Hi Data Team,

I have finished reviewing the files that comprise the data schema for Fetch Rewards. I was able to construct a workable ERD/schema, but upon conducting an exploratory analysis in R, there are a few points where I could use some guidance from the team.

The bottleneck of my process happened to be the nested json structure within the receipts data file. It is necessary for the company to log each of the items that exist on a given receipt ID and right now they are all stored in the "rewardsReceiptItemList". In other words, the structure follows as a table (receipt) full of many other tables (the contents of the receipt). When observing the item information that is listed on the receipts, I noticed an inconsistency with the structure.

It is apparent that each receipt logs a different set of data pertaining to items. For example, the item table of one receipt may store 18 variables, while another may have only 7. Furthermore, there are instances where those 7 variables are tracking completely different data than the 18 variables on the other receipt. This may cause a problem with analysis because if the type of data being tracked among instances is different, then it is impossible to compare them to one another. Robust analysis relies on a normalized table structure in which each row contains the same number of columns that track the same variables. From there, valuable trends and comparisons can be detected.

Considering this, I feel that it would be very beneficial for me to grasp an understanding of the data collection process. Specifically, how does the company obtain the data that I have reviewed? Are there certain portions of the process that are manually entered by Fetch employees or is all the information pulled in by software that automates the collection? Insight into the operations of the business will allow me to identify where changes can be made to ensure that Fetch rewards stores data in an orderly, normalized manner.

Please let me know if you have any questions for me- I look forward to hearing back from you!

Thanks,
Keenan Daly