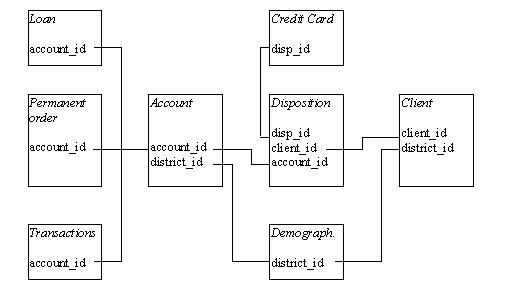
**A Capstone Project on Data Science and Analytics –**

This project will showcase an end-to-end solution of Data Science processes and technologies.

In this project the raw real-time data has to be stored in the staging area (Database) and from there all the ETL functionalities has to be performed.

1. Datasets Entity mapping -



With the help of this entity mapping, you have to prepare a Master table that takes into consideration all the related tables mentioned above. Make sure all the clients id has been considered in the master table which is common to all.

The Datasets descriptions are as follows:

* The data about the clients and their accounts consist of following relations:
* -relation account (4500 objects in the file ACCOUNT.ASC) - each record describes static characteristics of an account,
* -relation client (5369 objects in the file CLIENT.ASC) - each record describes characteristics of a client,
* -relation disposition (5369 objects in the file DISP.ASC) - each record relates together a client with an account i.e. this relation describes the rights of clients to operate accounts,
* -relation permanent order (6471 objects in the file ORDER.ASC) - each record describes characteristics of a payment order,
* -relation transaction (1056320 objects in the file TRANS.ASC) - each record describes one transaction on an account,
* -relation loan (682 objects in the file LOAN.ASC) - each record describes a loan granted for a given account,
* -relation credit card (892 objects in the file CARD.ASC) - each record describes a credit card issued to an account,
* -relation demographic data (77 objects in the file DISTRICT.ASC) - each record describes demographic characteristics of a district.

* Each account has both static characteristics (e.g. date of creation, address of the branch) given in relation "account" and dynamic characteristics (e.g. payments debited or credited, balances) given in relations "permanent order" and "transaction". Relation "client" describes characteristics of persons who can manipulate with the accounts. One client can have more accounts, more clients can manipulate with single account; clients and accounts are related together in relation "disposition". Relation’s "loan" and "credit card" describe some services which the bank offers to its clients; more credit cards can be issued to an account, at most one loan can be granted for an account. Relation "demographic data" gives some publicly available information about the districts (e.g. the unemployment rate); additional information about the clients can be deduced from this.