Hi, hope this mail finds you well. From the dataset presented, I have identified the following issues. In the CustomerAddress data, the attribute/column property\_valuation we need clarification on what it stands for or what it means in terms of the business.

While on the CustomerDemographic dataset, the column "default" contains invalid values and should be dropped/deleted because it has no effect on the whole dataset. The job\_title and job\_industry\_category has some missing values, this is about 656 rows with missing job title and industry category which holds only 16.4% of the total dataset. Before we could carry out analysis, the rows should be dropped. And the DOB column has few missing values; I suggest that, customers with missing a missing age and also has missing job title or job industry category should be dropped while we populate others with the average age of customers in the system. Also there are some consistency issues in the gender column, where some customer’s gender is represented by Female or Male then some are represented by F or M.

And the Transaction data, about 197 transaction has no record, i.e missing online\_order,brand, product\_line,product\_class,and product\_size values. This could be big problem as $213,695.05 worth of products was sold with no record.