

Graph Clustering and Partition:

Goal: To create clustering to evaluate experience and age of workers, prioritizing experience but also taking into account age. This helps hiring and evaluation of the experience of workers.

Main.rs: In main.rs, I set up the clusters, loaded the data and important dependencies, I also visualized the clusters and grouping using

clustering.rs: In clustering.rs, I transformed data, using a vector to use for clustering, I randomly initialized centroids using algorithm of clustering, and assigned points to clusters if close to center(centroid).

utils.rs: In utils.rs, I created load_employee_data to load in necessary data and unwrap so Rust can output correctly to readable format, I also iterated to print out each row

graph.rs: I imported and implemented employee graph to import graph and hasmaps

employee.rs: I deserialized here and put in a struct to be able to correctly read the file

centrality.rs: I imported ndarray and rand, a k-means function, and euclidean_distance to calculate distance and means.

Conclusion:

The output demonstrates each cluster centroids, but also information regarding each employee, including Name, Employee ID, Age, Experience, etc. but also the cluster it belongs to, separating different employees into clusters in which they have similarities. The Features used are age, salary and experience. As a result, cluster 1 has less experience than cluster 2, 2 has less experience than 3, etc. Experience is prioritized in my clustering, simply because they are more important for job hiring. Cluster mean experience is shown below as well as mean age.

Cluster 0: mean experience = 2.48

Cluster 1: mean experience = 4.24

Cluster 2: mean experience = 5.51

Cluster 3: mean experience = 7.15

Cluster 4: mean experience = 9.03

Cluster 5: mean experience = 10.49

Cluster 6: mean experience = 12.66

Cluster 7: mean experience = 14.94

Cluster 8: mean experience = 16.44

```
Employee Salary Data (1974-1980)
Cluster 2: Mean Age = 39.29, Mean Experience = 5.51
Cluster 0: Mean Age = 26.37, Mean Experience = 2.48
Cluster 4: Mean Age = 40.68, Mean Experience = 9.03
Cluster 9: Mean Age = 55.52, Mean Experience = 18.26
Cluster 5: Mean Age = 43.01, Mean Experience = 10.49
Cluster 7: Mean Age = 40.65, Mean Experience = 14.94
Cluster 8: Mean Age = 44.82, Mean Experience = 16.44
Cluster 6: Mean Age = 41.77, Mean Experience = 12.66
Cluster 1: Mean Age = 34.91, Mean Experience = 4.24
Cluster 3: Mean Age = 40.80, Mean Experience = 7.15
f:/opt/anaconda-root/src/final project/final proj1
```

Things to improve:

The project fails to recognize how much different features impact the clusters, as one feature like experience, may be more important than salary or age, especially when one requires experience for hiring more than age, although it may be a factor. The exact amount to scale age to experience is unknown and needs to be improved upon.