScps

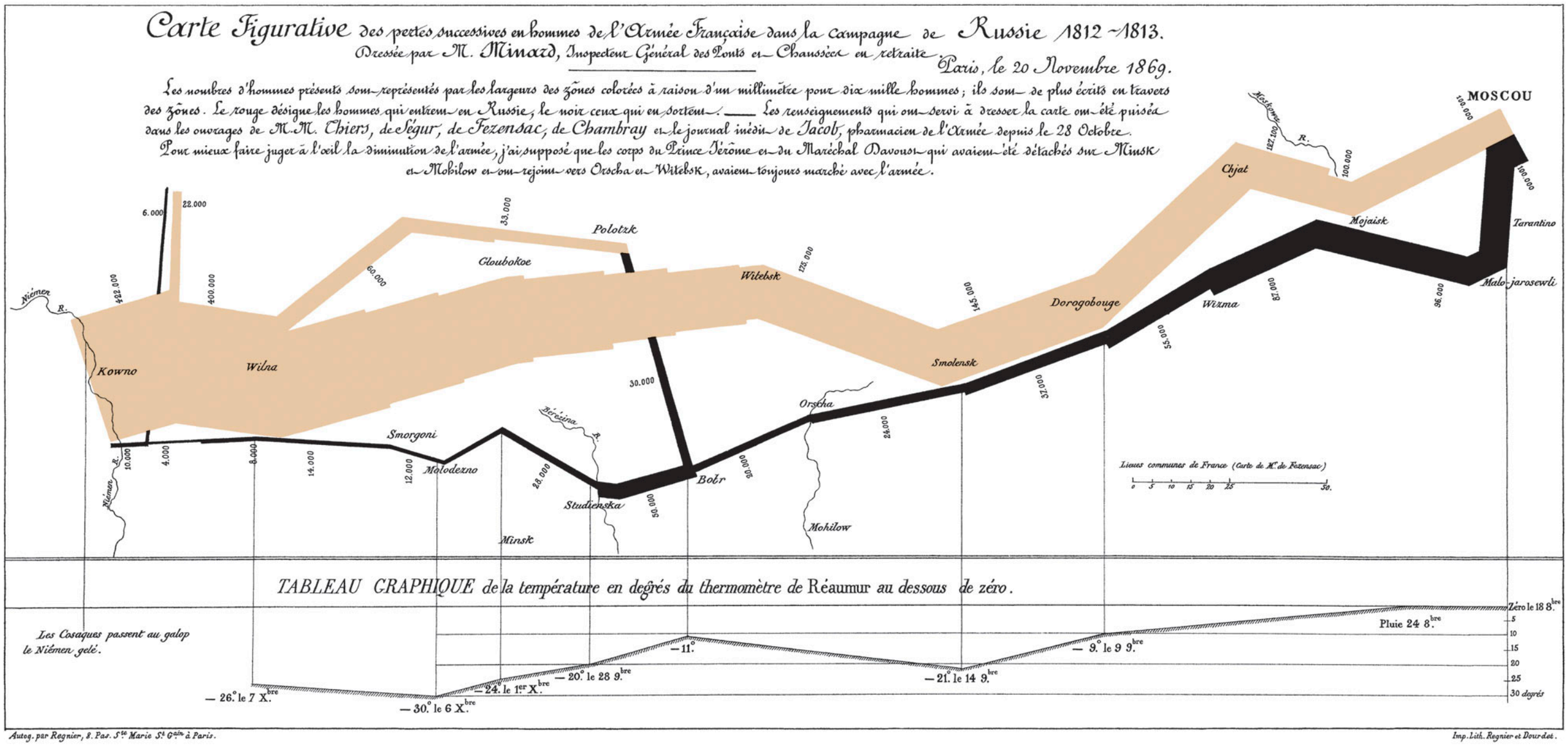
Plots

We’ve spent some time talking about the physical parts of graphics before, but now lets talk about the softer side of plots.

Quality over quantity: similar to what we’ve discussed previously, but a couple of really good plots are better than a bunch of individual (often one variable plots).

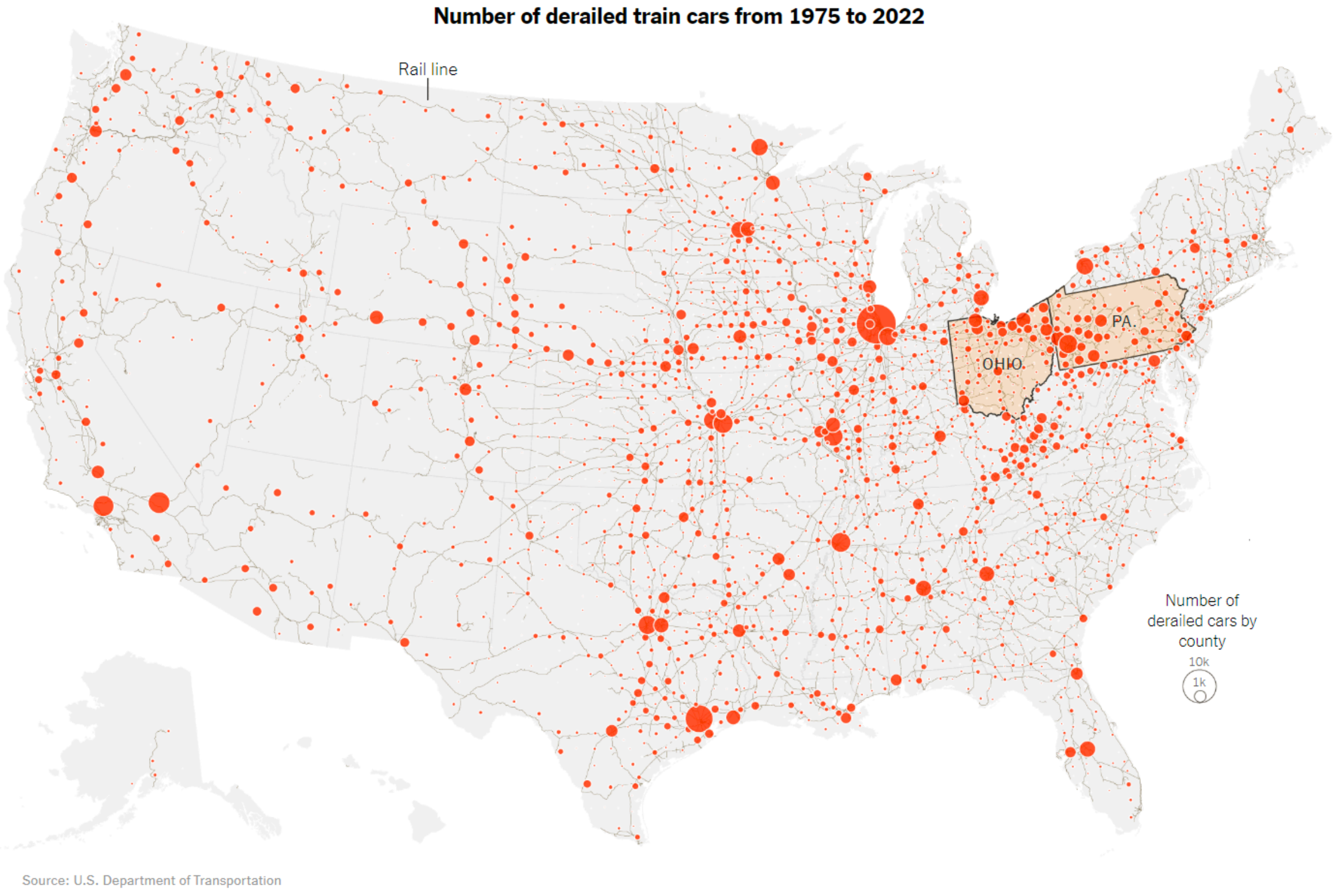
Context is key: This is important in both plot creation and discussion of the plot. Everything should be about the context. The aesthetics of the plot should enhance the context of the plot. The discussion of the plot should be almost entirely about the context of the plot.

Plots



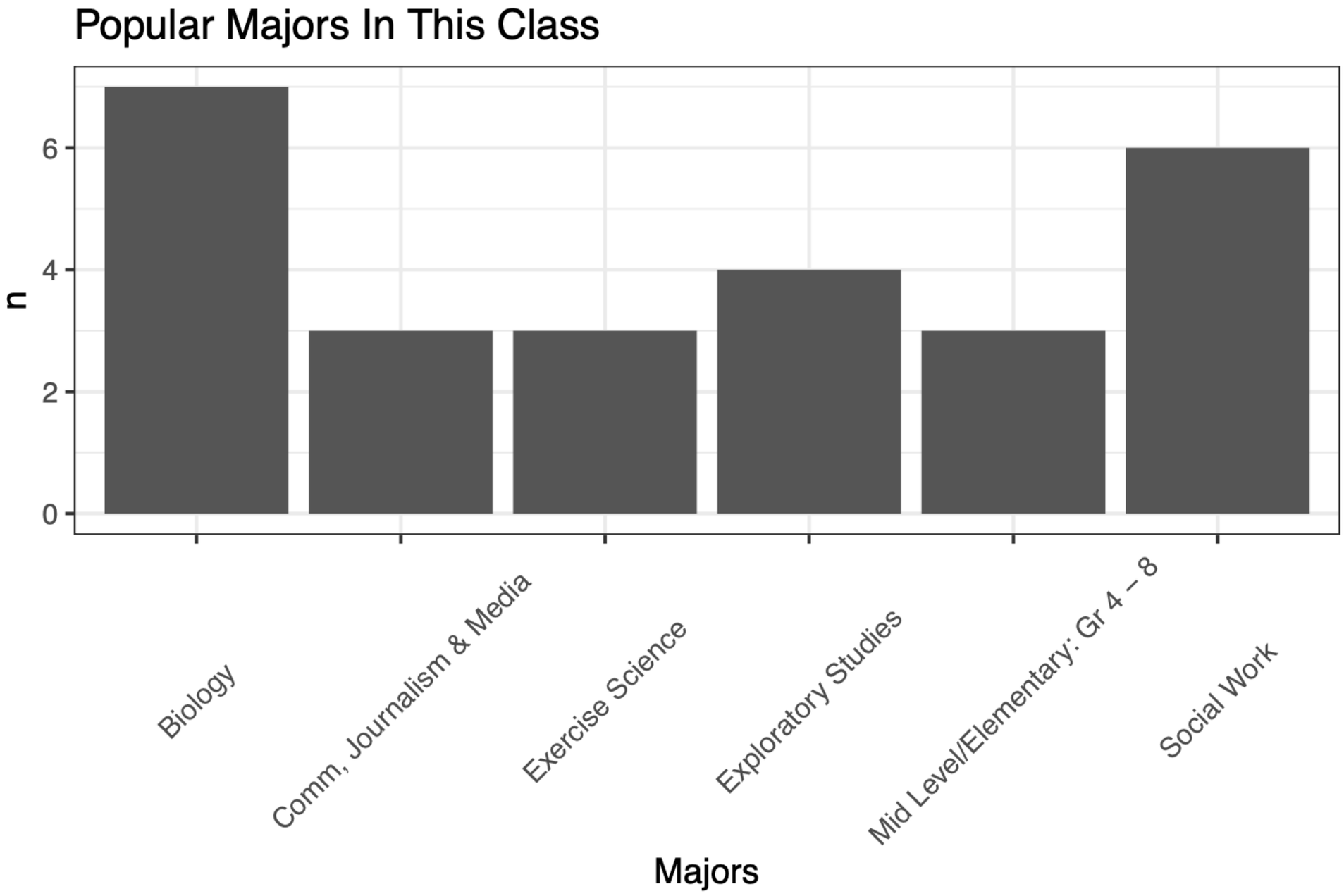
https://edspace.american.edu/visualwar/minard/

Plots



https://www.nytimes.com/2023/03/02/learning/whats-going-on-in-this-graph-march-8-2023.html

Plots


Popular Majors In This Class
n
0
2
4
6
Biology
Comm, Journalism & Media
Exercise Science
Exploratory Studies
Mid Level/Elementary: Gr 4 – 8
Social Work
Majors