Anthony Cunningham

DATA 824

**Assignment 8**

Exercise 1

The following chart is an example of a graphic with a large amount of chart junk:

Chart

Description automatically generated with medium confidence

(found [here](https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=00040Z))

This graphic uses an unnecessary amount of ink to fill in the bars, which hinders the viewer from detecting where (on the x-axis) one bar ends and another begins. All of that ink could have been saved and would have enhanced the plot if lines separating each individual bar would have been used instead.

On the flip side, here’s an example of a graphic with minimal chart junk:

Chart

Description automatically generated

(found [here](https://homepage.divms.uiowa.edu/~mbognar/applets/bin.html))

This graphic is efficiently conveying the probability density function with little ink: each bar is completely filled with a basic color, each is separated and aligned at its midpoint with its corresponding x-tick, making it easier to visually separate each bar’s probability.

The following chart is an example of a graphic with a high Lie Factor:

A screenshot of a video game

Description automatically generated

(found [here](https://www.storytellingwithdata.com/blog/2012/09/bar-charts-must-have-zero-baseline))

A classic example is depicting a bar chart with a non-zero baseline, as shown above. This results in a high Lie Factor because, visually, the taller bar is many times larger than the smaller bar (~4x in this case), while the actual effect size is ~13%.

Exercise 2

Here’s an example of a chart that holds to the expressiveness principle:

Chart, line chart

Description automatically generated

Since there’s a clear ordering (by month), it makes sense to use a line chart to track the y-axis measure from month to month.

Conversely, here’s the same information being conveyed in a chart that violates the expressiveness principle:

Chart, histogram

Description automatically generated

This bar chart is incorrectly used with the x-axis measure, which would be more appropriate as a nominal variable (“Month” should be considered an ordinal variable).