Review For Exam Number 1

Generally, the analysis fully tackled the task of providing household analyses of sustainability factors by HOA and not HOA households. I'm impressed with how much was accomplished in the limited amount of time and the quality of the plots! A lot of cool functionality for ggplots was used that I didn't know were possible, so I certainly learned a lot from reviewing this.

While the general idea of the prompt was addressed (), there were areas where I felt that the general analyses could be improved to be used as a future template. A generally good idea is that it is helpful to comment code so that it's easier to follow what's going on. This also makes it easier for someone to go back and check their own work. Also, spacing out code and indenting makes it easier to organize ideas and thoughts into chunks.

Also, don't be afraid to be direct and explicit in analyses. If I did not take the exam, I wouldn't know what questions you were trying to address. Adding one or two sentences in the beginning will help the work seem more cohesive.

With a little more time, it might be nice to do an analysis at the neighborhood level for receycling habits, since the prompt mentions that current literature lacks multi-indicator analyses at the neighborhood level.

Specific Comments:

1. Page 0: There were 14 missing observations for tree data, however these were not cleaned out of the sample. The analyses, therefore, may not be representative.

2. Page 0: Before beginning your analyses, it's always helpful to summarize the data so that readers have a clear picture of the dataset that's being analyzed. This could be in the form of a graph, table, etc. Maybe just comparing the count of HOA to non HOA households could be valuable insight!

3. Page 7, line 4: Great job remembering to specify the "Recycle" variable as a factor!

4. Page 7, line 6,7: I've never seen names() but I'll definitely use that in my own code in the future for ggplots, very clever.

5. Page 7, line 11: Good job changing this to a factor. I'm curious, can you use as.factor() or does that raise an error?

6. Page 7, line 19,20: Awesome use of these functions. I'll definitely need to use these also for my own plots!

7. Page 8, Figure 1: Be sure to be consistent with your spelling of trashbin so that it is either one word or two everytime you type it.

8. Page 8, Figure 1: Good job being reserved in your analyses. Without relative frequencies, it's hard to draw too many conclusions.

9. Page 8, line 12: Is it necessary to multiple all values in a prop table by 100? I think you may be able to get your point across without doing this.

10. Page 8, Figure 2: Is it possible for you to make your ggplot code more concise/applicable to other figures you may want to make in the future?

11. Page 9, Figure 2: With a little more time, it'd be great for you to go more in depth on your analyses.

12. Page 9, line 5: I didn't know you could do this, cool!

13. Page 10, line 2: Do you need to multiply the frequency by 100?

14. Page 10, Figure 3: You mention the numerical values of "HOA" and "Garden", but these aren't seen anywhere in the figure. You did a great job making the plot, so no need to specify what these values are since they don't show up in the figure anyways. Also, it seems like more than a few homes that are part of an HOA have gardens. Be confident in your figures and look at the differences more closely!

15. Page 11, line 6: I'm confused about the comment. Are you ommiting the 14? If so, why weren't these observations omitted from other graphs?

16. Page 12, Figure 4: Good job mentioning the skewness of the data. Would a bar plot be the only good way of illustrating the data? A density curve could provide more insight.

17. Page 13: Awesome job remembering to reference ggplot2!