**Interactive Data Visualization Portfolio Reflection**
**Katie Donia**

**Final Project**

While exploring the US census data for the first time, I wanted to dig around for the most unfamiliar data I could find and stumbled upon a specific section for "island areas". When I clicked in, I saw surveys for Guam, the Northern Mariana Islands, American Samoa, and the US Virgin Islands. I was immediately intrigued, because I knew absolutely nothing about most of these places, not even a vague sense of location. At first I thought I could make a project that compares and contrasts the islands territories with each other in an attempt to find what makes each one unique. However, I found that when I was explaining my project during the ideation sessions, I kept gravitating back to distance and the crossing of physical space as the most interesting dimensions of the US island territories. Just how far away are they from the US, and if they are so far away, why does the US own them? And who ends up living on a Pacific Island and where do they come from? Coming out of the ideation sessions, I decided that I wanted my project to focus on this concept of distance by making use of d3's geographic projection capabilities. My first big design choice was to start from a world map centered on the Pacific Ocean and to have small buttons over the islands that, when clicked, zoom the view all the way in until the island takes up most of the screen.

My second big design choice was what census data to pick to best convey the idea of migration and where people have come from to settle on an island territory. I picked out three census questions that could offer insight into migration; race, place of birth, and language spoken. I thought place of birth could indicate migration that has happened in the last generation, whereas race could indicate much longer term patterns of migration. Language spoken is largely bound to both race and place of birth, and as such doesn't offer much more information than race and place of birth already do. Ultimately, I only implemented one stacked bar chart for race. If I were to continue working on this project, I might add the bars for place of birth and language spoken, although I do think race is the most interesting data and useful in learning about a wide range of Pacific Islander races I had never heard of, which is definitely in line with the planned function of the project.

I originally wanted the project to look at all four US island territories, but I realized, with the help of the critique session, that my project would be more interesting in a narrative approach rather than an exploratory one. If I am interested in a narrative about distance and space, then I should do my best to convey that narrative to my audience. At this point I also realized that if I was going to go the narrative route, then each of the islands should have a unique and specific narrative, and in that case I should maybe just pick one island that has the most interesting narrative possibility. Thus, I decided to focus exclusively on Guam and to develop Guam's narrative through specific text about Guam's history. That is also when I designed the guessing game that my project begins with. I think the guessing game is a super fun way to illuminate just how little most of us know about Guam and how far away Guam is from the US.

The stacked bar charts that display the census data were actually the quickest part of the project for me to design. Once I began to understand the visual freedom that d3 offered me, I decided I wanted to go all out on creating a final project that, most importantly and above all else, looked and felt really cool. I wanted my project to be dynamic and mobile in a way that would be impossible for Tableau. Most of my time in the planning stage was spent thinking not about how to display the data, but about how to make my project the most user friendly and intuitive as possible. I spent a lot of time deciding exactly what I wanted to happen on the screen at every possible click and hover. I spent the longest amount of time working on the interactivity and transitions within the stacked bar charts, particularly the way that some race categories break down into smaller subsections and how the bar chart visually transitions from being out of the total population to out of the smaller section that was just clicked on. I wanted the transition to make sense such that it is visually intuitive what is happening during the transition and what is being displayed after the transition. I wonder if I accomplished that or if it is confusing to users.

The biggest design feature that I had planned, but did not get to implement was to do more zooming and traveling around the world map to show the countries/areas that the different races in the stacked bar chart are indigenous to. I wanted there to be a button that pops up when you click into a specific race on the stacked bar chart that zooms the map out such that both Guam and the location that race is indigenous to appear on the screen together, again for the purpose of visualizing the distance that people traveled to settle in Guam. I think this would go a long way in furthering my distance-based narrative and would take the project to the next level by exploring a lot more of the Pacific Ocean than just Guam.

**Tutorial**
I chose to fix up my scatterplot tutorial, which I had already improved upon from tutorials 2 to 3. The first thing I did was change the title of the graph. Originally, I just gave the graph a basic descriptive title, but this time I wanted to create a more succinct narrative and needed a title that would quickly clue the audience in to what the graph is trying to illustrate. I like using questions as data visualization titles, because the audience knows right away that the graph is going to be trying to answer the question. So, I made the title into what is essentially my research question: "How Much Effect Does the Quarterback Have on a Football Game?" The next thing I did in service of this narrative was flip the x and y axes. According to how my question is worded, I am interested in using the quarterback's performance as a predictor for outcome of the game, so without a doubt qb passer rating should be the independent variable and team points scored should be the dependent variable, which is not how I had it arranged before. I also added to and rearranged the tooltips a little, again in service of best answering my research question. Lastly, I added a much needed legend to explain the color and size scale of the points in the scatterplot.