

Dear [Client (or) KMPG Team],

I hope this email finds you well. I wanted to take a moment to update you on the data quality of the project. I have identified several issues, and the quality of the data can be significantly improved by a few changes.

Summary statistics of the three datasets provided by Sprocket Central Pty Ltd.

<u>Table name</u>	<u>No. of records</u>	<u>Distinct Customer IDs</u>	<u>Data Received</u>
Customer Demographic	13	4000	4000
Customer Address	6	3999	3999
Transaction	13	3494	20000

Some of the data quality issues encountered in the datasets are as follows:

- **Inconsistent Values:**

- **Eg:** () { _; } >_[\$(\$())] { touch /tmp/blns.shellshock2.fail; }
- ÿ¼ï½¥â¿ã¾â²(ï½¿ââ¿âï½¿)â±â¿ï½¥ï¾

In customer demographic, default attribute has been corrupted.

- **Blank columns**

- Eg: online order, brand, product size, job titles, ..

Some of the columns in the data sheet are blank

I believe that there is still room for improvement, and i will continue to work on enhancing the data quality.

If you have any questions or concerns, please feel free to contact. I would be glad to assist you.

Best regards,

KMPG virtual intern student

