Data 115: Introduction to Data Analytics
Week #1 Assignment
Fall 2021
Grace Okamoto

1. The first two charts make it the easiest for me to understand the grading policy. Although I am more used to the style of the first pie chart, the second pie chart is also easy to understand. However, for me, the first pie chart wins, because it has the labels relatively close to the slices of pie, whereas the second pie chart requires a legend to be frequently referred to in order to understand what compromises what percentage of the grades. Being someone who gets confused and overwhelmed when there's a lack of linearity or direction for where to first direct your eyes, I appreciate these two more than the other charts provided.

2.

a. Average rainfall in Washington State, locations visited by your ex, internet browsing history, etc.

b. Math, computing, statistics

c. R, Microsoft Excel

d. Public health issues regarding healthcare affordability metrics, analysis to discover which groups of people have the highest priority for being serviced first

e. Professional, scientific, and technical services, finance, insurance, health care, social assistance

3. First, raw data is collected. This seems like a relatively straightforward process, however it could be difficult to collect data if responses are only limited to a specific group of people, or if the sample size is so large that the data collection will take a long time. Then, the data is cleaned, meaning that the measurements in it are standardized and inspected for errors. Next is exploratory data analysis, where they try to seek correlations between variables in their data. If this is possible to find, then a model or algorithm can be created using these correlations. To make this data understandable, the data is then transformed into something less numerical and more visual, to communicate more effectively in simpler terms the breakdown of the data. Lastly is to make this data usable to companies to increase efficiency of a process or give greater insight into their demographics.

4.

a. One of the interesting topics in the paper I read pertained to the relationship between Data Science and other fields of science. It was interesting to witness the complexity in the field and the creativity that is needed in order to solve the complex issues that arise.

b. I chose the Data Science Bachelor's Program from the University of California, Berkeley. The prerequisites for their program are fairly similar to those of WSU's lower-core requirements. The other courses involved in their curriculum seem to be much more lenient in terms of the variety of courses available to students that would apply to their degree. It also seems as if there are less courses required to take overall from UC Berkeley. Instead of the tracks offered at WSU, UC Berkeley has a wide variety of "Domain Emphases" offered, which only requires two courses to taken. The number of their Domain Emphases greatly outnumbers those of WSU's nine specialized tracts. Overall, there's much more options and leniency offered in UC Berkeley's program, which makes sense because the people at UC Berkeley probably have their lives together more.