ML ASSIGNMENT 4

Video link: <https://drive.google.com/file/d/1oPKUEB3YldDE0MefRDLPjkxt25Z3U7JI/view?usp=sharing>

1. Apply Linear Regression to the provided dataset using underlying steps. a. Import the given “Salary\_Data.csv” b. Split the data into train test partitions, such that 1/3 of the data is reserved as a test subset. c. Train and predict the model. d. Calculate the mean squared error e. Visualize both train and test data using a scatter plot

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

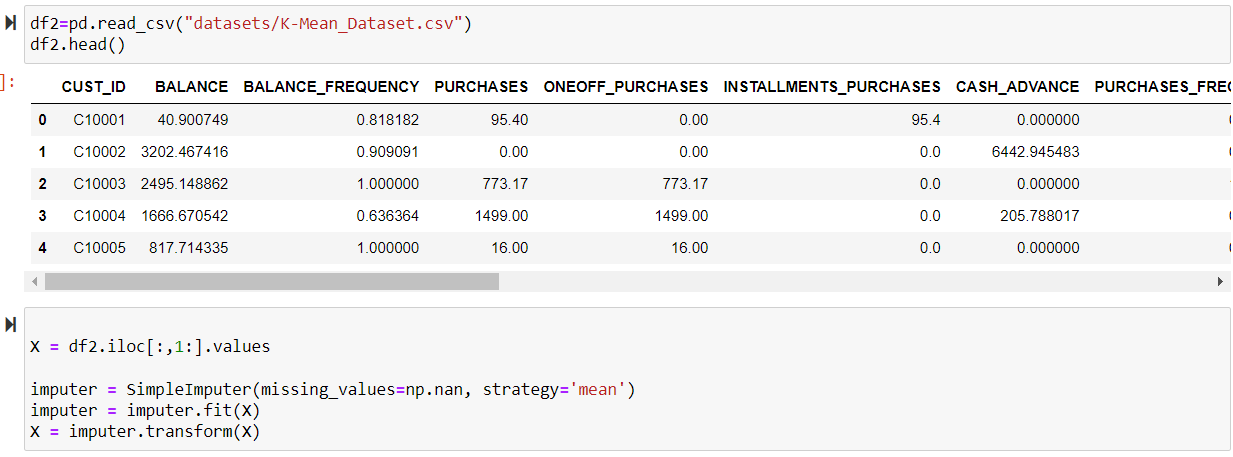
Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

1. Apply K means clustering in the dataset provided: • Remove any null values by the mean. • Use the elbow method to find a good number of clusters with the K-Means algorithm • Calculate the silhouette score for the above clustering



Graphical user interface, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

1. Try feature scaling and then apply K-Means on the scaled features. Did that improve the Silhouette score? If yes, can you justify why

Graphical user interface, text, application, email

Description automatically generated