Consider the work you just completed in the Data Engineer Exercise. Please write no more than 2 paragraphs explaining to a client how you approached this task. Remember that clients do not need to know every process detail, but do want to understand how and why your choices contribute to our overall strategy and any benefits of the end product.

Assumption: I understand this task as I would explain to a client why I came up with two data sets: people and acquisition fact, and how it could be helpful to the client, instead of explaining how I processed the data the way I did because they are explained in the codes.

Although The Constituent Information dataset includes a lot of interesting attributes about the constituents such as gender, income, birthday, most of the values are encrypted and meaningless. The only information that might be helpful is source, which is the source of data such as Google, Twitter, Facebook, etc. Similarly, from the email data set, we can get the email, created date, updated date, and primary email. From the subscription dataset, we can get the subscription status. We can link these information from the three dataset based on shared attributes ‘cons\_id’ and ‘cons\_email\_id’. From the people dataset, another piece of information that might be helpful is the number of constituents acquired on acquisition date. So we count the number of acquisition on each date in the dataset to come up with the acquisition fact dataset.