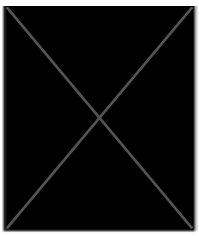


[redacted]*Data Analyst Assessment*

Attached is sample data from survey responses for users who visited grocery retail stores between 08/01/2021 and 10/01/2021.

- Users with multiple store visits are allowed to give the survey multiple times.
- Each question in the survey is associated with the location asked as the first question "which location did you make your purchase today".
- Question code q2 is asked multiple times in the survey to the same user based on the product category from which the user made a purchase.

Below is the structure of the Data

SURVEY_ID Unique Id assigned to the Survey

USER_ID Unique Id assigned to the User

QUESTION_CODE Code associated with the question

QUESTION Question Text

RESPONSE Response to questions by User

PRODUCT_CATEGORY Product associated with question code q2

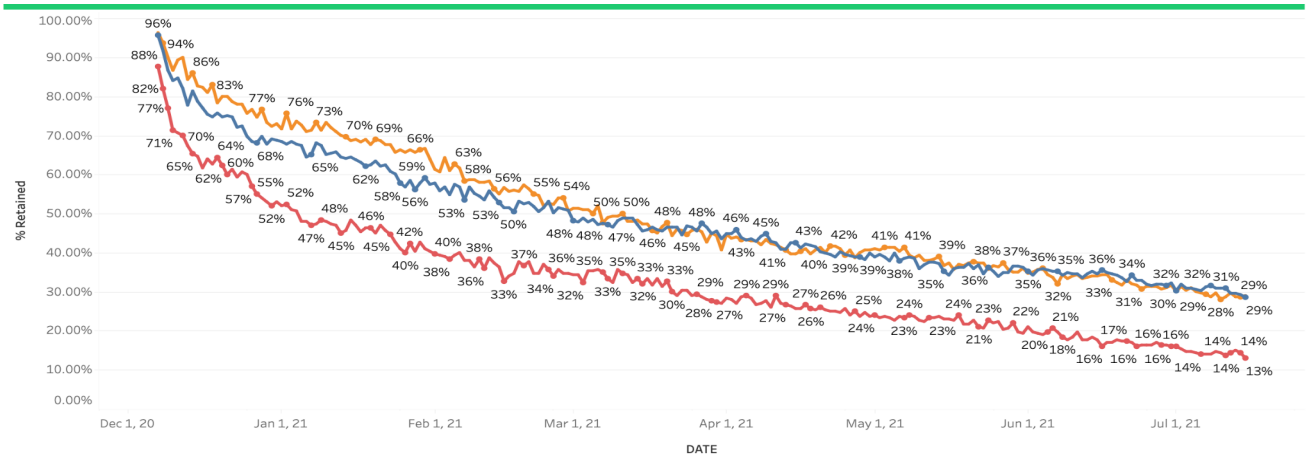
SURVEY_RESPONDED_DATE Date when the user responded to the survey

Assessment: Please find answers to the below questions using SQL queries and visualizations. Use any available tools to share visualizations such as Excel, PPT, or BI tool - Preferred. Please share the process if you are able to find the answers using any alternative ways.

1. Do you observe any anomalies in the data?
2. How many distinct locations/retail stores are part of the survey.
3. Transform the data using SQL query to be consumed by BI tool based on below requirement. Share SQL Query & Visualization. *(Tip: you can use transformed data to answer questions 4 to 8)*
    a. Visualize trip type ratio by location.
    b. Use time_of_the_day and buyer/non_buyers as filters for the visualization. The user is termed as a buyer if the response to q3='i was able to buy everything on my list' else non-buyer.

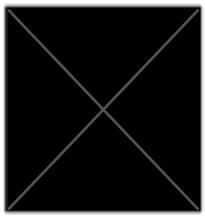
[redacted] *Data Analyst Assessment*

4. Assign a sequence/order using the window function to all users in which they gave surveys and list all users who gave more than 2 surveys(Share SQL query).
5. Visualize q2 with the goal of showing important factors in the purchase decision of users by location for buyers vs non-buyers.
6. Calculate & Visualize recommendation_score for each location mentioned in the data using responses of q4 for buyers. Sort the visualization by top score.
    a. Recommendation score: (Yes Responses- No Responses)*100
7. What is the most important factor for purchasing among buyers from Walmart?
8. Which two locations/brands are the most preferred location among buyers who purchased "Coffee/Tea" and what is the most important factor driving the purchase?
9. Which is the most preferred location for "Grab & Go" among all users.
10. As a part of your project on user engagement predictions, you need to find a correlation between 2 columns (number of views & subscription date). Which plot is the best to explore this scenario?
11. Below is the 7-month retention chart of the user categories shared in the table below, what are your recommendations to the growth director who is looking to retain more users with a limited budget.



- Users who were not paid anything
- Users who received $1 incentive every week for sharing their opinion via surveys.
- Users who received $0.5 incentive every week for sharing their opinion via surveys.

[redacted]



12. Mention at least 2 ways of time management when you are assigned 3 below tasks that are expected to be finished in a week. All three have equal strategic importance.

Task a: Straightforward but will take approx 2-3 days to finish according to your plan.

Task b: Straightforward & needs a quick query update.

Task c: Problem-solving task with multiple unknowns, unable to predict the time that will be required to finish the task.