**Scalend Data Science Internship Assignment**

**Q1)** Given is a sample dataset containing bank transactions data. Some of the data elements are transaction id, customer id #, mode of transactions, etc.

We would want you to come up with algorithms (single or combination) that can segment users into different categories.

Make suitable assumptions.

Dataset attached – sample\_data\_test

**Q2)** If there are two datasets given (ideal situation - 2 different tables from a database), autodetect relationships between the two data sources by establishing cardinality and other relationship among the columns.  
Ex: 2 tables Table\_A and Table\_B have 2 columns that have same data: Table\_A.customer\_id is similar to Table\_B.id

Dataset attached – Bank\_branch\_transactions ; Bank\_customer\_db