January 21, 2022.

To whom concern,

I am Alma Cecilia Villarreal Duarte, intern of data quality of KPMG. I was looking to datasets you have send to us, and I would like to give some observations about the quality of them.

First, to work with the data and facilitate its manipulation with the programs, I recommend separating the sheets into 3 csv files, and removing notes, comments, deleting empty formatted cells and default columns.

Also, reviewing the ages of the clients, I found that there were people with more than 100 years old, which represents a problem in the reliability of the data, so I would suggest eliminating those records.

Another point to mention is the data entered must also be standardized. For example, in the gender variable, the user should select only F for Female and M for Male, and in case he/she wants to enter other data, mark error. If working with these data, the values should be modified and, in the cases, where they are empty or do not correspond, the empty data should be filled in.

This also applies to cities; you must choose whether to use acronyms or the full name. In the case of working with these bases, the full names should be replaced by their respective acronyms.

In addition, to compare zip codes, he would recommend changing the data to numeric type.

Likewise, for the missing data in the customer demographic sheet, he considered that it will have to be eliminated because it does not present sufficient information. However, for the transaction sheet, the missing data can be filled in. Similarly, for the job industry column of the customer demographic sheet, you can change the records with n/a to empty cells and then fill the data with the information obtained.

Finally, I would like to point out that there are more addresses than customers with demographic data, so I consider that they should have the same number of records and should match the customer id.

For the moment this is all from my side, if you have any questions, please send me an email.

Sincerely,

Alma Cecilia Villarreal Duarte.