Hello John,

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. The table below highlights the findings after the data assessment, please let us know if you have any questions related to the findings presented.



Below we have provided detail description of findings and mitigation methods and suggestions,

1. Accuracy
   1. Date of birth data is not accurate there are some outliers
      1. *Mitigation: remove outliers*
      2. *Suggestions: create new feature age, age feature will help to easily understand errors*
2. Completeness
   1. There are some missing data
      1. *Mitigation: apply missing value treatments*
   2. customer\_id is incomplete among customer\_demographic and customer\_address
      1. *Mitigation: use only records which are available on both sheets and filter out all other incomplete data or update spreadsheets with correct data*
3. Consistency
   1. Inconsistency gender data in customer\_demographics
   2. Inconstancy state data in customer\_address
      1. *Mitigation: replace data to remove inconsistency*
4. Validity
   1. There are some records with product\_id = 0
      1. *Mitigation: check database whether there is 0 product\_id*