

Dear Manager,

RE: REVIEW OF DATA QUALITY OF DATASETS

After a thorough assessment of the three datasets here are some data quality issues identified:

- Data Completeness issues are evident in the **Transactions and Customer Demographics** sheets. These are evident as columns: 'last name', 'DOB', 'job title', 'job industry category', 'default' and 'tenure' in the Customer Demographics dataset all have null values, same with columns: 'online order', 'brand', 'product line', 'product class', 'product size', 'standard cost', 'product first sold date' in the Transactions Dataset.
- Data Validity issues are evident in the **product first sold date** column (Customer Demographics Dataset) as the data therein aren't in date format. **Default** column in Customer Demographic dataset also have data in special character format. This validity issues will need further clarification.
- Data Consistency issue in DOB column in Customer Demographic sheet where a customer has a value of '1843' as Date of Birth.

Kind Regards

Andrew