Hello XYZ,

Please find below the total quality assessment of the raw data given.

Accuracy:

1. One of the dates under DOB is 1843-12-21, which looks like a data error.
2. Some of the values under gender have values like ‘F’ or ‘M’ instead of ‘Female’ or ’Male’. Also there are several spelling mistakes under this column.

Completeness:

1. Out of 335,994 fields, 2,587 fields are empty resulting in 0.76% of the data fields being unfit for the analysis.
2. The data under Default tab has gibberish values. Hence, this column can be made of no use.

Data Consistency:

1. The two tables i.e Customer Demographic and Customer Address and inter related with common field customer\_id. There are 4000 records in Customer demographic whereas 3999 records in Customer address. There are **4** records of customers that are present in Customer Demographic that are not present in Customer address. Similarly, there are **3** records of customers that are present in Customer address but not Customer Demographic. These 7 records are not fit to be a part of this analysis because of the incomplete data.

Data Currency:

Much on this can not be said as we don’t know the source of the data. If we do, it would have been possible to determine if the data given is current.

Data Relevancy:  
 1. There are several data columns like address, default, property valuation, country, deceased\_indicator that are not relevant for the analysis.

Data Validity:

1. The data under Default tab has gibberish values. Hence, this column can be made of no use.

Data Uniqueness:

1. This can not be determined because if we have 2 date fields with same data or 2 name fields with same name, there is a chance that the same person has purchased the product or there are two people with same name or birthday.