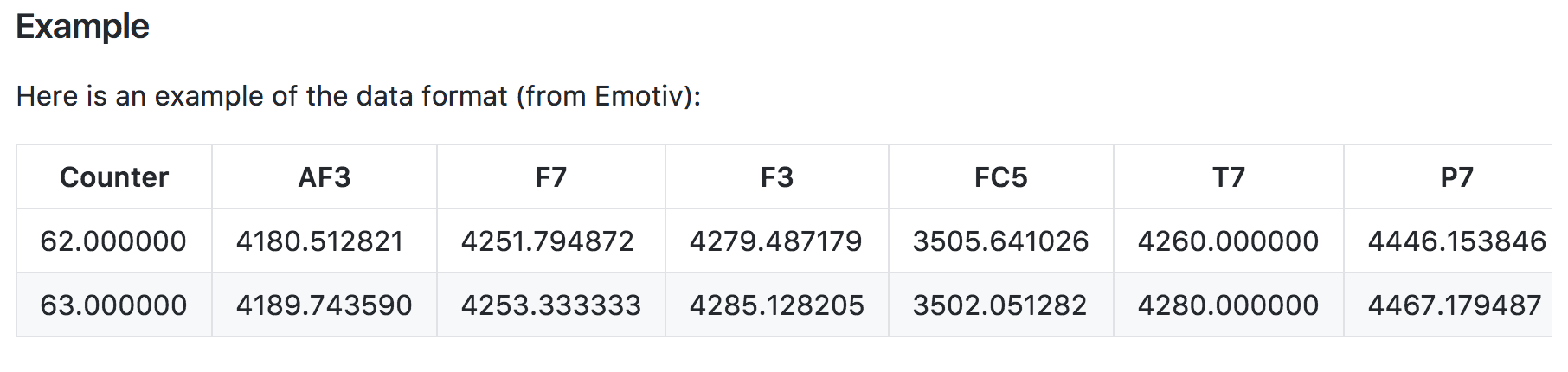
**1. Description of the Data Domain and Storyboard**

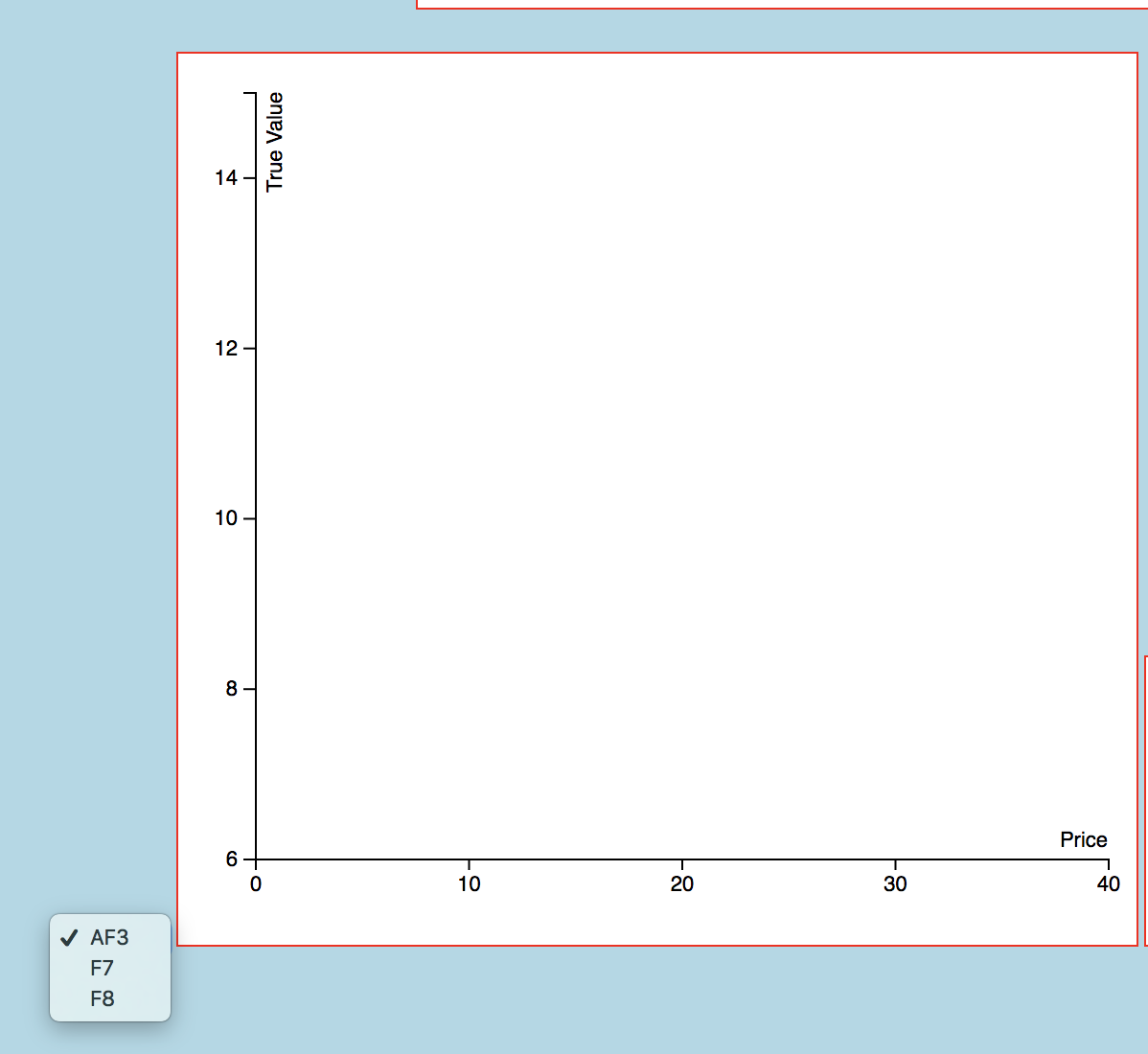
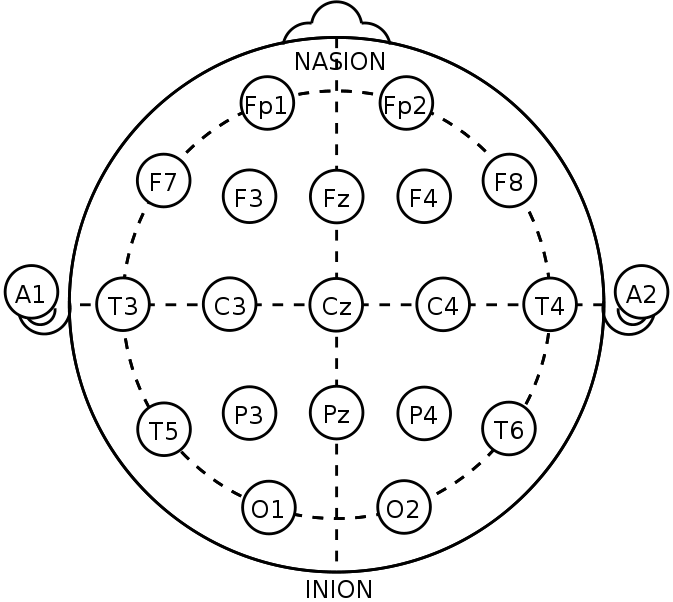
The data domain I have chosen for this assignment is electroencephalogram (EEG) data. As a brief overview, EEG is a non-invasive method of measuring brain signals through attaching external sensors onto the electrodes on the human scalp. These electrodes, or channels, measure brain activity over a period of time and the results per channel are corroborated to determine the subject’s sensitivity to stimulus. The uses of EEG data varies from measuring attention span and symptoms of epilepsy, to determining brain damage from concussions in sports. In particular, the individual alpha frequency (IAF) peak has been the most common tool used to study variability of EEG rhythms across subjects. The dataset for this assignment was provided by Dr. Andrea Stocco of the Cognition and Cortical Dynamics Laboratory (CCDL), a part of the UW Psychology Department. Currently, my Capstone group is looking to develop upon Dr. Stocco’s existing data analysis work on Github and creating more accessible data visualization is included in our goals. Thus, I have taken the opportunity to explore the EEG domain further through this assignment.

In planning for creating a visualization tool for this assignment, I used Dr. Stocco’s existing data pipeline on Github ([Source](https://github.com/uwccdl/qeeg)) as a guideline to visualizing EEG data. The particular repository in question leverages R functions to perform the necessary transformations of the EEG time-series data into a frequency table of log values. Using the “analyze.logfile” function converts data in the Emotiv Epoc format into three primary tables: summary, spectrum and coherence. The function also outputs N number of spectrograms and N\*(N-1)/2 number of coherence plots in the PDF format. However, it is obvious that the amount of data output from this single function can be overwhelming for the journeyman EEG researcher. Thus, for the scope of this assignment, I propose to create a tool to compare spectrum data in the form of line graphs, across different channels of the brain. The main function of the tool is to be able to visualize individual data points as well as pinpoint the IAF of a spectrogram. I feel that the Javascript D3 library will serve me well for this particular undertaking as it easily takes in the tab-separated table values provided in the repository, and can construct line graphs in the scalable SVG format.

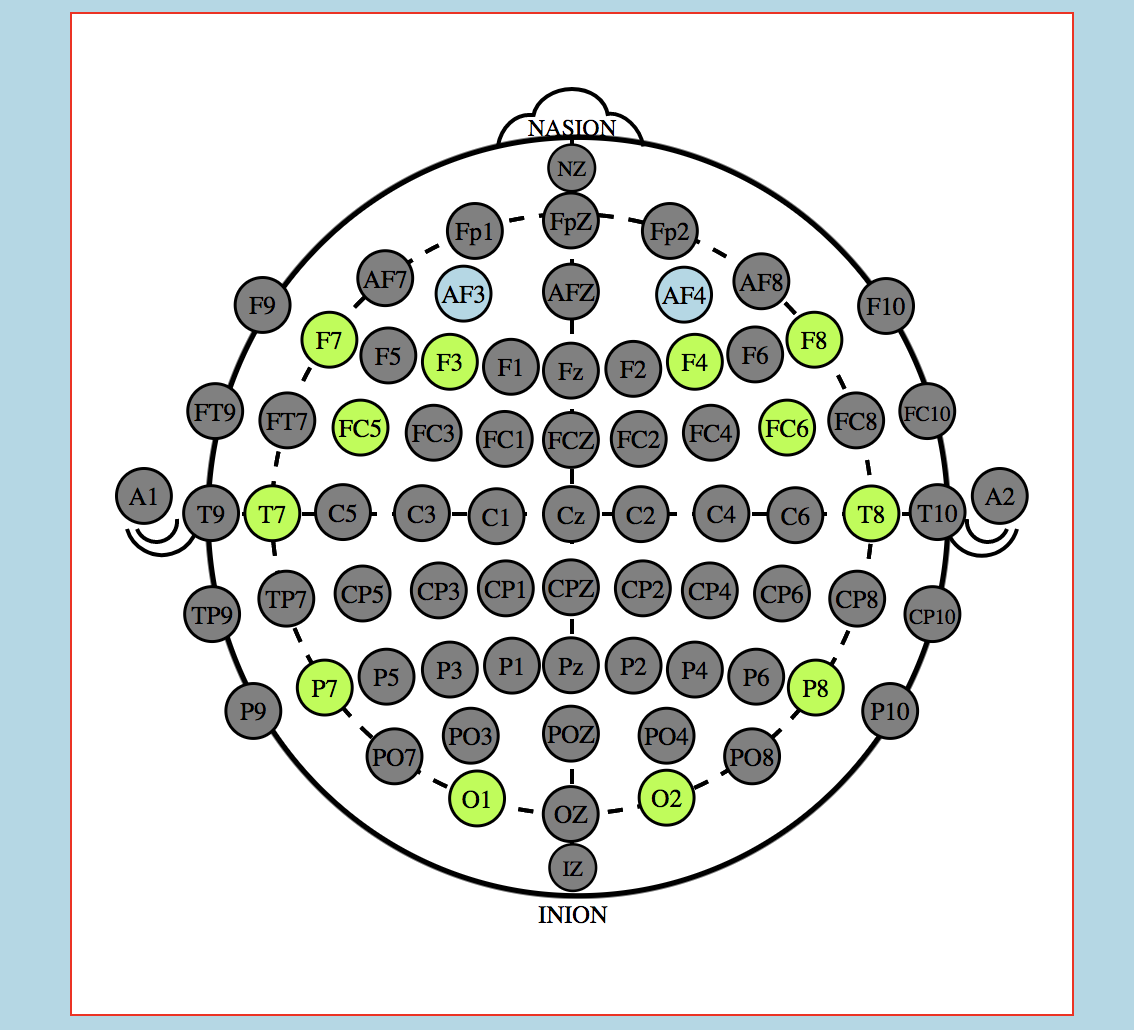
**2. The Iterative Process**

Going into the assignment I realized I had very limited knowledge of EEG data analysis. While I knew the different channel attributes in the data table corresponded to the electrodes on the brain, the nomenclature is pretty vague, as seen in the image below.

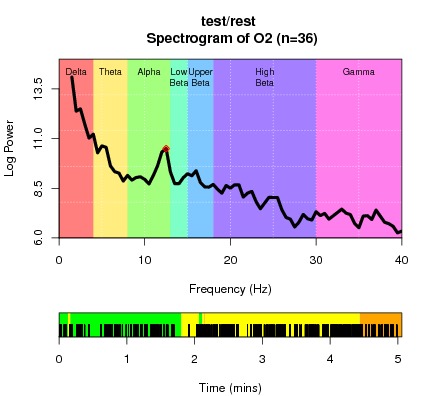


My initial idea was to create two dropdown boxes using the class code that would allow the users to compare spectrograms of different channels side-by-side. However, I ended up questioning myself: Where is AF3, F7 or F8 located on the brain? The brain is a fascinating construct, yet it is difficult to break down different the different portions without getting too scientific. I decided to look into the origins of the nomenclature and look up diagrams of the distribution of electrodes on the brain, and discovered the 10-20 system.

According to Wikipedia, the 10-20 system is “an internationally recognized method to describe and apply the location of scalp electrodes in the context of an EEG test or experiment.” The locations of the electrodes tend to be fixed so that repeated studies can be conducted an their results corroborated with each other. The diagram on the left is a simple, lower-resolution featuring 21 channels. However, the data in Dr. Stocco’s repository uses channels like AF2, which meant that a higher-resolution diagram is used. I also discovered that the diagrams on Wikipedia were originally in the SVG format, meaning I can manipulate the color and behavior of the individual circles in the diagram. To this end, I created a clickable 10-20 diagram for the visualization tool.



The logic is simple: there will be a prompt to pick two of any green nodes (grey means data is unavailable for the particular channel). The blue circles on the diagram on the right shows that the user has chosen two channel datasets to look at, and at the same time they can also see which particular part of the brain they are looking at. This allows them to easily distinguish between symmetrical, left/right, frontal/rear lobes and particular electrodes located near the ear. While the given data only has 14 channels data, the tool can be expanded upon to feature more channels when necessary.

 The second issue I encountered creating this visualization tool was visualizing the spectrums. EEG data analysis primarily converts time-series data into a frequency band which are color coded respectively for Delta, Theta, Alpha, Low-Beta, Upper-Beta, High Beta and Gamma frequencies. While aesthetically pleasing, the color distracts users from seeing the real important data, which is located in the alpha region in order to determine IAF peaks. Thus, I resolved to highlight the alpha band only in the comparison tool.

Lastly was the consideration of adding another layer of interactivity. While, the PDF file outputs from the Github repository already pinpoints the highest IAF peak (as seen in the green-band red dot in the diagram above), there is also the possibility of multiple IAF peaks in the alpha band that might be worth looking into. To add another layer of interactivity, I decided to create a focus object with tooltips on the line graphs that shows the data points per 0.5 Hz.

The final product allows users to select two desired channels to compare with a high resolution 10-20 diagram, which will then display two graphs side by side for comparison. Both graphs have the ability to view the individual points on the line and the alpha-band region is highlighted for IAF measuring purposes. The reset button deselects the current channels and allows the users to choose 2 new channels data to look at.

**Final Commentary**

A significant portion of the development process (5-8 hours) was spent reading the R code on Github to understand the functions and variables used for EEG data analysis. As part of Capstone I also worked on breaking down the function to sub-functions that can be put on API endpoints. To this end I considered working with the data that was not part of the existing output. However, I decided to settle with the example spectrum.txt data for my final iteration of the tool.

Programming the tool also took around 10 hours of development time, largely a mix of trial and error in getting the SVG elements in the Wikipedia image to behave according to my vision with basic Javascript, and leveraging the D3 library to visualize differently from Dr. Stocco’s diagrams. One particular issue I ran into was online guides having older versions of D3 syntax which complicated my debugging phase. The final report took around 2 hours to write.