

Quiz #4: Categorical Data

Milica Cudina

2021-09-14

Problem 1. (10)

The data set at hand concerns “pre-existing medical conditions of 92 children involved in a study on the optimal duration of antibiotic use in treatment of tracheitis, which is an upper respiratory infection.”

(2 points) First, you should read the data from our csv file “antibiotics.csv” into a data.frame:

(2 points) Now, create and print out the table of counts of different conditions present in the above dataset.

Now, you feel that your table is not pretty enough. Try the following command (don’t forget to un-comment the line of code once you insert the name of your table):

```
#knitr::kable(name-of-your-table-goes-here)
```

(6 points) Now, you want to plot a bargraph of the data. You do **not** want all of your columns to be gray and you **do** want all the labels of the bars to be visible. To figure out what the available colors are, you can use the command `colors()` in your console. In class, we have used the `barplot` command. To be able to adequately represent the data at hand, you will need to make some modifications to what we did in class. Please, format your barplot so that it satisfies the following properties:

- your chart must have a descriptive *title*,
- the barplot is *horizontal*, i.e., bars stretch from left to right,
- your labels for both axes are readable without tilting your head, i.e., they are readable from left to right,
- your labels for the *conditions* are must be readable in their entirety, i.e., none of the text gets cut off due to inadequate margin sizes.

This [barplot customization tutorial](#) is sure to be useful.

Problem 2. (5 points)

Remember that The University of Texas has a subscription to *The Economist* and that you can enjoy their charts and articles whenever you want.

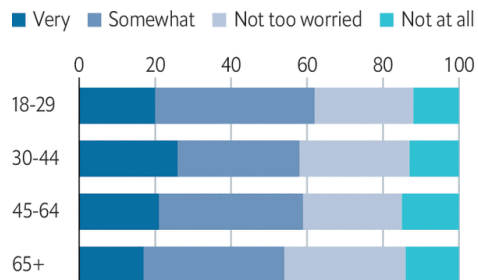
Consider the following charts which appeared in *The Economist* in late 2020:

```
knitr::include_graphics("bar_graph.png")
```

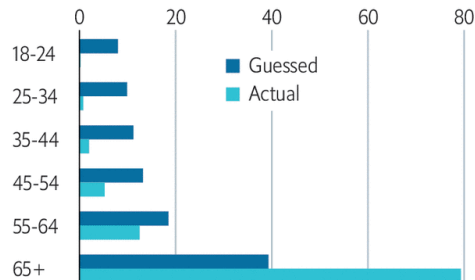
Youthful pessimism

United States, covid-19 by age group

Worry about contracting covid-19, %



Actual and guessed* shares of deaths, %



Sources: YouGov; Franklin Templeton-Gallup Economics of Recovery Study; Center for Disease Control; *The Economist*

*By a poll of all ages

The Economist

According to the above charts, which of the following statements is **not** correct? **Why? Justify** your answer by going through all of the statements below and providing supporting evidence of their veracity (based on the provided charts, of course).

- More than three in five 18- to 29-year-old Americans are “very” or “somewhat” worried about contracting COVID-19.
- Young people do not fall ill with the virus as often as older people.
- The findings **suggest** that many people underestimate the age of a typical COVID-19 victim.
- Young people are **underestimating** the number of COVID-19 victims in their age group.
- People aged 65 and older account for more than a half of the COVID-19 victims.