

## **CSE 564 LAB 2B REPORT**

**By Anujay Ghosh**  
**SBU ID: 115362293**

The following are the interesting observations on top of the ones mentioned in the demo video:

We can observe that the multi-dimensional scaling plot function has a "random state" variable, which takes in a random number as the seed for initializing the centers. Without defining a consistent random state number every time the browser is refreshed, you will notice that a different kind of graph is generated with the data points being spread out differently each time. This doesn't mean that there is a correct or wrong multi-dimensional scaling plot generated by the software or the Flask server, but in fact, it's because of the random state not being consistently defined.

On observing the MDS data plot, which has been segregated by cluster colors (blue and orange), signifying cluster 1 and cluster 2, at first glance, it looks like there are more points in the orange cluster. This is because the orange cluster looks spread out while the blue cluster looks concentrated in a small space. But on observing the opacity of the data points, you can notice that the blue points have a very high opacity- they're very dark, and their opacity is close to or even completely one at some places, whereas most of the orange points have very light opacity, signifying that most of the points in the blue cluster are projected to be very close to each other and very high in intensity, whereas the orange cluster points are really spread out and, in fact, low in number compared to the blue cluster points.

Moving on to the second data plot, which shows us the variables (attributes) in MDS, I have signified each attribute with a triangular data point and the name of the attribute next to it. The distances here are calculated using one minus absolute correlation distance, whereas when compared to the previous data plot, the previous data plot had points separated by to each other by euclidean distances.

The third graph in the project is the parallel coordinate plot, where you can observe the skills of different axes very significantly. Some values are categorical with the whole thing, such as the month number, and some are extremely high in the number of distinct values, such as attributes like start date and end date. Some numeric attributes are in the range of thousands, some in 100,000, and some even go on to multi-millions. Since we have the plot lines according to the key means clusters, we color code them in blue and orange. If you carefully look at the bank name access, we notice that all of the orange cluster data points come from State Bank of India, which is exponentially bigger than all the other four banks mentioned in the graph. The number of ATMs off-site and on-site taxes next to each other at the bank names, and we see that State Bank of India happens to have a very high number of ATMs in the country, where the other banks' ATMs are in single digit thousands versus 30 to 40,000. Observing the grandfather, we see that most of the plot lines then towards the lower side of the graph, forming a pyramid shape, signifying that even though there are high numbers in the scale, they stick to the lower

end of the ranges that we have mentioned. So most of the entries call numeric values on the axis that are not that high. We can also observe the orange lines for a very high tendency to go upwards in the number of debit cards access, signifying that State Bank of India has a very high number of debit cards to the citizens of the country of India. In a way, this can be an indication that since the number of debit cards issued by State Bank of India is so high, the government is trying to promote cashless forms of transaction and encouraging the citizens of India to use plastic cards, which do not deal with the complexities of microloans and debt but, in fact, allow you to access your own money without having to use cash as legal tender.

It can be inferred that the financial trends in the country of India have been moving towards a cashless economy. This is evident from the fact that the government is promoting the use of plastic cards through the high number of debit cards issued by State Bank of India, which is the largest bank in the country. The increasing number of credit cards and POS over the years also suggests that the country is moving towards a more digital and convenient form of banking, purchasing and spending. Additionally, the fact that the blue cluster in the MDS plot has a high concentration of data points with high opacity suggests that there are certain financial trends that are more common or prominent in the country. Overall, the financial trends in the country seem to be moving towards digitization and convenience, with a focus on promoting cashless transactions.