Annie Qin
ECS 163
Feb 27th, 2020

Project 1 Report

For my visualizations, I visualized the Stack Overflow Developer Survey. I chose, a donut chart to visualize the main identity of the respondents, a spider chart to show the attributions of different languages used among different group of people, and a parallel coordinator to show four attributes regarding ages, and coding experience background. My visualization is pretty easy to use. You click on the donut chart to chose which group of people you want to visualize, and the other two charts will filter their data to the group of people you chose. There are a lot of interesting aspects by the visualized results. For instance, students are preferred to use python more than the other languages, but for professional developers, SQL seems like the most common language that used. In addition, there are one person whose age of first coding is 80 years, but he has been coded for 50 years, and has been code for professional purpose for 50 years. He is definably one of the outliers and very unexpected. In addition, a lot of student first visit stack over flow almost the same year of their first code experience. But as for professional developers, the results are very vary. The reason why I chose the donut chart to show the identities, is that not only I want to show different identities of the respondents, but also the attribution of the respondents among different identities. From that donut chart, we could know that people who are professional developers are the group that take the biggest part of this survey. For the spider chart, it perfectly showed the attribution of the popularity of different languages, which is exactly the result that I was seeking. For the parallel coordinator, one of the biggest advantage is that it could show different attributes, and it could show the relationships between these different attributes as well. For some challenges were that, it was a big set of data, it was hard to come out appropriate visualization design at first. I tried a lot of different method. For instance, I first tried a tree map to do one of the visualization, but that technique seems not fit to the logic of my data. Then, I tried to use a heat map or a matrix chart, but it is also not working for my data. I took a long time to figure out how to transform my data to a matrix chart, but it didn't work out. Then I started to look for if there is other techniques, and I found the spider chart, which is a perfect fit for my logic. In addition to that, the data was pretty big, and thus the cost was pretty high also, it took a pretty fair time to run at first. Thus, I tried some

techniques to at least improve it a little. It was a hard process also. Overall, I enjoyed the assignment, although there were some hardship. Below is the final product screen capture.