The good, the bad, and the ugly.

This is a recitation for the visualization week.

With great power comes great responsibility.

There are many ways to visualize the same data.

You have just seen how to make quite attractive visualizations with ggplot2, which has good default settings, but judgment is still required from the user.

For example, do I decide to vary the size of a point or do I vary the color of a point?

It is worth noting at this point that Excel and other similar programs can also be used to make perfectly acceptable visualizations, or terrible ones.

The tool can help but it's ultimately up to the user it to make decisions.

So what is the difference between a good visualization and a bad visualization then?

I would argue that a good visualization clearly and accurately conveys the key messages in the data.

A bad visualization will obfuscate the data either through ignorance or malice.

So what does this mean?

Visualizations can be used by an analyst for their own consumption to gain insights into the data.

Visualizations can also be used to provide information to a decision maker and/or to convince someone of something.

Now, a bad visualization can hide patterns that could give insight or mislead decision makers.

This is where the malice part comes in.

So today, we will look at a few examples of visualizations taken from a variety of sources.

We'll discuss what is good and what is bad about them.

Then we will switch into R to build better versions of them for ourselves.

But I want you to think for yourself in this presentation.

You might not agree with all the points I make or my opinions about these visualizations.

Visualization is inherently subjective and the right visualization will depend on the situation.

So use your own judgment and think about what I talked about before with a good visualization and a bad visualization.