

**Some helpful resources:**

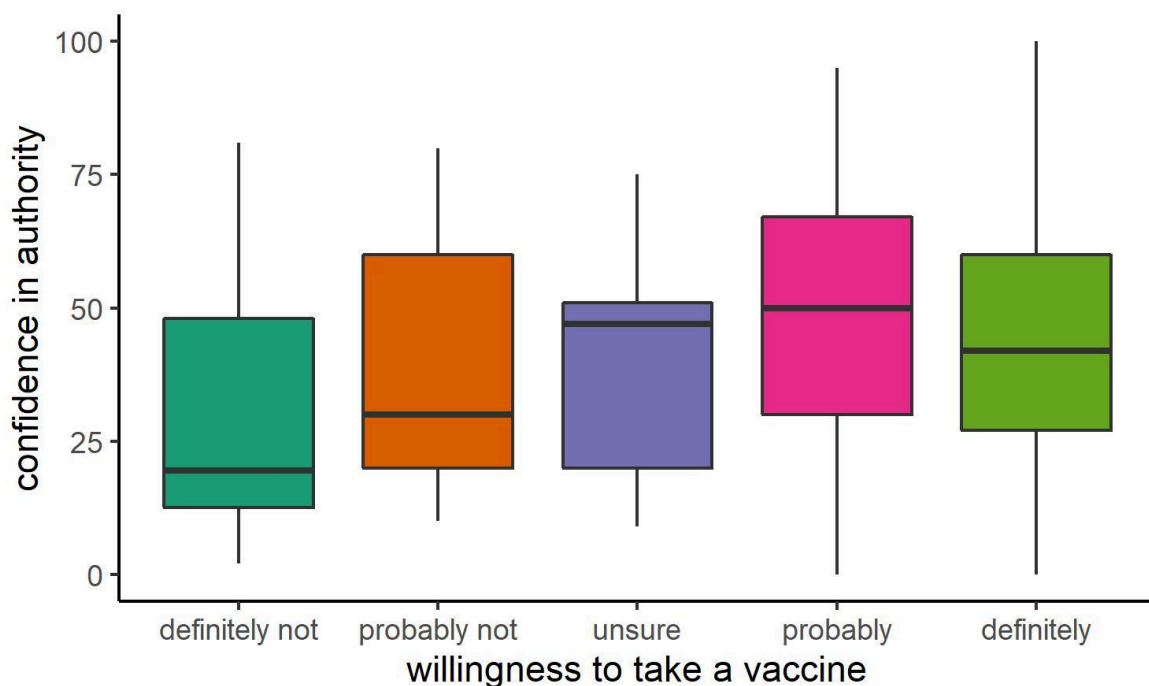
- Remember you can get help on any function by typing `?`  followed by the function into the console.
- ggplot cheat sheet:  
<https://www.maths.usyd.edu.au/u/UG/SM/STAT3022/r/current/Misc/data-visualization-2.1.pdf>
- R-graph gallery: Help choosing and creating types of plots (note: only use the ggplot section on each page) <https://www.r-graph-gallery.com/index.html>

**Hint:** For some questions you may need to make a variable into a factor and order the levels so that the axis values make sense. Remember, when you make a factor, R will put the levels in alphabetical order by default, but sometimes this order doesn't make sense for your data! Look at `?factor` to see how to set the levels by hand.

**Hint:** You may need to use the skills from last week cleaning data to make these plots!

1. **Open an R script and load the `covid_attitudes.csv` data.** Make sure you have loaded the tidyverse package and set `options(stringsAsFactors = FALSE)`

2. **Recreate the following plot**



Some **hints** before you begin plotting:

- Make sure you remove missing data (e.g., observations with NAs) from your data frame (think about the function we used to do this during session 3)
- Use the function `table(covid_attitudes$Q35.take.vaccine.)` to see how the responses are currently spelled. Do you notice any errors or any inconsistencies with the plot you are trying to recreate? (You should notice 2 things!) Use `mutate()` and `case_when()` to correct the spelling by creating a new variable called `Q35_take_vaccine_corrected` that is spelled correctly.
- Check the levels of your new correctly spelled variable `Q35_take_vaccine_corrected`. They might not be ordered in the same order as the plot above, but that's an easy fix! Use `mutate()` to create a new variable called `Q35_factor` and use `factor()` to reorder the levels. The `factor()` function only REORDERS the levels of your variable, it doesn't allow for renaming! That's why we had to do the previous step 😊

Some notes on this plot:

- This plot uses data from `Q35.take_vaccine.` and `Q101.confidence_in_authority`
- It uses the Brewer color palette `Dark2`
- It uses `theme_classic()`
- Hint: You can remove the legend from a plot by adding `+ theme(legend.position = "NULL" or "none")`

**3. Save the plot using `ggsave()`.** Set the width to be 5 inches and the height to be 3 inches.

**4. Create a plot of your choosing to answer a question you have about the data!** You can come up with your own question(s) as a group, or feel free to use one (or more) of the following questions:

- *Does the belief that scientists understand covid vary by type of community ?*
- *Does attention to news relate to confidence in the US government? Is this true for all age groups?*
- *Does attention to news affect expected symptom severity? Does this relationship vary with trust in hospital news?*
- *How do age and community relate to willingness to take a vaccine?*
- *How does education level relate to belief that scientists understand covid?*

There are SO MANY different questions you could ask with these data! Have fun exploring the data and looking for patterns!

**5. On [this JamBoard](#), make a new slide for yourself and post your finished plot(s), your name, and the question(s) you explored!**