Dear Andrew,

I hope you are doing well. I am writing about the data quality assessment we recently completed on the data sent to us for analysis.

I would like to point out some issues with the data and make some recommendations on how to mitigate them.

The data quality assessment was made on the measures of Consistency, accuracy, completeness, uniqueness, Data Validity, etc.

**Transactions Data:**

1. Issues with the completeness (missing values) of the data.

2. The product first sold date is not a date but some numbers.  
3. List price and Standard cost do not serve any relevancy.   
4. Missing Quantity column  
5. Product class does not clarify what it refers to.  
  
**Customer Demographics:**  
  
1. Gender column has multiple values and spelling mistakes like "U", "F", "Femal", and "M".    
2. Check the year values in date of birth.   
3. The column "default" is not interpretable.   
4. Issues with the completeness (missing values) of the data.   
  
**Customer Addresses**  
1. There are customer ids that are not present in Customer demographics data.   
2. The format of state names is inconsistent. Either they should be all short forms or full names.  
3. Property valuation is not interpretable. Is it in millions or thousands?

Recommendations:

1. Be aware when inputting the date of birth values.

2. If you are using Excel, use filter options to make sure that the data is consistent.

3. Double-check or cross-check the values that are inserted into the cells

4. Make clear formats for each of the columns, you can create restrictions or rules that will allow the data to be inserted into a certain format only.

5. Highlight the empty cells.

If you have any questions regarding this, please do not hesitate to ask.   
  
Regards,

XYZ