**NHMRC : National Australian Research Grants Interactive Data Visualization**

Randi Mendel

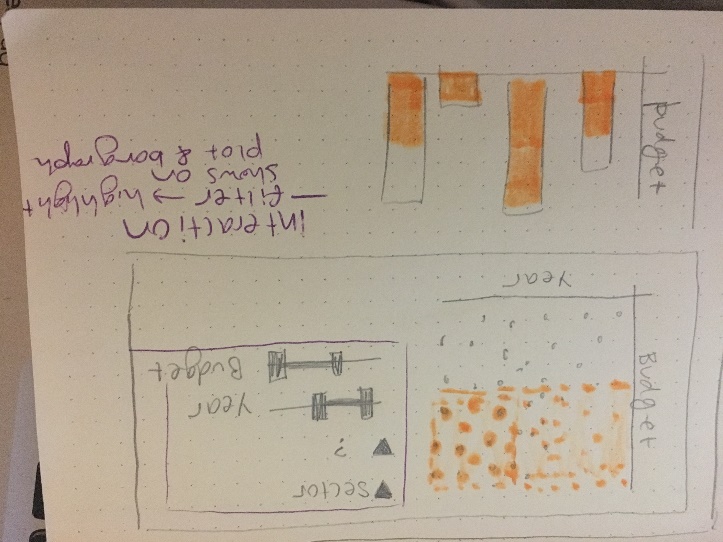
**The Data: NHMRC Research Grants**

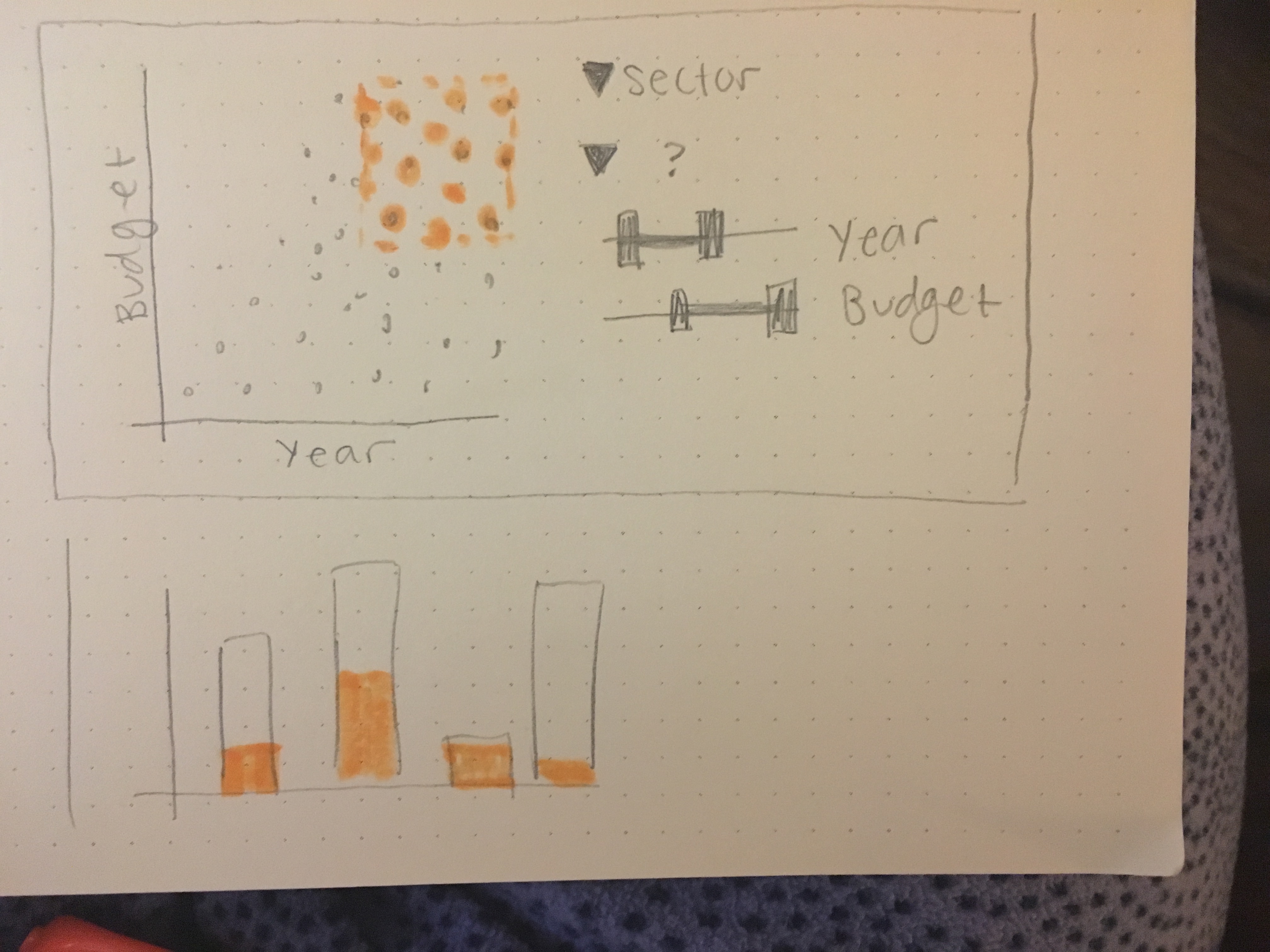
Data Source: <https://www.nhmrc.gov.au/grants-funding/research-funding-statistics-and-data>

This is where I received my dataset on funding for Research in Australia, which has all been compiled over years of projects, including a number of aspects of each of these projects that has been gathered. I chose to focus in on an interesting aspect of this data that I was curious about – what institutions were running these research projects, which is the “sector” in this dataset.

Based on the vast number of research projects that exist in this dataset, and the overwhelming amount of information, plotting all projects has a rich amount of information to start with, however there are some interesting trends here to be explored, which is why I wanted to begin by plotting all these projects by budget, but also offer the option to filter down the amount of data being selected. As you can see in my initial draft of the visualization, in addition to showing a point for each project, encoded by coloring each section, given that this was overwhelming to actually analyze each sector, I chose to use a bar chart to most easily compare the budgets that were allotted to research projectstaking place at each given type of institution, which is given by the Sector on the x-axis of the bar chart in my final visualization.

**Storyboarding and Domain (part 1)**

This was my storyboard of what I planned my visualization to look like when a range of values were selected with the sliders on the right. I wanted to show all the data, but have the selected region highlighted that fits in the filters.



As seen in the second image, when a larger region of years is selected, the highlighted region would be distributed farther along the x axis, which would also be mirrored on the bar chart, which shows that portion of the total budget is covered by the highlighted area.

Data Displayed:

* Project as a point on the scatter plot, showing the year of the project and the budget for that project encoded by the location of that point on the graph
* Budget totals by sector that hosts the research, which is linked to the selected project
* Filter by sector in plot and select particular range

Interactions:

* Highlighting by range selection
* Filter by dropdown options
* I did not draw this on the storyboard, however I also wanted to be able to show the transition from the application year 🡪 start year 🡪 end year, by moving the points on the graph to a new position on the x axis

**My Implementation of the Data Visualization (Part 1)**

I chose to demonstrate the data in the same form as my storyboard, with a scatterplot of the data with year by grant budget for each project, along with a bar chart of the total budget for a full sector, however there are a few differences in the interaction with this data from my original plan, based on further exploration of both natural associations made by a user, along with other aspects that made more sense or became clear upon building the visualization in d3. The implementations of interaction, along with the reasoning for these choices are as follows:

* Filter by sector: This dropdown allows the user to view any sector by itself on the plot of the projects, in order to compare the distributions, which removing some of the information that might crowd the visual.
* Filter by Open/Closed: This dropdown allows the user to look at projects that are still open, vs those that are closed, and have already ended. There is a slight implementation issue with this feature, that it is losing a portion of the data, especially noticeable in the
* Select Year Animation: There is button available to select the year to display on the x-axis of this visualization by application, start year, or end year. This animated between these phases in order to allow one to see the length of particular projects and even visually compare projects by the amount of time (represented by displacement in the x direction) of the points.
* Brushing: I implemented this in a different way that my original intention of filtering. I had originally planned to select ranges by highlighting, however given the detail of this data, once learning more about brushing, I chose to implement domain and range selection in a different way for a number of reasons, as follows: (1) a user clicking and dragging on the image gives them a better visual idea of the range that they are choosing for quick exploration. (2) They get to zoom into that region in order to see more detail on those data points, this cannot be done by just highlighting the region, since the points will still be small, causing significant occlusion.

**Changes and Feedback (Part 2)**

After part 1 feedback, I realized that I needed to make the following changes to my second iteration:

1. Data was unclear, and it was hard to interpret the meaning. This applied to various aspects of my project, namely the following:
   1. **Axis labels**: Without having axis labels, it was unclear what each point in the data actually represented. That being an issue, readers could not easily interpret the meaning of that data. For this, I added axis labels for x and y axis’ on the scatter plot to clarify the data, along with just the y axis on the bar chart, because I believe that the labels of sector made this fairly clear.
   2. **Titles**: I updated the title on the scatter plot, and made it transition names to fade in and out so that the reader would have a better chance of seeing a change in the title when the year transitions between application, start, and end years of the research project.
   3. **Legend**: I added a legend to show the color encodings of the sector that were present in both the bar chart and the scatter plot. I changed the opacity so that these all matched one another. In the first iteration, I thought that people would be able to associate the color of the bar chart with the scatter plot, but having the bar chart follow the scatter plot made this difficult. I still have the legend off to the far side in order to fit all the information within a single view of the screen, however I believe it is fairly intuitive for the reader to see this legend.
2. The range of my data was not representative of the data. I discovered that the budgets of projects were stored as a string and was setting the maximum to a smaller value so that not all of the information was displayed. I solved this by editing the axis, and having it scale correctly to always display all the data.
3. I updated the layout of the visualization so that all of the data could be viewed at one, which also helps the changes in both visualizations to be more clearly linked to one another.
4. I moved the clear brushing button to be added above the scatter plot because the location makes it more noticeable to the reader when they brush the plot into a smaller view.
5. I also changed the axis’ to be defined at the beginning of the program so that they would not be redrawn with updates to the filters or brushing of the data.

**Development Review**

Working the development of this visualization was a fairly long and difficult process. I found that the structure of the functions and placement of variables was the most difficult part. When I began this visualization, I started from the in class example file, applying my own data, but restarting variables to a new structure to try to link the two visualizations to one another. However, in my early iterations of this visualization, I struggled with where to define variables, and where to draw particular aspects of my visualization, which made it difficult to adjust them.

The most time consuming part of this project was learning the workings of d3, and trying to apply their functions to the structure of my code. Often the hardest issues were minor things, such as updating the formatting of the axis’ of my graphs. Doing research to find and apply concepts in d3 proved to be more challenging than I expected, and more time consuming. In total, I spent around 10-12 hours per week in the last 2-3 weeks of the quarter in order to complete this project. I was fairly unfamiliar with JavaScript, which caused this project to be time consuming.

Overall, towards the end of the project, becoming more familiar with d3, these tasks were easier, however I was not able to restructure the functions to make the brushing a transition effectively, which I would have liked to do had I had more skills in d3 and more time to complete this project. Overall, developing this program was a good experience and helped me to learn the materials more effectively.