Title : Data quality issues

From : KPMG Analytics, Information modelling team

To: manager at Sprocket central pty Ltd

Hi, this is Alex smith

I am an Analyst at KPMG Analytics, Information modelling team

I have received your data and after looking and analysing the dataset there are some data quality issues. This issues are mainly focused on accuracy, completeness, consistency, timeliness, validity and uniqueness.

* The data is not accurate, as we look at gender table in the customer demographics there are spelling mistakes and incomplete gender name. This can be mitigate through correcting incomplete formats.
* Also the date consist of many incomplete record such as the job\_title and last name column have blank records, this helps us in knowing that the data is incomplete. We can reduces this through proper insights of that particular record.
* There are several records missing in the dataset considering the customer demographics table most of the data have missing records this leads to the inconsistency of data. This can be mitigate through finding and filtering records
* In the Date of birth column of the customer demographics there are some years of data is missing such as up to 1953 the data is consistent but after that it falls to 1944,1940 and even the year has gone to 1843 this data is not in proper timeliness thus this can be reduced by proper analysis of all the records.
* There some incorrect classification values for gender. This shows that the data is not valid properly. This can be reduced by recognizing the gender type
* Looking at the uniqueness of the every customer have unique id’s such the transition \_d in the transition table and the postal code in the address table considering all this records the data is unique.

I hope you have got the insights of the data quality if you have any queries please reach out by replying to this mail.

Thanks and regards

Alex smith