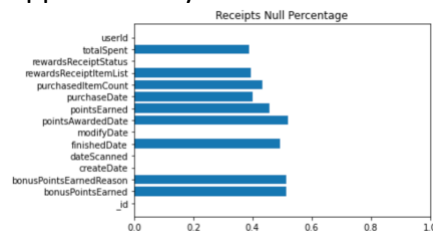


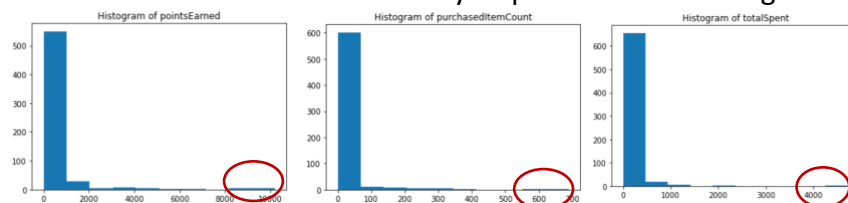
Hi everyone,

This is Jenny. I wanted to provide some updates and raise a few questions regarding the Exploratory Data Analysis project I have been working on recently. Please find the details below:

1. I kindly request a clearer definition of Receipts Items List. This will enable me to structure an optimal Receipts Items Data Schema.
2. During this project, I have identified several data quality issues that need attentions:
 - a. Incorrect Data Type: Currently, all date columns are assigned the data type "object" instead of the appropriate datetime formats. I recommend reviewing our database to ensure that all date fields are stored consistently.
 - b. High Percentage of Missing Data: In the Receipts Data I have noticed that approximately half of the data is missing in 9 out of 14 columns.



- c. Presence of Multiple outliers: Some receipts show customers earning 4000+ points, purchasing 600+ items, or spending 4000+ dollars. I would like to confirm if these values are correct or if they require further investigation.



- d. Duplicate Records: The Users Data contains a significant number of duplicate records (57%). I suggest removing redundant entries and ensuring that all primary keys are unique.
3. Can we schedule a meeting for a better understanding of your goals and business needs? This will allow us to align our efforts more effectively and ensure that our analysis and insights align with your expectations. Additionally, your domain knowledge would greatly assist us in addressing data quality concerns and refining the analysis further.

Thank you for your attention to these matters. If you have any questions on any of the points mentioned above, please do not hesitate to reach out. I look forward to discussing them further.

Best regards,
Chieh-Yin (Jenny)