Hello sir/madam,

Received the datasets provided by you on 01/03/2021 the analysis report of the data given is as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Total records** | **Unique records** | **Receiving date** |
| Transaction | 20000 | 3493 | 01/03/2021 |
|  |  |  |  |
| Customer demographic | 4000 | 4000 | 01/03/2021 |
|  |  |  |  |
| Customer address | 4003 | 3996 | 01/03/2021 |
|  |  |  |  |

I would like to highlight some issues with the datasets provided by you which are

In dataset of transaction there are many additional customer ids with same records while unique id’s are less so here is the possibility of data loss please confirm it once.

Many columns in the dataset have empty values also, have inconsistency in data. So we are continuing the further operations on this data by cleaning it and standardizing it. Glad to operate on your data. Looking forward for a great journey ahead.

Thanking you.

Gajanan Tayde

KPMG