Dear team leader,

I saw the data you provided me, This data need a lot of prepressing and modification before analysis to have good insights from it. I will provide them by using **Data Quality Framework Table in this 3 datasets (Customer Demographic - Customer Addresses - Transaction data in the past three months):**

**# In** **Customer Demographic data :**

* **Last\_name column has NAN data (not completeness)**
* **Gender column has M, F, U. femal and that is wrong choice must be only female or male (not relevancy, not validity)**
* **DOB column has wrong data because people can leave to just 100 year then rows that has people live from before 1923 must be wrong ( not acurrency, not validity)**
* **DOB column has NAN data (not completeness)**
* **Job\_title has NAN data (not completeness)**
* **Default column has unimportant and wrong data ( not acurrency, not validity)**
* **Tenure column has NAN data (not completeness)**

**# In Transaction data:**

* **Last\_name column has NAN data (not completeness)**
* **Gender column has U and that is wrong choice must be only female or male (not relevancy, not validity)**
* **DOB column has NAN data (not completeness)**
* **Job\_title has NAN data (not completeness)**
* **Address column has zeros in the address as the number of the house and that is wrong because there is no house numbering by zero (not Accuracy)**
* **address and postcode column has wrong values because different places must have different postcode but there is in the data duplication in postcode with different places (not Accuracy, uniqueness, not consistency)**
* **property\_valuation column has single numbers but that is not normal, the normal to be by thousands at least (not Accuracy)**

**# In Customer Addresses data:**

* **Address column has zeros in the address as the number of the house and that is wrong because there is no house numbering by zero (not Accuracy)**
* **address and postcode column has wrong values because different places must have different postcode but there is in the data duplication in postcode with different places (not Accuracy, uniqueness, not consistency)**
* **property\_valuation column has single numbers but that is not normal, the normal to be by thousands at least (not Accuracy)**

**And the rest other data has all Data Quality Frameworks. So my recommendation to fix all this before analysis to have good knowledge from our data and help our customer.**

Best regards,

**Menna Reda**