**KPMG Virtual Internship**

**Sujit Kumar Rai**

**B. Tech (CSE)**

**Mail**

Data Quality Issues

* Accuracy
* Most of the column don’t have consistent value un their column. i.e. DOB column consist text formatted value and some column must have integer type but they also contain text values.
* **Completeness**
* Most of the data fields are blank.
* **Consistency**
* In most fields data are not of same data types such that DOB column in Customer Demography also consists of text formatted data.
* **Currency**
* The Age of some customers are more than 110.
* **Relevancy**
* The transaction and address sheet contain customer id more than 4000 i.e. irrelevant.
* **Validity**
* The customers more than 100 year of age are not required.