Preprocessing:

Used R for preprocessing. Took the database and found all the symptoms occurrence in the time frame. Each change in symptom over time is recorded as one entry. That is, if depression is followed by anxiety for one patient (depression anxiety) is one entry. This time series pair is calculated for each patient. We can use the subset of this preprocessed data for the visualization (only 16 different shapes can be used).

Visualization:

New\_data.csv have all the symptoms transitions for all the patients with date encountered.

Specs.csv is the sample data for processing.

Test.html is the d3 code which takes the specs.csv and maps to the new\_data.csv occurrences.

I am not able to plot the visualization. The processing and encoding is correct, but have some issue in plotting.

Critical Evaluation:

Pros:

It’s a good temporal visualization to see the relevance between the symptoms. We can conclude saying that if this is the current symptom, patient might probably get the next one. It will be like precautionary measure which can be taken to avoid the next one.

It also helps in identifying the kind of disease it might lead up to.

Cons:

The limitation of only 16 different structures limits to check the propagation for all the different kinds of symptoms. Color is one of the aspects which can be used (within shapes) to distinguish symptom transition.