

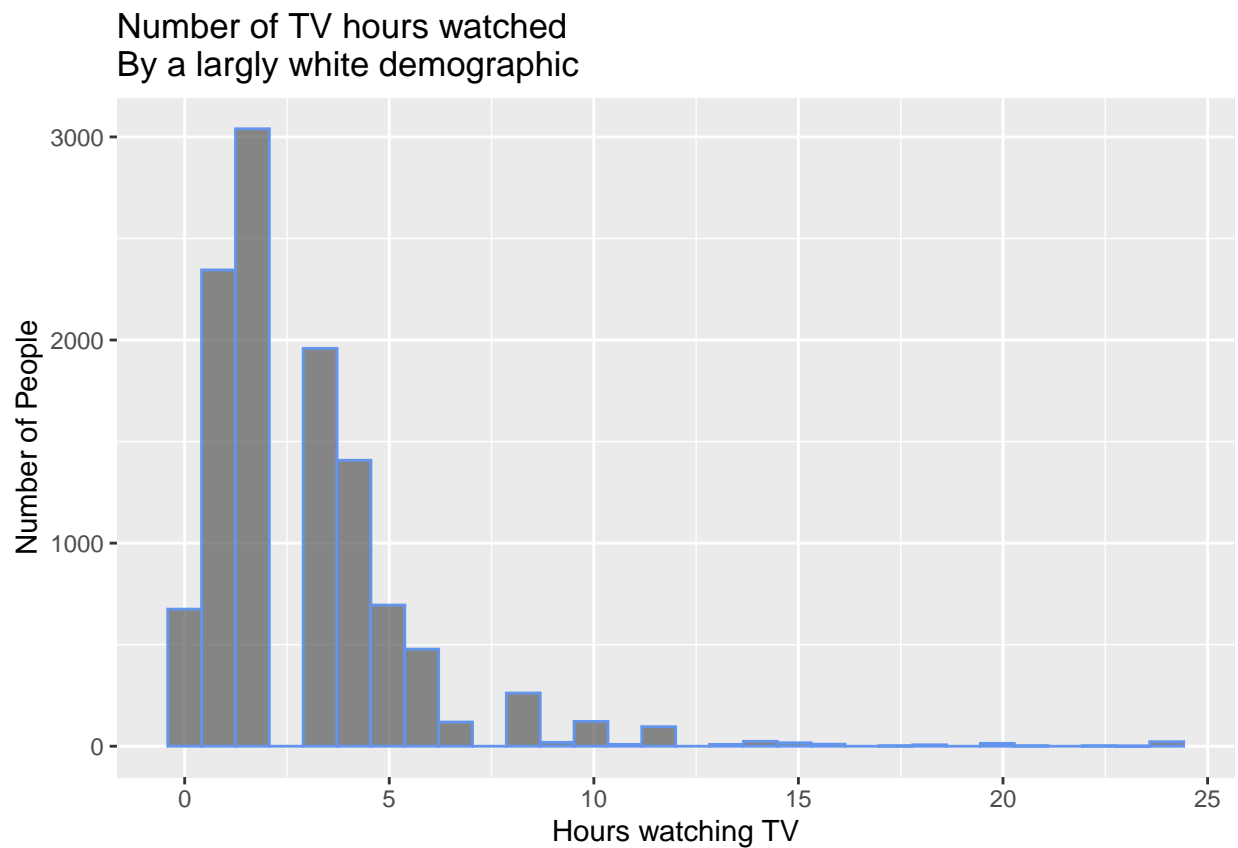
Lab 3

Tess

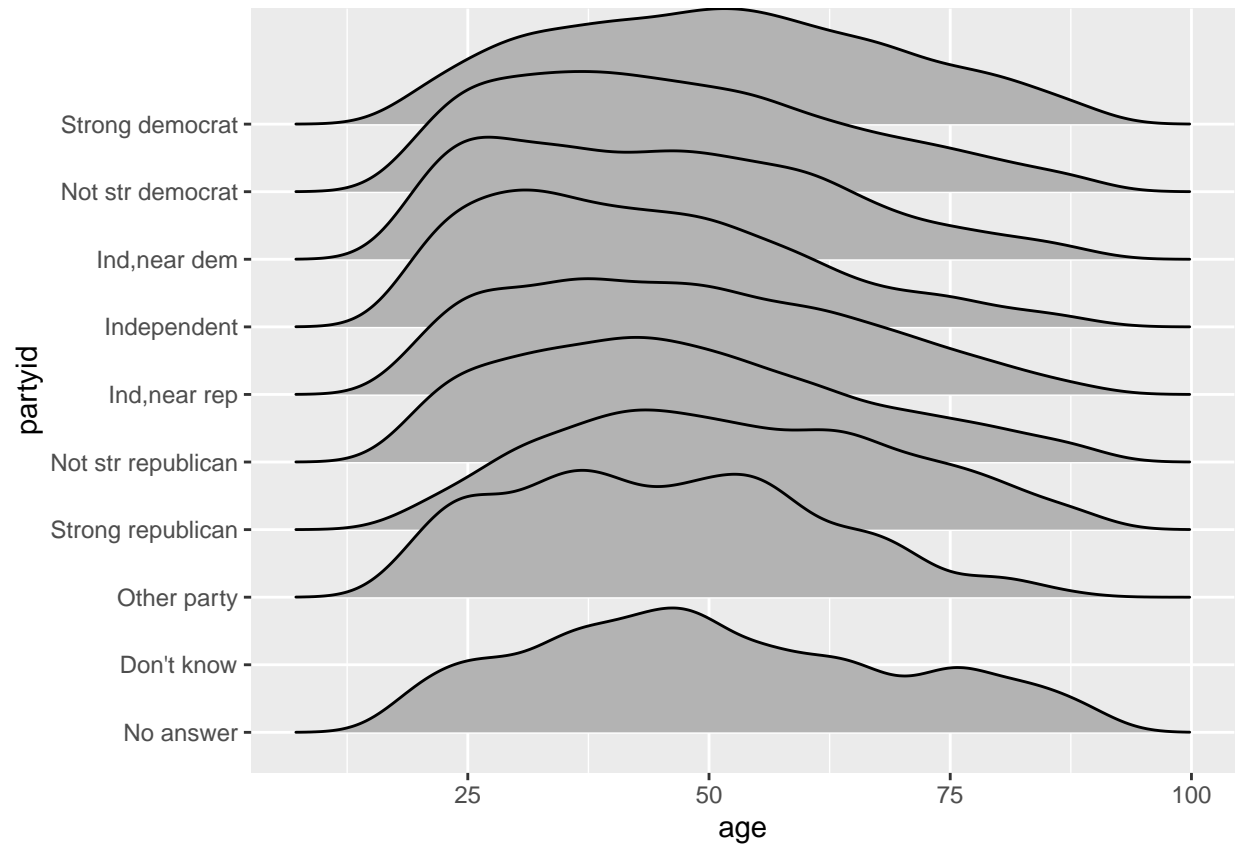
1. Run the following code to load the *tidyverse*.
2. Run the following lines of code to view and read about the data we'll use for this lab.

Comment out line 26 after you run it, but putting a `#` at the start.

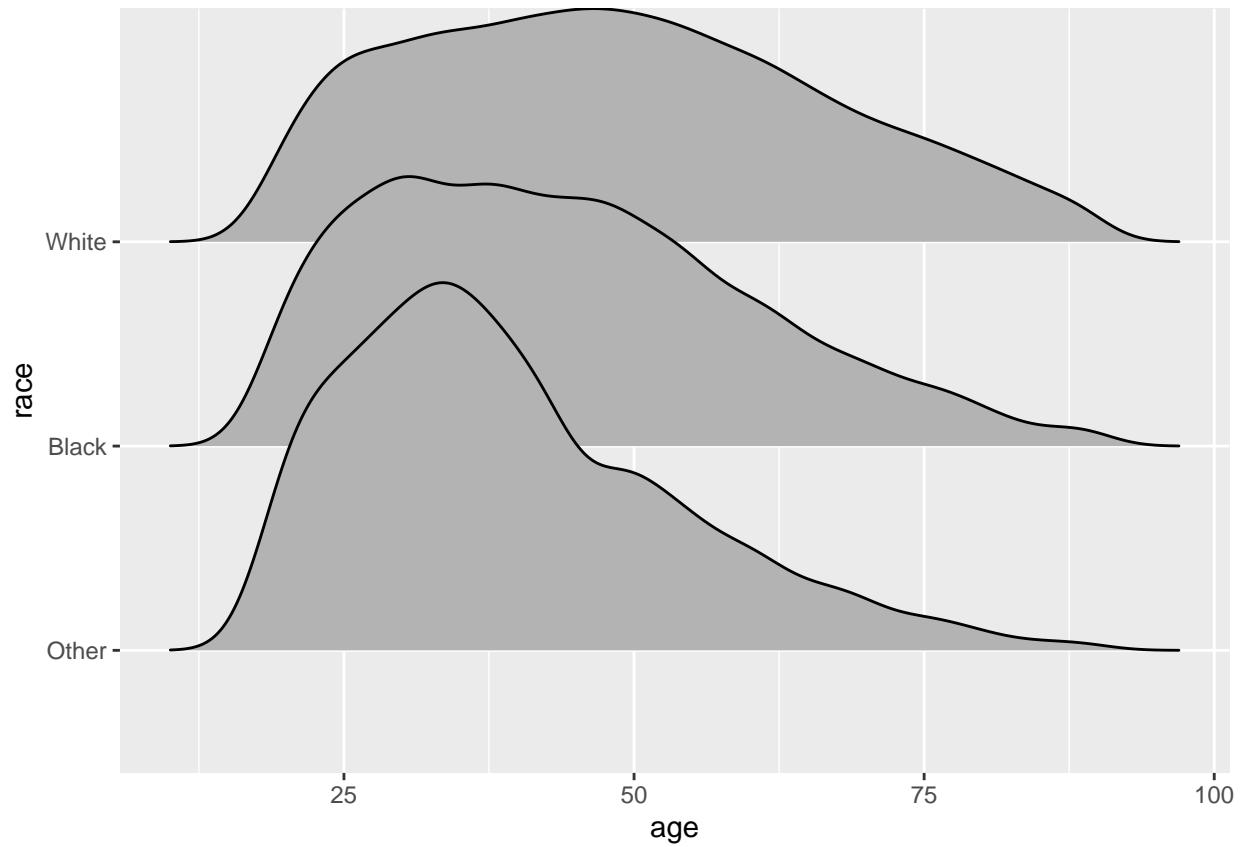
2. Use `ggplot` to create a histogram of `tvhours`. The plot should:
 - Include a small amount of transparency so gridlines are visible under the histogram.
 - Be a different color than the default
 - Include meaningful axis labels
 - Have a title and a subtitle



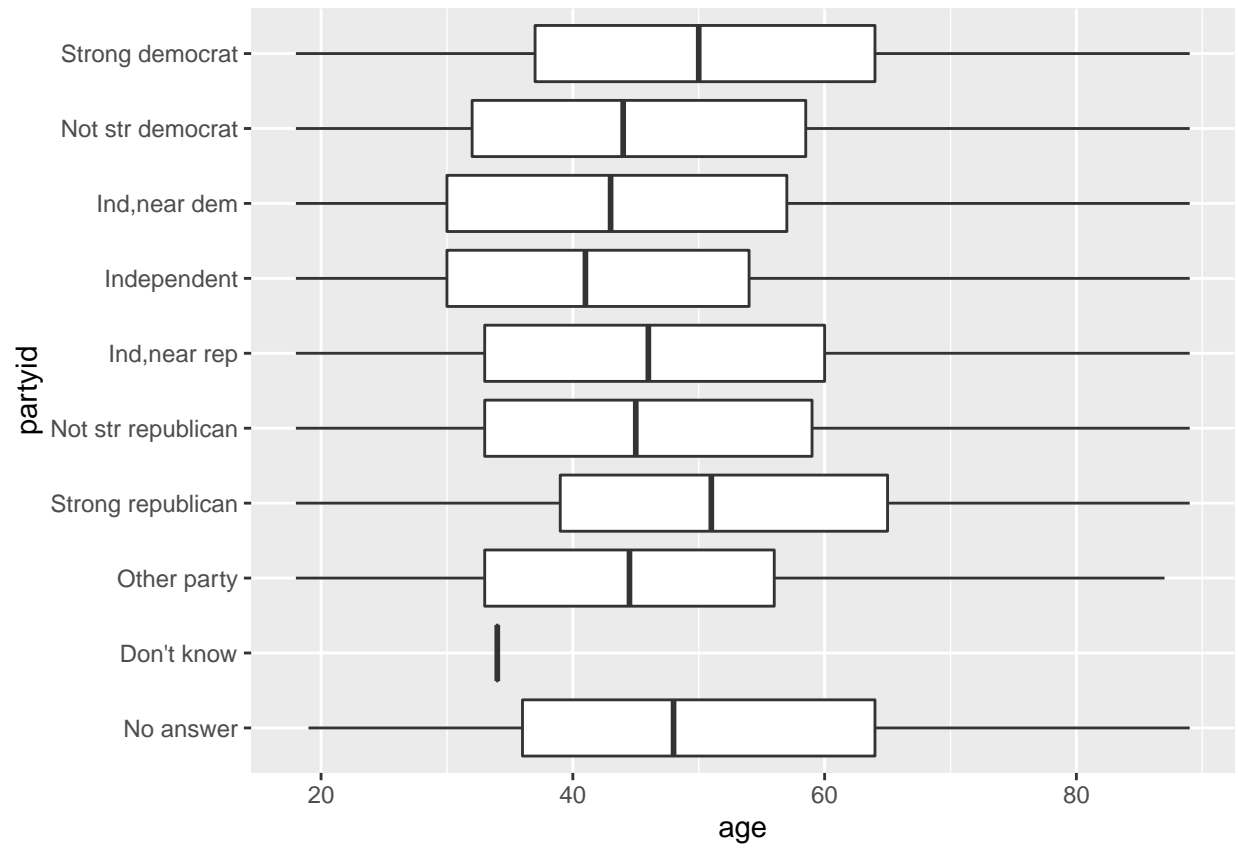
3. Create a ridgeline density plot of `age` by `partyid`. Do you notice any strong differences by party affiliation?

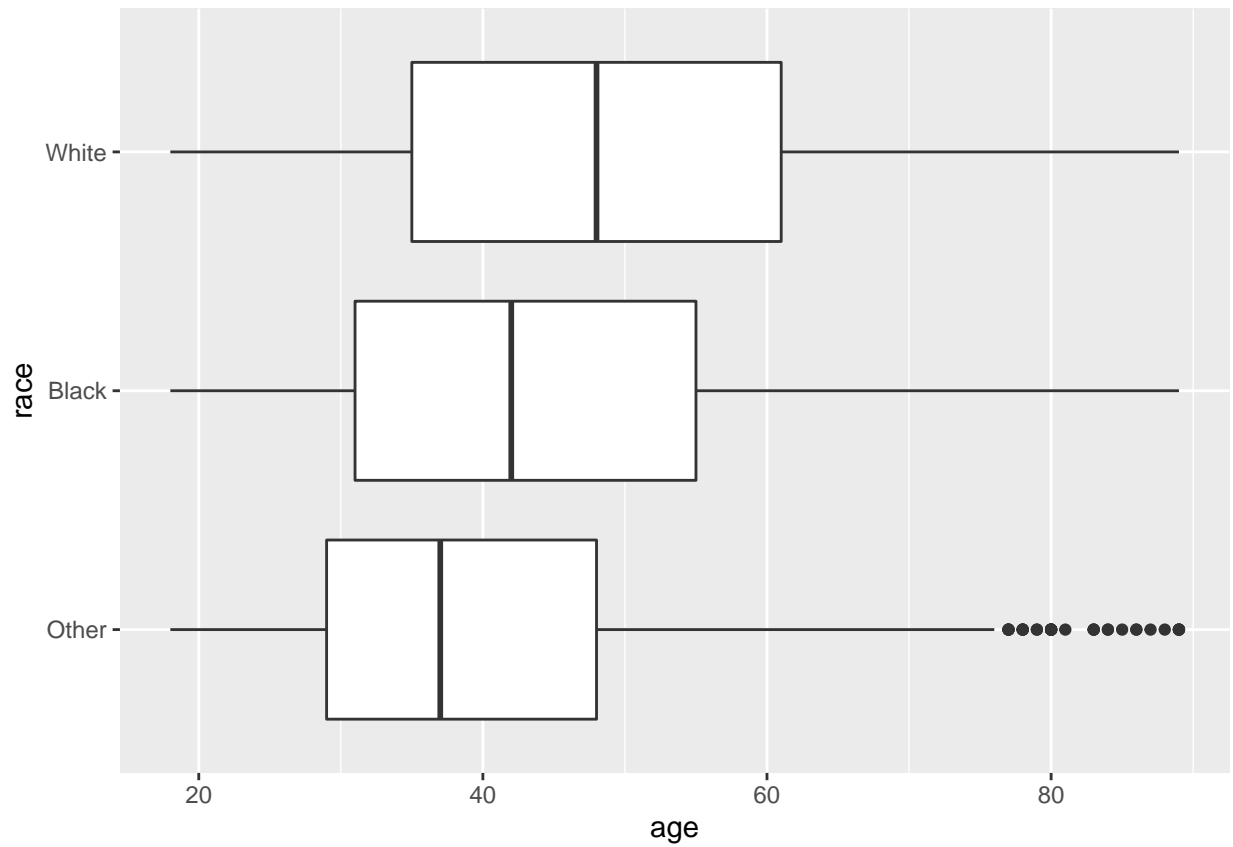


4. Modify the plot from Question 3 to produce separate ridgeline densities by **race**. Do you see any patterns?



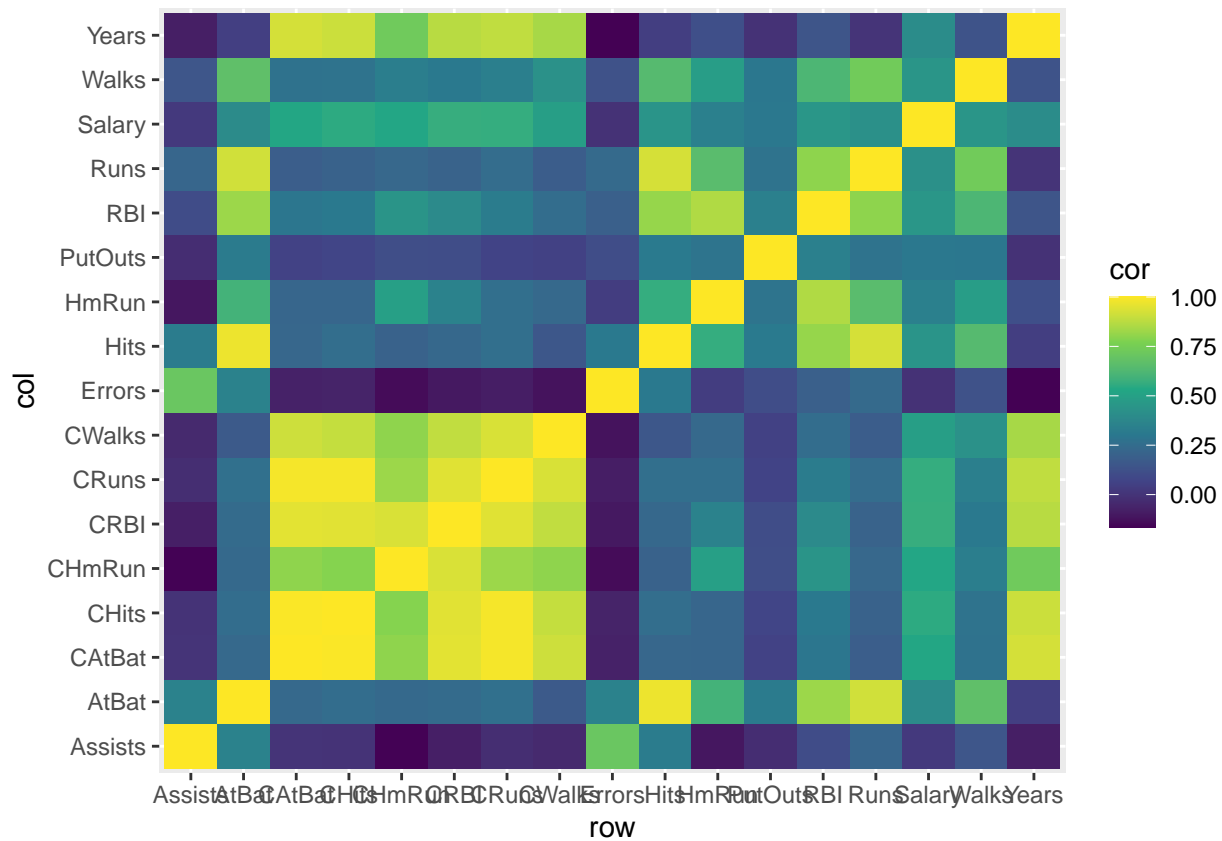
5. Replicate the plots from Questions 3 and 4, but this time displaying the relations with boxplots (i.e., one boxplot from Question 3 and one boxplot for Question 4 in the chunk below). Which do you think displays the relation more clearly? Why?



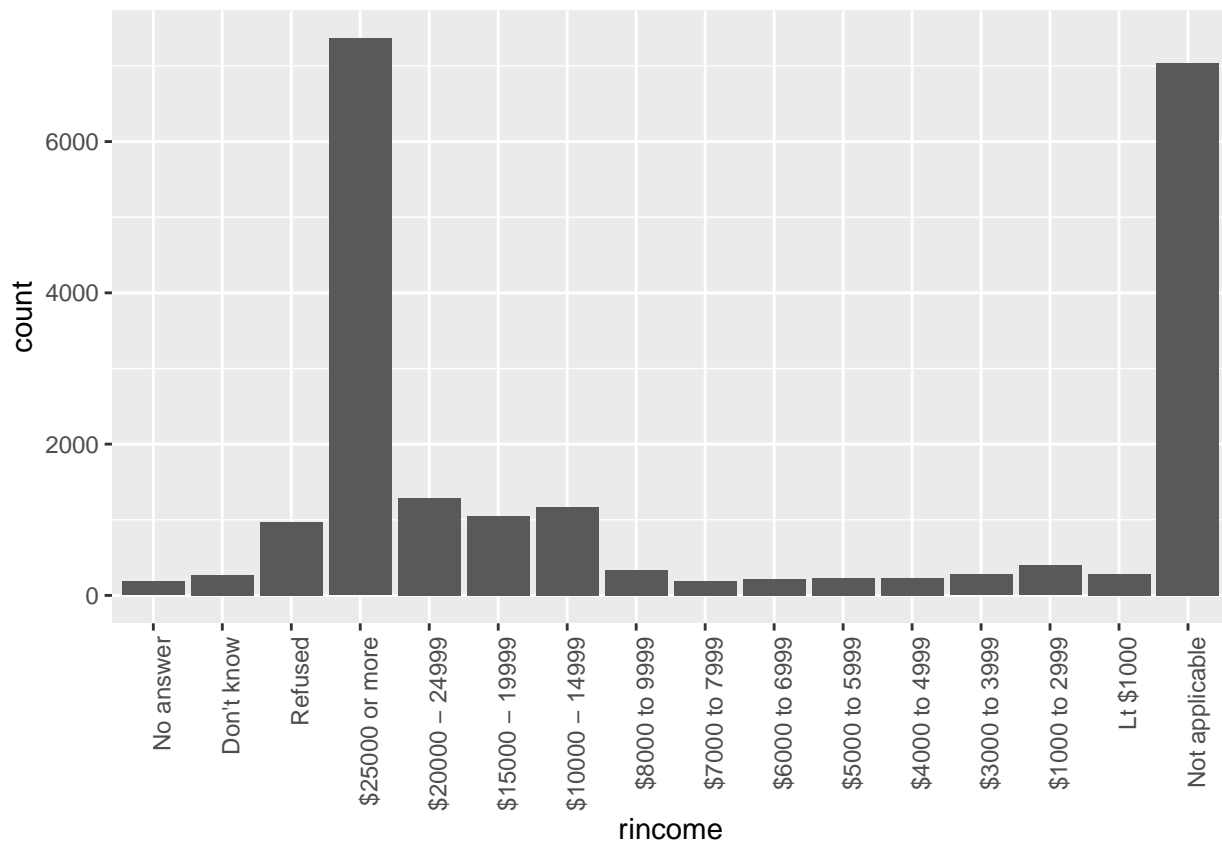


6. Run the following lines of code to get a correlation table as a data frame.

Use the correlation table to create a heatmap using `geom_tile`. Change the colors to one of the four viridis palettes. What does the heatmap help you see?



7. The code below will produce a barplot.



Use the code below to create a new data set - `gss_cat_new` - that redefines the `rincome` variable so that *Not applicable* is displayed first, followed by *No answer*, *Don't know*, *Refused*, and then each income category from lowest to highest. Once you've redefined the factor variable, re-produce the plot above by running the `ggplot` code at the bottom of the chunk. The plot should have *Not applicable* closest to the y-axis, then *No answer*, *Don't know*, etc.

