Dear Client,

Thank you for providing us with the datasets from New Hope Clinic. The table below highlights the summary statistics of the datasets received. Please let us know if there is any discrepancy in the figures.

|  |  |  |  |
| --- | --- | --- | --- |
| WORKSHEET NAME | NUMBER OF RECORDS | UNIQUE CUSTOMER ID | DATE DATA RECIEVED |
| Patient Demographics | 257 | 248 | 25/05/2021 |
| Patient Age | 257 | 248 | 25/05/2021 |

The data quality issues encountered and the methods used to rectify the identified inconsistencies are listed below.

**Uniqueness**

Both tables provided contained duplicate customer ids. The following duplicates were found in both tables: **1613(2), 5525(2), 6738(2), 7131(2), 17162(2),18866(3),22087(2)**

**Completeness**

Columns like Patients suggestions, underlying conditions , patients condition contains a large number of empty values. These columns would be excluded from our analysis.

**Keys**

The following keys represents why customers choose to visit the clinic

PROXIMITY- Closeness to home and work,convenience

CUSTOMER SERVICE- one stop for my need, without referral,I feel welcome,efficient service

PROXIMITY AND CUSTOMER SERVICE-closeness to home, good service,family hospital for years

REFERRAL- Referral,HMO

NEW PATIENT-First time visiting