**Data cleaning and quality observations**

**Following are the data quality issues identified in the source data files:**

**Receipts data file:**

**Sizeable fraction of values missing in 'bonusPointsEarned', 'bonusPointsEarnedReason', 'finishedDate', 'pointsAwardedDate'.**

**Users data file:**

**More than half of the records are duplicate.**

**Brands data file:**

**Sizeable fraction of values (more than 50%) are missing from topBrand and categoryCode variables.**