

Lab 11 - Data, Aesthetics, & Geometries

Jackson Nahpi

November 9, 2017

Complete the following exercises below. Knit together the PDF document and commit both the Lab 11 RMD file and the PDF document to Git. Push the changes to GitHub so both documents are visible in your public GitHub repository.

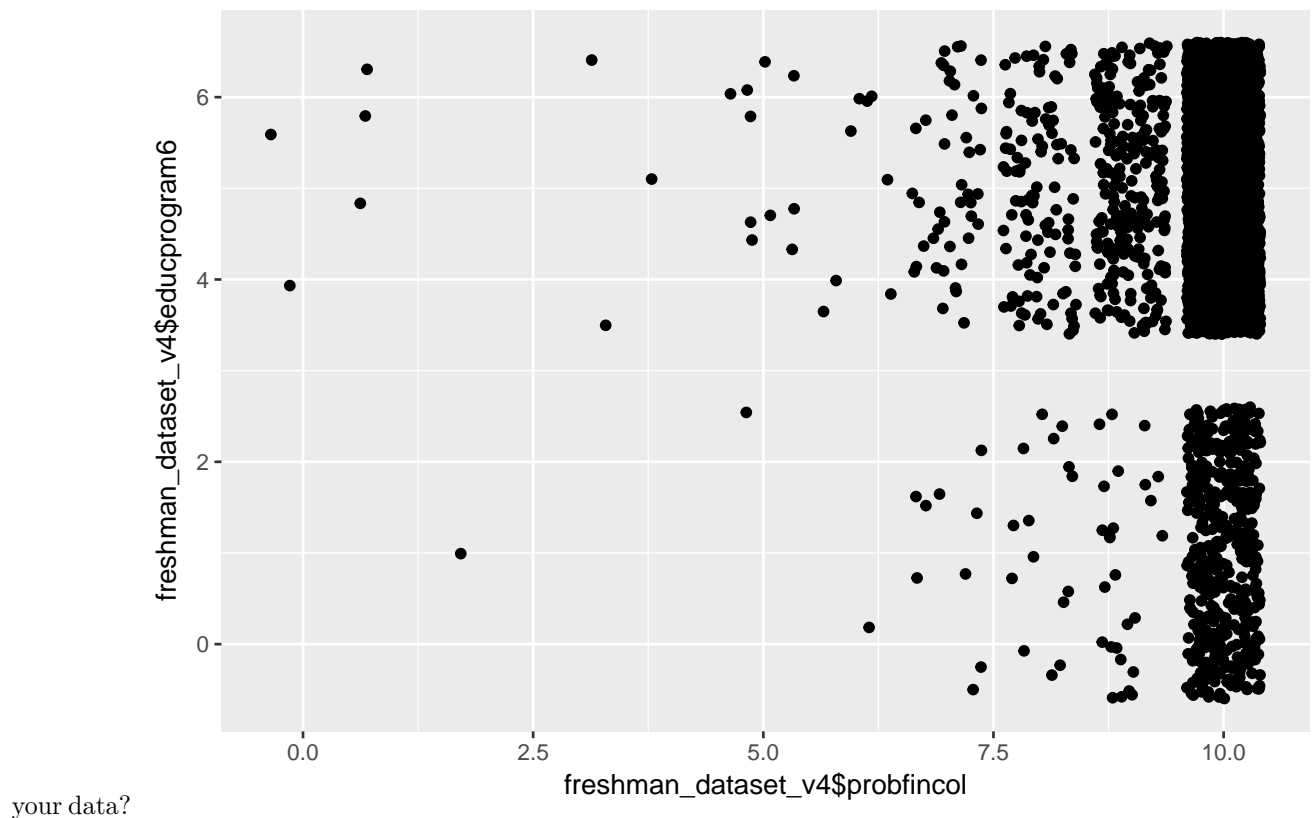
1. Which variables in your dataset are you interested in visualizing? Describe the level of measurement of these variables and what type of geography you think is appropriate to represent these variables. Give your reasoning for choosing the `geom_()` you selected.

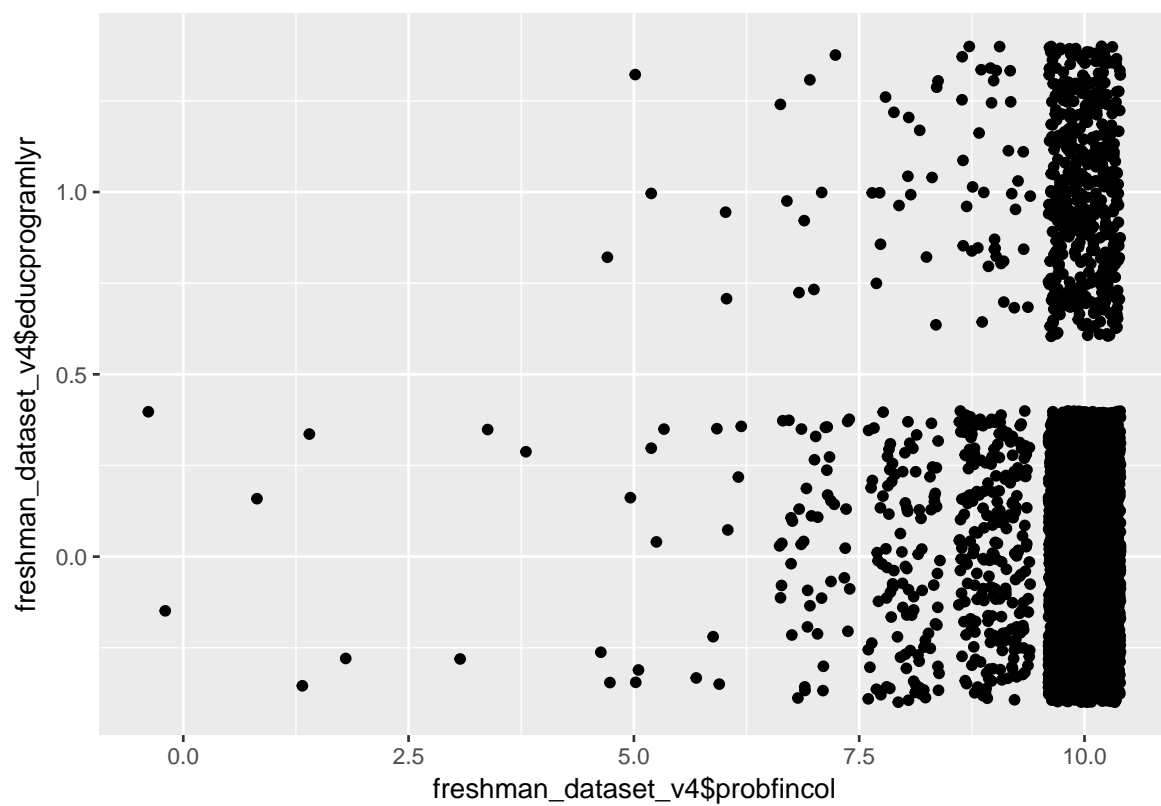
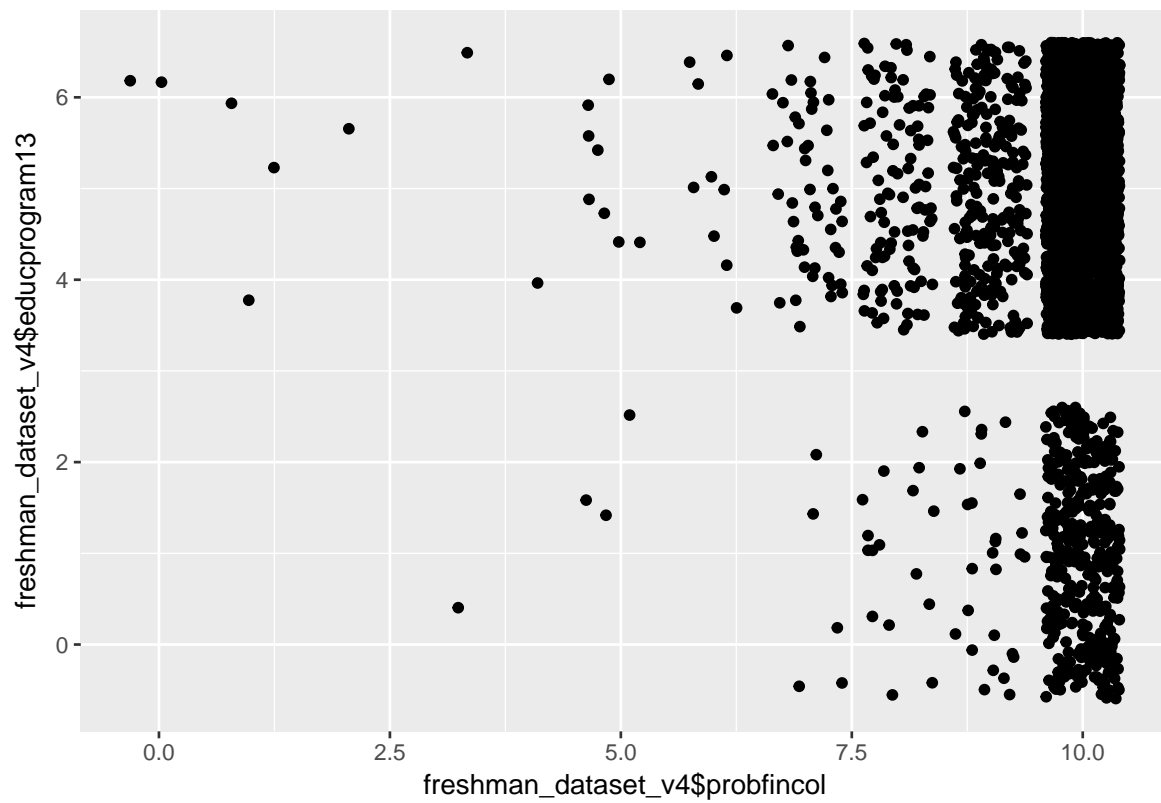
I would like to visualize the relation between educational programs and college aspirations.

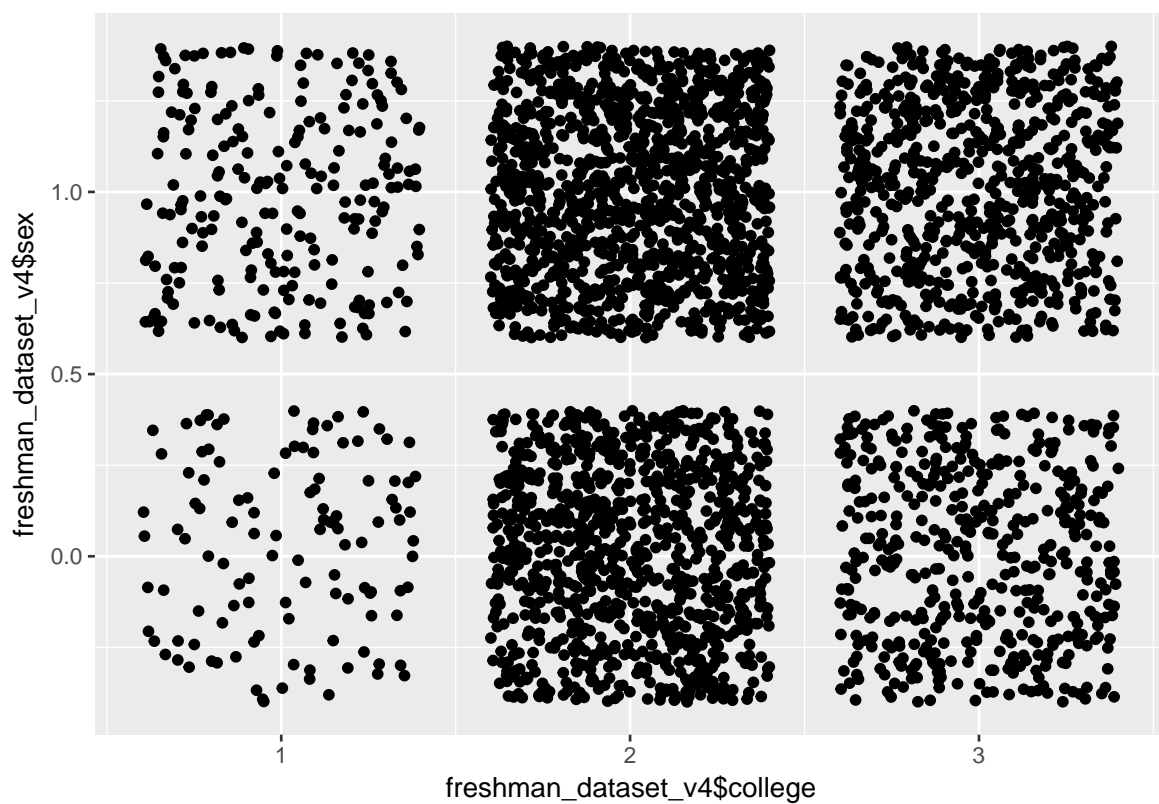
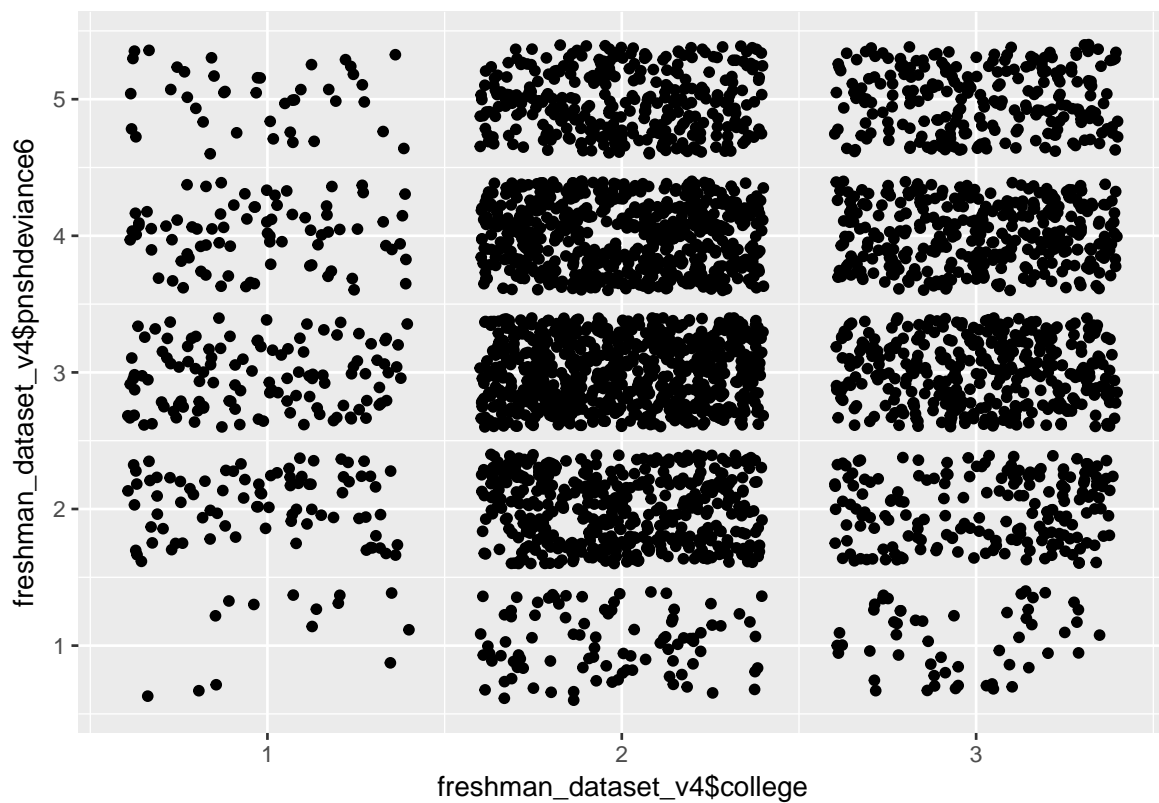
2. Is your data in the proper format to visualize the data in the way you want? Why or why not? *If you need/want to change the structure of your data, do it below.*

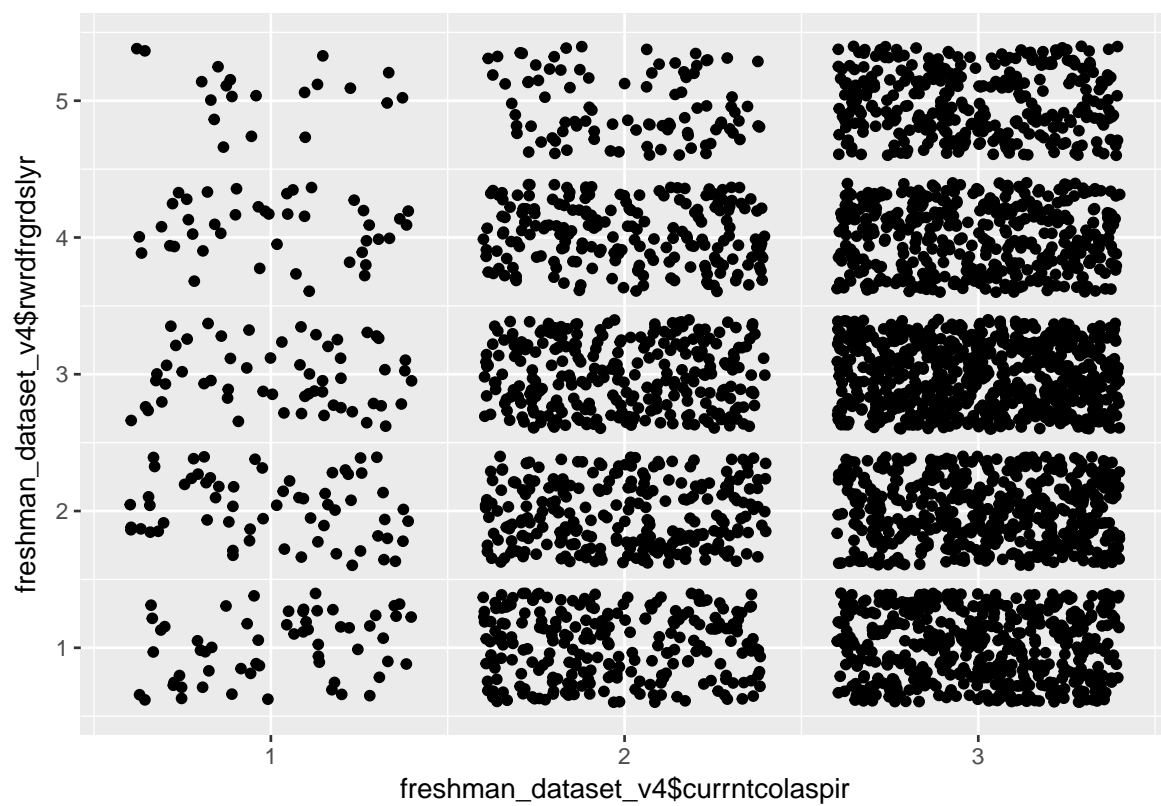
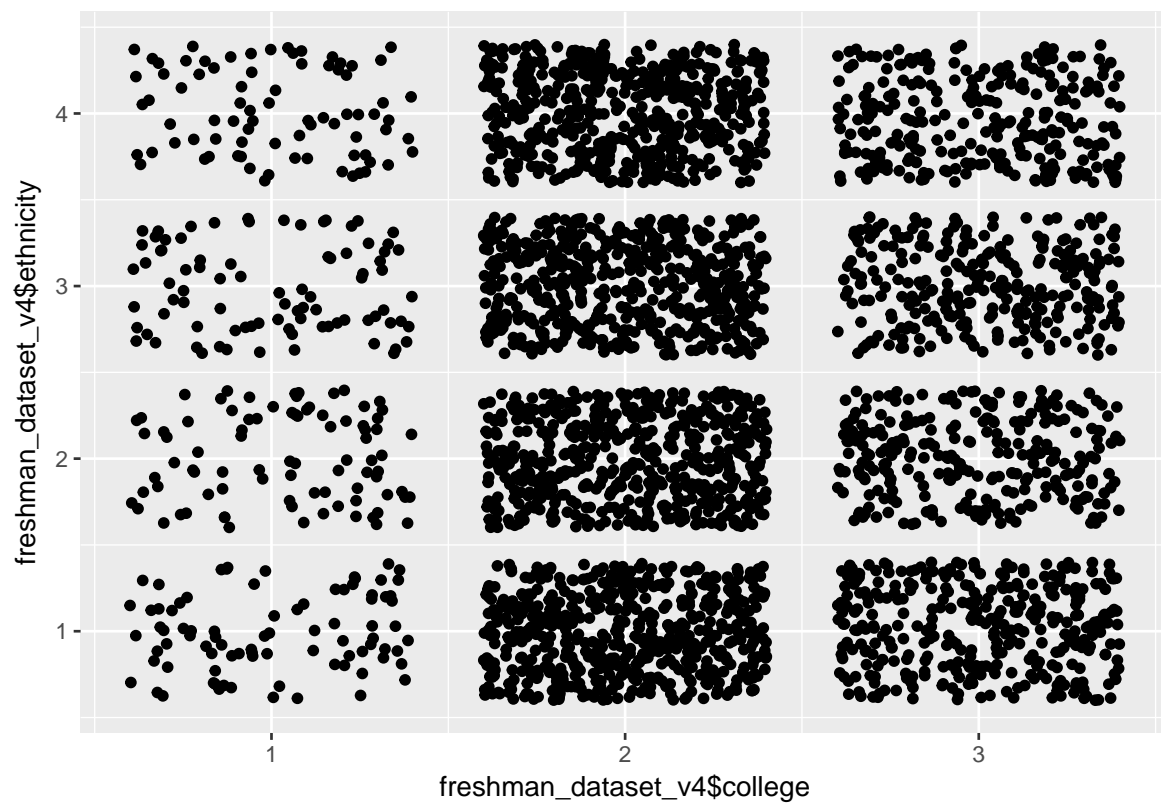
Yes my data is in proper way to visualize it since I want to look at my data in its categorical form.

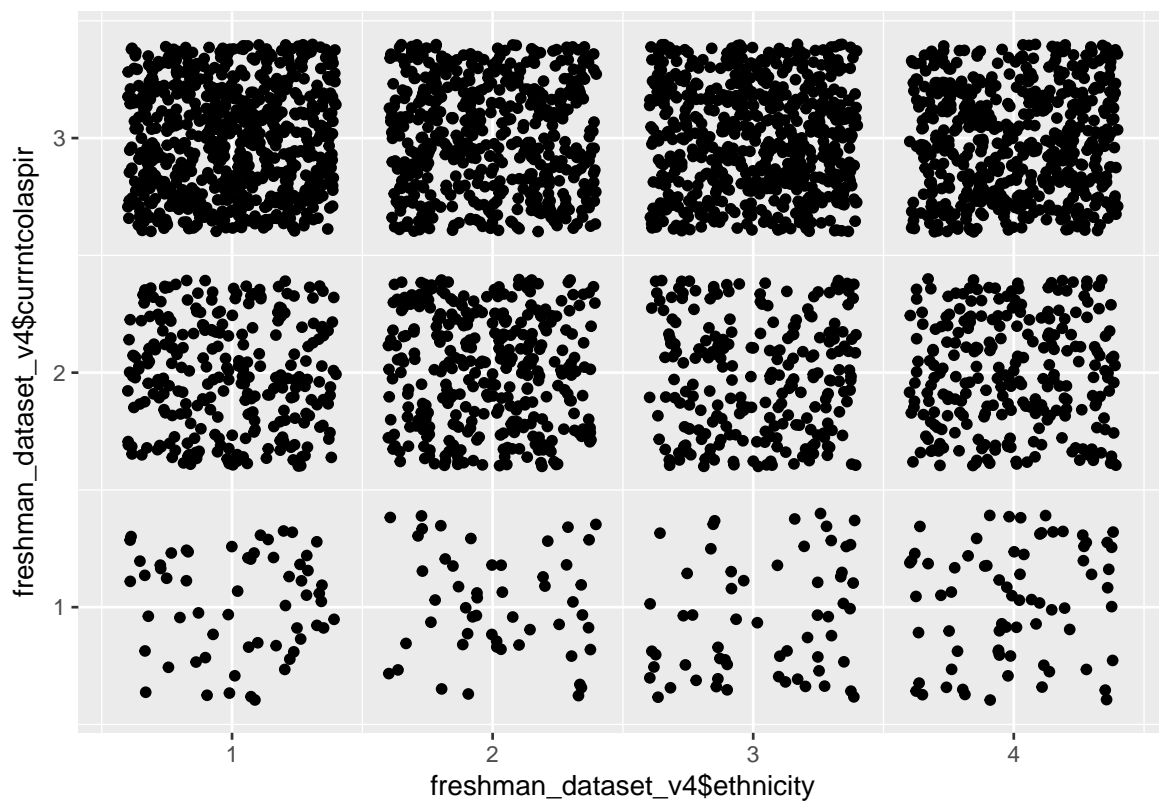
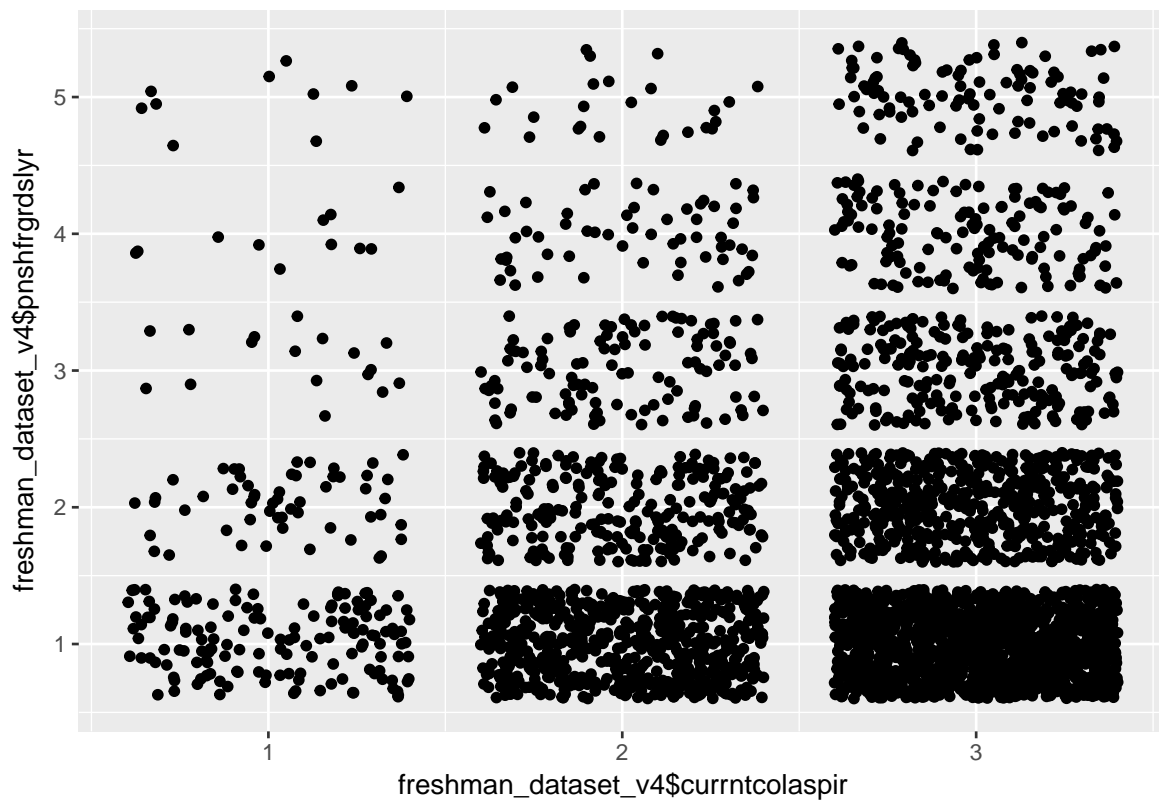
3. Create at least two different exploratory plots of the variables you chose using the skills we covered in class today. What types of mapping aesthetics did you choose and why? What do these plots tell you about







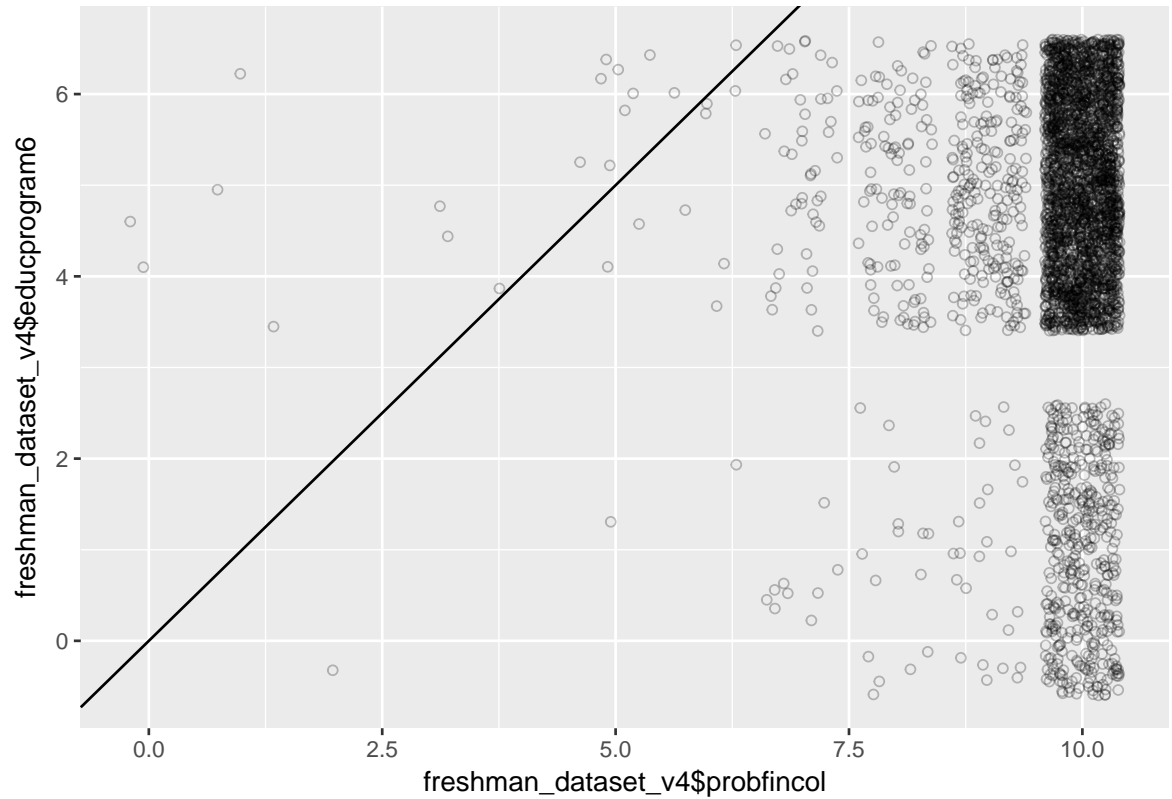


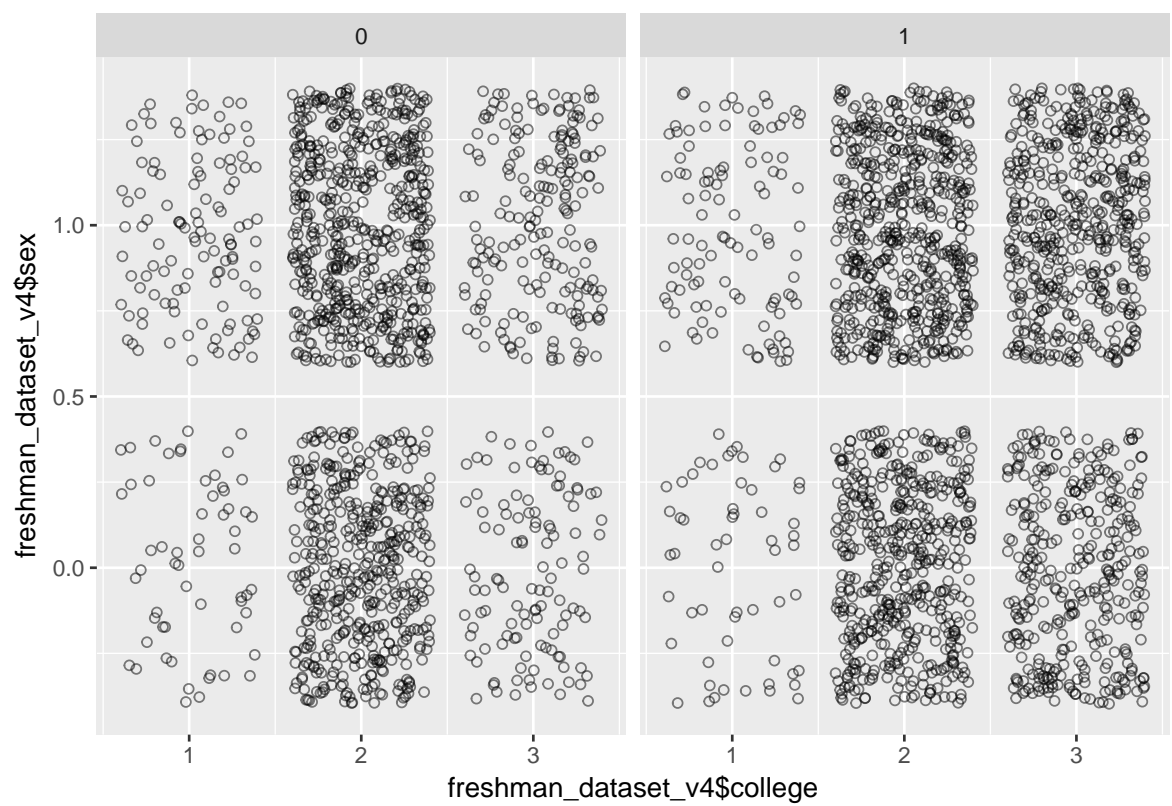
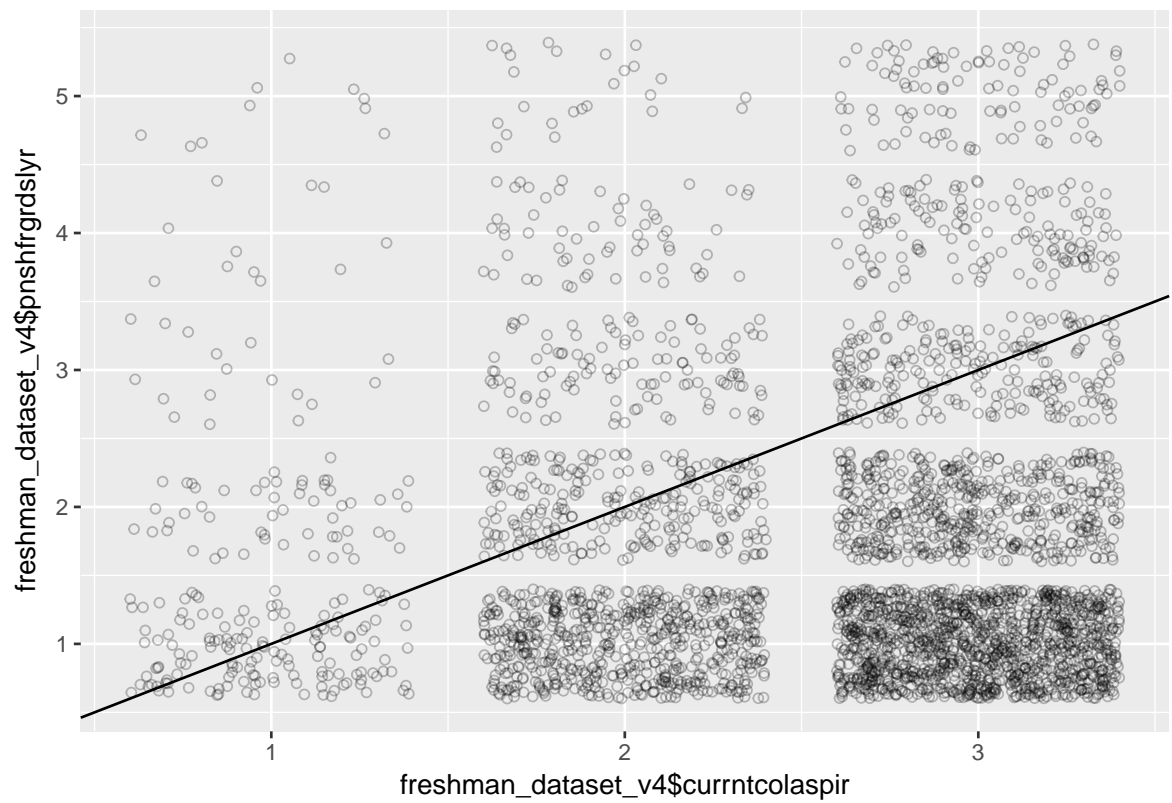


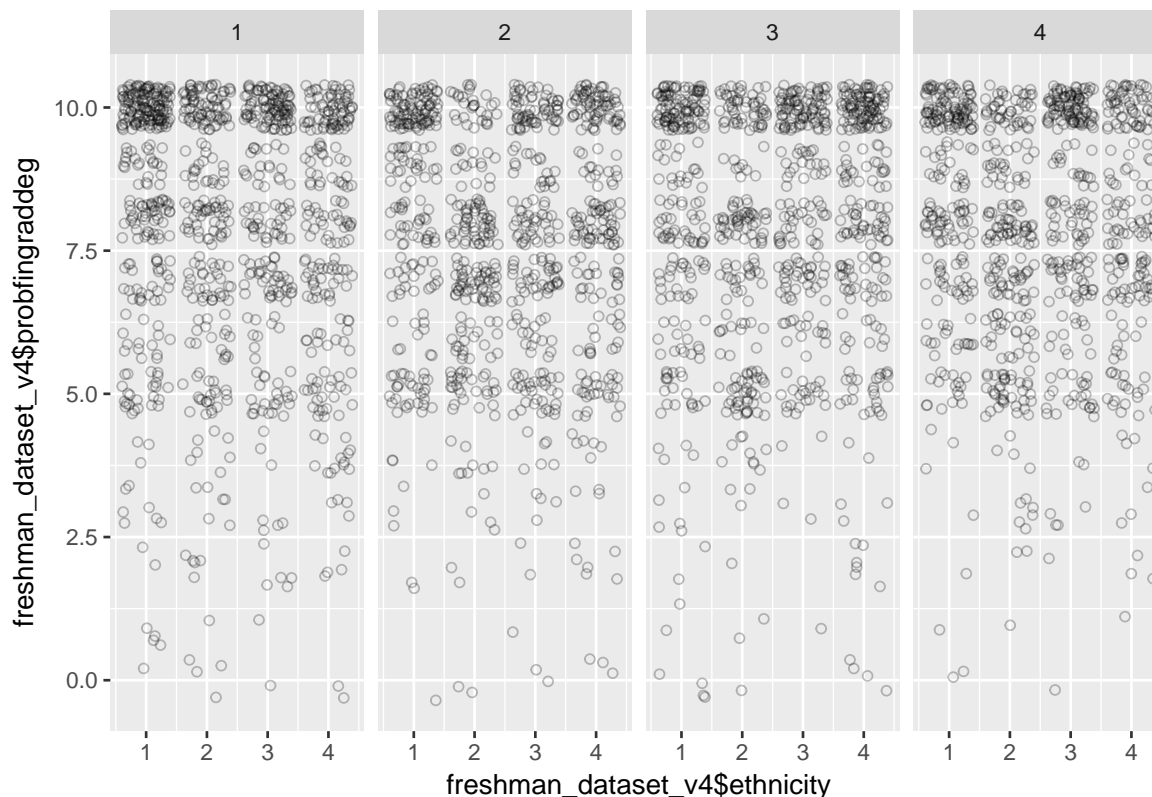
I did dotplots since my data is solely categorical. Most of the plots relating to college and certain cultural capital aspects given by their parents just show how the distribution is shifted depending on

the cultural capital generated by adults in the household.

4. Create at least three variations of the plots you've already made by modifying some of the arguments we covered in class (i.e. `position`, `scale`, `size`, `linetype` etc.). Do any of these modifications help you understand your data better? Why or why not? Do any of them create a misleading interpretation of the relationships between your variables? If yes, how so?







Yes some of them such as the table comparing sex and college choice show some slight relationship, however this may be misleading due to the low number of values the variables can take since they are categorical. However, there is a larger amount of females in my data so non representation is not an issue. The first two plots I altered are very good at showing trends and being useful in that regard so I may keep altering those ones later on. The ethnicity compared to probability of finishing grad school may be related but I need to alter the visual to better see the chart.

5. From the plots you've created thus far, do any of them seem appropriate for a general audience? Why or why not? If so, what do you think you'd still need to do to make them more suitable as explanatory visualizations?

I think the plots in question 4 that show comparisons between probfincol and educprogram6 would be good for my data since many of them are of the same format so I need to work on adding more to them and finding the best manner to present them and try to make the trends they are showing more clear. I also think that the plots showing facets of ethnicity and sex on different aspects are very interesting as well I may consider doing a density, heat plot, or maybe size adjustment of the points to show significance to better examine the differences in it.