*From,*

*KPMG Data Analytics Department*

*To,*

*Sprocket Central Pty Ltd*

*Hello Team,*

*Thank you for providing us with the three datasets inorder to review the data quality to ensure that it is ready for the analysis of second phase. After thoroughly evaluvating the data using the Microsoft Excel Software, we found some quality issues in data which is provided.*

*Following are more indepth descriptions and findings based on the data quality dimensions*

***Accuracy***

* ***DOB*** *was inaccurate for “****CustomerDemographic****”*

*Recommendation: Need to filter out outlier in DOB*

* *Agriculture was spelled incorrectly in* ***job\_industry\_category*** *for “****CustomerDemographic****”.*

*Recommendation: Need to replace word* ***Argiculture*** *to* ***Agriculture*** *using replace tool.*

***Completeness***

* *In some fields* ***(last\_name,gender*** *etc****)****and some other values under* ***customer\_id*** *were omitted.*
* *Blanks in* ***job\_title*** *for “****CustomerDemographic”,*** *in* ***inline\_order*** *and* ***brand\_column*** *for “****Transactions****”.*

***Consistency***

* *The gender specifications in the field “****gender”*** *from “****CustomerDemographic”*** *were not consistent.*

*Recommendation: better to provide dropdown menu in order to avoid human typing errors. Recommend to use F for Female and M for Male and U for Unknown.*

* *Representations of* ***state*** *in “****CustomerAddress “are*** *not consistent. For instance, NSW are labeled as New South Wales.*

*Recommendation: better to rename to short names .* ***NSW/VIC*** *can be mentioned instead of New South Wales and Victoria*

***Currency***

* *The people mentioned with* ***deceased\_indicator*** *as* ***Y*** *are not considered as active customer/current customer for* ***”******CustomerDemographic”*** *.*

*Recommendation: Recommended to remove all the customer details under* ***Y*** *category in* ***deceased\_indicator***

* *Incorrect date format in* ***product\_first\_sold\_date from “Transactions”.***

*Recommendation: better to provide correct date format*

* *Better to change current format to currency format (include dollar sign) in* ***list\_price from “Transactions”.***

***Relevancy***

* ***Default*** *column is not relevant for the company analysis from* ***”CustomerDemographic”****.*

*Recommendations****:*** *As the default columns and details are not useful and non-printable format it is better to delete the details.*

*Above mentioned details summarises the data quality issues identifies through the first stage of the data quality issues.*

*Hope it finds helpful for the second phase. Please let us know if you have any questions regarding the same*

*Kind Regards,*

***Ann Maria Joseph***

***Data Analyst***

***KPMG***