**Epoch Cores Selection - Round 1**

**Task 1**

**Name: Aparna. S. Kaushik**

**Objective**

You have been provided with a dataset containing information about various pincodes across India, including their corresponding longitudes and latitudes (clustering\_data.csv). Your task is to focus specifically on the pincodes of your home state

**Process Followed:**

1. **Data Filtering and Preprocessing:**

* Converting the 9th and 10th column using convert\_column function while loading the dataset.
* The 9th and 10th columns contain data of Longitude and Latitude and convert\_column checks if it contains a float data type if not replaced by NAN value.
* Then, the data containing NAN was removed in the 9th and 10th columns.
* Filtered data for my Home State (Tamil Nadu).
* Tried removing duplicate data in the filtered data, but there was none.
* Using the percentile concept and some trial-and-error methods, I removed the outliers present in the data. As some points were not even close to the state.

1. **Data Visualization**

* Visualized the data using cartopy.
* Using this only, I was able to filter the outliers.

1. Implemented k-means clustering from scratch.

**Inference and Insights:**

1. **Based on Longitude and Latitude**

* Plotted the result
* Tried to derive some inferences.
* My inference from the clustering is that:

1. The whole state mostly has the Delivery and BO office types.
2. It does a decent job of clustering such that clusters have equally distributed data points.
3. But all clusters have almost the same observations.

So, the inference can be that all the regions of the state have similar features.

1. **Based on Longitude, Latitude and Pincodes**

* Plotted the result
* Tried to derive some inferences.
* While visualizing the map, we can say that it is a poor clustering. But when we see the clustering inferences, we can say that the data points are separated according to the division.

But still, BO office type and delivery type is prominent. Thus, we can conclude that the whole state is consistent in terms of office type and delivery type.

**Reference Links:**

* https://www.youtube.com/embed/4b5d3muPQmA?si=2MWRRsMuAEiicsEy