Hello,

This mail is in regard to the data quality assessment in the Customer Demographic dataset of Sprocket Central Pyt Ltd. There are certain issues with data regarding accuracy, completeness and duplicate values. Some of these are pointed out below:

1. In gender attribute, values vary like Female, Male, F, Femal, blank etc. All these can be categorized in 2 values only, Male and Female.
2. There are multiple records with 0 purchases over a century. There is no need to maintain such records.
3. Multiple attributes have some blank values like Job title, Job Industry category, Last name have blank values or n/a which doesn't help the analysis.
4. Certain attributes like decreased indicator, default only have garbage data that does not make sense.

I believe if we work on these issues of dataset and then start analysing patterns of customers then we will be able to develop a better and optimized marketing strategy for the organization.

Thank You;