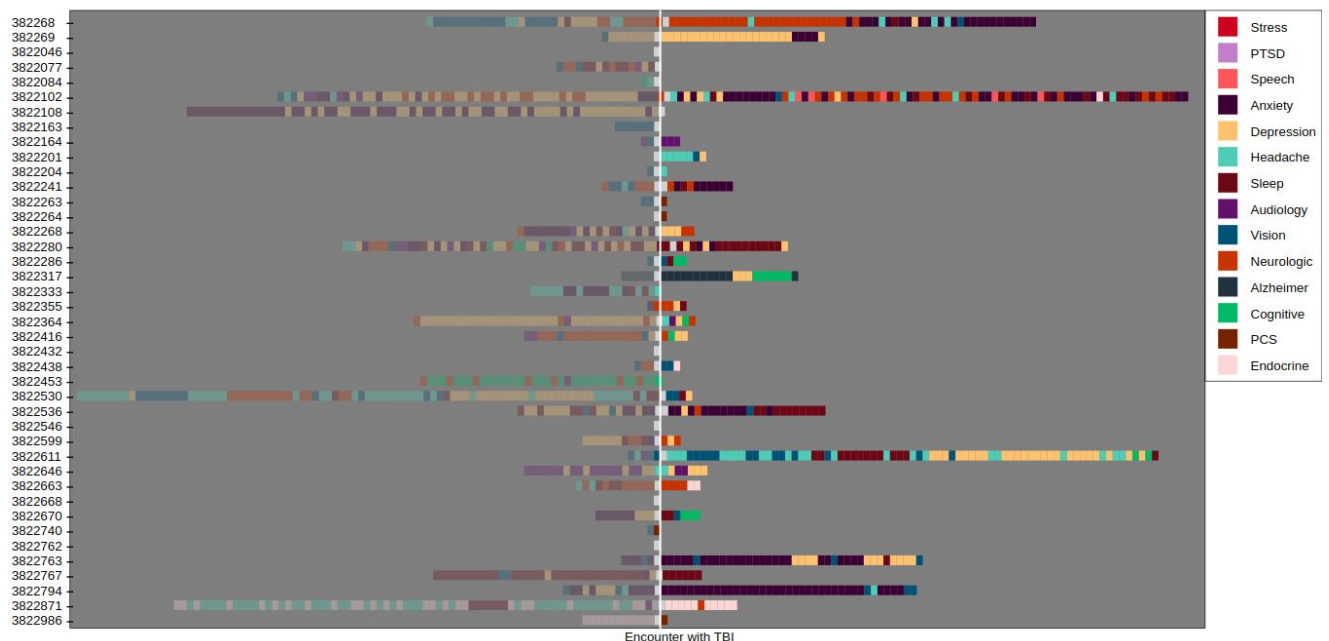


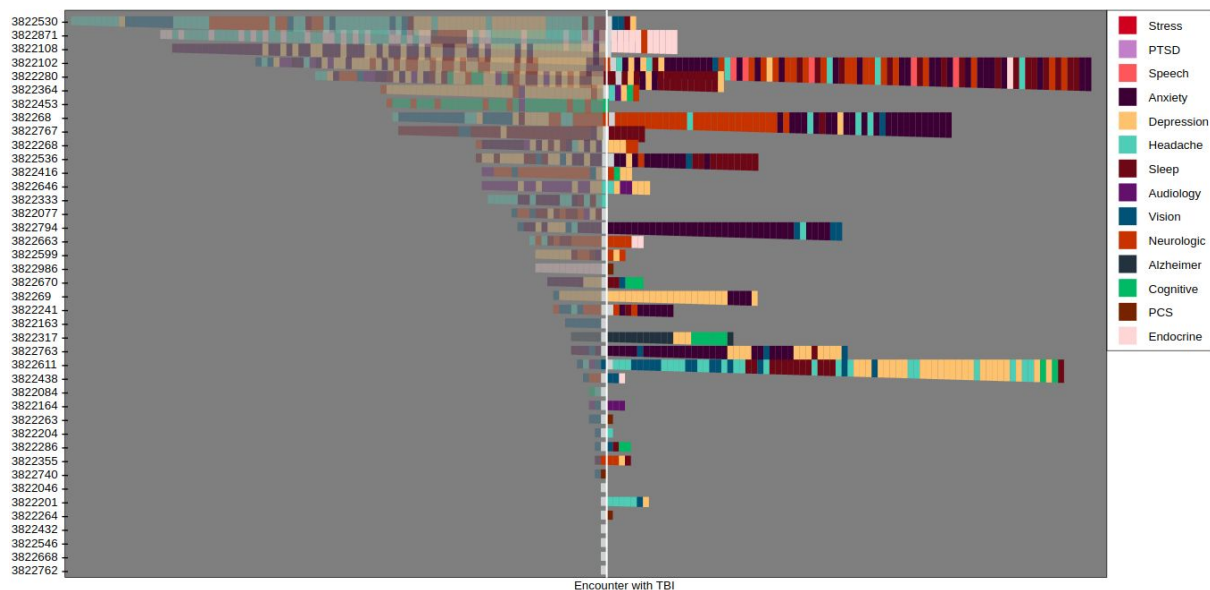
Assignment-3 Part Two

Data visualization task: For this task two different visualizations of the Electronic Health Record dataset. The **first graph** is the baseline graph that the assignment asks to implement. Each horizontal component describes the encounters of the for a single patient. Only the encounters that showed symptoms have been shown. All the patients have been aligned along the encounter where the TBI occurred. (Some patients have light grey encounters in the middle. These are TBI encounters where no other symptoms were found).



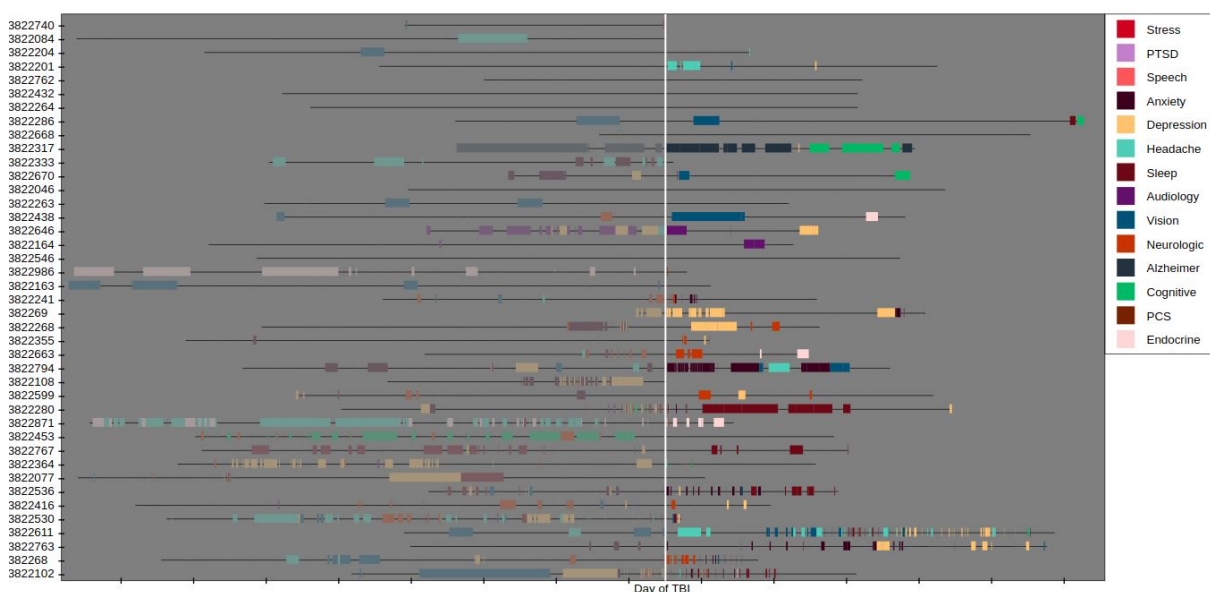
Graph: Baseline Graphs

This graph has two modes of interactivity. 1. The patients' order can be changed according to different criteria namely, 'ID Number', 'Earliest first encounter before TBI', 'Effective Position of TBI in timeline' and 'Number of encounters'. 2. The graph can be switched to 'incremental size' mode where the size of each encounter square increases with the order of encounter in the timeline. This is useful because the patients have been aligned according to TBI encounter and 'incremental size' mode tells whether the encounter is an old (small square) or new (big square) encounter. The graph below shows such an example for the **Baseline Graph**.



Graph: Baseline Graph. Patients are ordered according to 'Effective position of TBI in timeline' and the 'incremental size' mode is switched on.

The **second graph** is one of the extra tasks where the graph shows all encounter (even without symptoms) and the graph shows the actual timeline (the day on which encounter occurred). The ticks on the x-axis are separated by a distance of 1 year (365 days). For this graph, we have the same patient ordering interactivity (4 buttons). Each encounter is either represented by a rectangle (with symptom) or a line (without symptom). The rectangle width and line length represent the time between consecutive encounters. Tiny ticks on the timelines represent the actual encounter (this is useful for symptom-less encounters).



Graph: Temporal Graph

Note: Since d3 v5 does not have support for 'category20c' we use random colors for the symptoms.

We now use these two graphs to answer the questions in the assignment part 2.

Change task 1: How have the encounters reduced from year 1 to year 2? (you may report that 30% fully recovered, yet the rest progressed to a more severe stage.)

Change task 2: How has the frequency of encounters (hospital visits) changed from year 1 to year 2?

Change Task 1: For 25 (~61%) patients the second year showed improvement as the symptoms did not appear in the subsequent encounters (eg. 3822599). But for some cases like patient 3822871 symptoms (Headache and Endocrine in this case) are seen throughout the first two years. The temporal graph is more useful in this case.

Change Task 2: For 18 patients the frequency of visits doesn't change much (eg. 3822077). For other patients the frequency increases (eg. 3822364) or decreases (eg. 3822599).

Observations

Pros:

1. This visualization captures the temporal aspect of the encounters.
2. A cohort can be analyzed by adding in another ordering such that patients with similar symptoms are together (using metrics like Jaccard coefficient etc.)
3. The temporal graph (temporal.html) also facilitates analyzing year to year changes.

Cons:

1. The dataset contains information about 41 patients. Putting everything on the same screen might be a problem if the user's screen is not big (for instance I use a 14-inch laptop).
2. Some patients have a very large number of encounters. This makes reading the information on timeline difficult.
3. The time between consecutive encounters varies a lot. Due to this, some encounters are long rectangles while some appear as thin lines. (width represents time b/w consecutive encounters)

Change task 1: How have the encounters reduced between the first and the second three months?

Change task 2: How has the frequency of encounters (hospital visits) changed from the first to the second three months?

For these tasks, we change the x-axis ticks to 90 days separation (instead of 365 days).

Change task 1: For 11 (~27%) patients the encounters get reduced (eg. 382268). (use temporal3month for both the tasks)

Change task 2: For 20 (~49%) patients the encounter frequency does not change much (eg. 3822077). For other the frequency increases (eg 3822333) or decreases (eg 382268).

Extra Tasks:

Temporal Graph

Ordering for Patients

Note: 4b is in file 'DesignDiscussion.pdf'