**Hierarchical clustering:**

Dataset:

Purpose of this mission:

In this I have used a dataset of the mall. It contains following details about the customers of the mall

Customer ID

Gender

Age

Annual Income

Spending Score which ranges between 1 to 100: it’s a metric made by the mall

Lower the score, less the customer spends

More the score, more he spends

As a member of analytics team, I try to understand the patterns in the customers

As opposed to regression and classification, where we know what to predict

Here, I do not know what to predict

**How to identify**:

Step 1:

Make each datapoint a single-point cluster -this will form N -clusters

Step 2:

Take the two closest data points and make them into one cluster

Step 3:

Take two closest clusters and make them into one cluster

Step 4:

Repeat the above step (step 3) until there is **only one cluster**