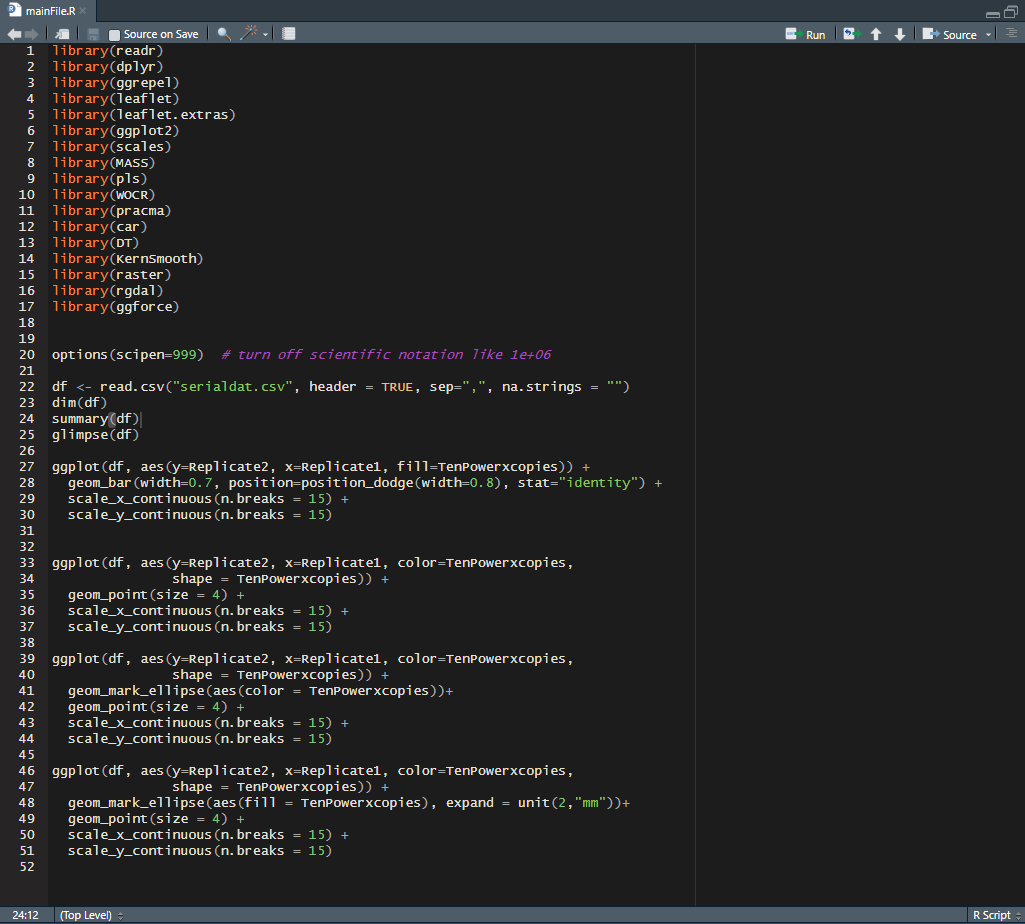
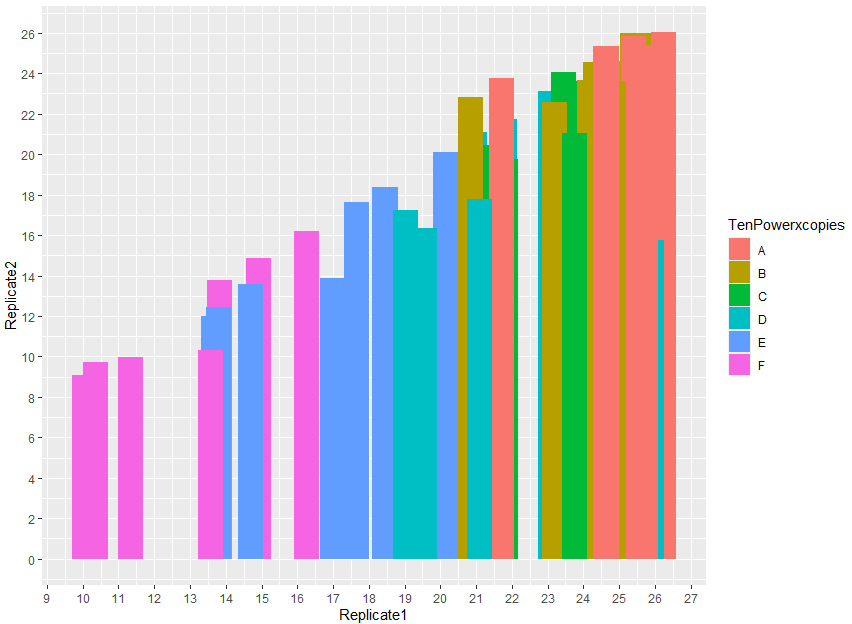
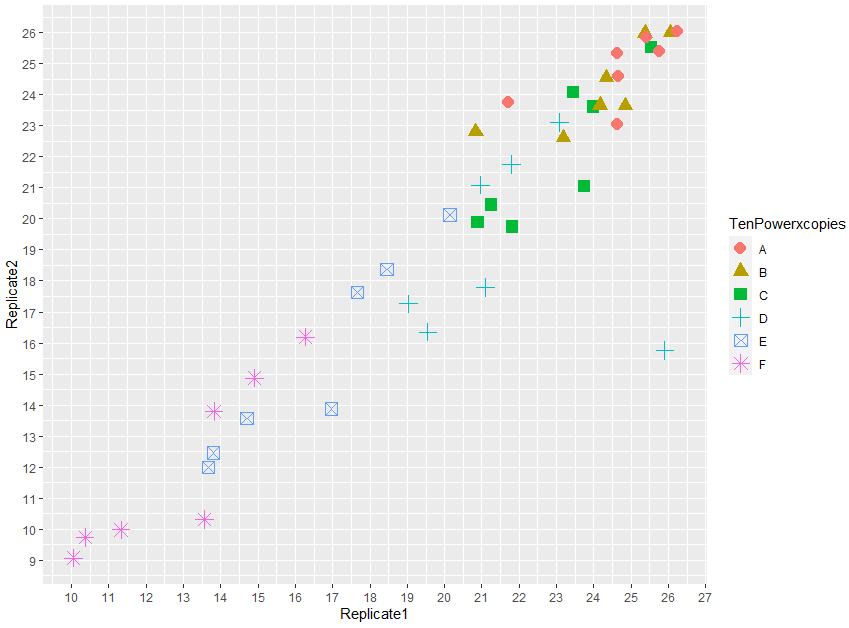
***HW 6: Use the provided data at the website (***[***https://ourworldindata.org/coronavirus/country/united-states?country=~USA***](https://ourworldindata.org/coronavirus/country/united-states?country=~USA)***)***, ***the attached transcription data,* *or All of Us data to complete a visualization of association for continuous outcomes. Choose at least two continuous scale variables and a categorical variable for grouping. Then, display the association in some manner that allows the viewer to see if association between the continuous variables differs across the categorical groups. I will not grade you on your selection of variables, but instead the quality of the dataviz.***

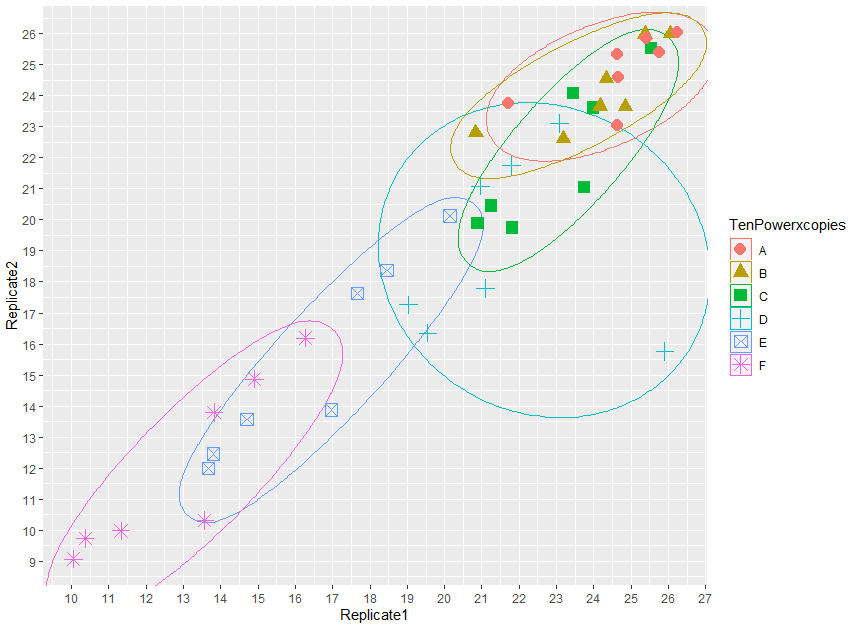
I used the dataset provided by Dr. Wagler on HW 5, serialdat.csv. I added a categorical variable with values A, B, C, D, E and F to represent variable “SUMOvar”- I did this in order to get 6 groups in my dataset. My first attempt was to create a bar chart and then a scatter plot. The bar chart gave me a better idea of the distribution of the values so I decided to combine the 2. My final result is a scatter plot with a clustering technique using ellipses.

Below is my source code and the various plots I did until I reached my final version.









**Below is my final version, my best visualization:**

