Hello,

This is Dennis from KPMG, Data analytics team. After an in depth review of the data you provided us with, the data quality analysis showed that there were some issues with regards to your data quality.

Some of the errors encountered in the dataset were;

* Missing values in the dataset which is as shown below

|  |  |
| --- | --- |
| **FIELD** | **NUMBER OF MISSING VALUES** |
| Online order | 360 |
| Brand | 197 |
| Product line | 197 |
| Product class | 197 |
| Product size | 197 |
| Standard cost | 197 |
| Product first sold date | 197 |
| Last name | 29 |
| DOB | 17 |
| Job title | 106 |
| Job industry category | 165 |

* Inconsistent data type within the same attribute
* Ambiguous data such as the default column in the customer demographic dataset.
* Incomplete information for example where gender was given some alias U.
* Incorrect data formats for example in the date columns

The team came up with some ways to mitigate these issues, some of the mitigation measures included:

* Missing values could be reduced by setting up systems to reject saving incompletely filled data.
* Specify exact formats for every piece of data to ensure consistency across every source your organization uses.
* Replace gender “u” with “Unspecified” to reference to unspecified customer gender.

Thank you for choosing KPMG.

Regards,

Management.