Fetch - Data Analyst Take Home

In this exercise you will:

Demonstrate how you reason about data and how you communicate your understanding of a specific data set to others.

Requirements:

The exercise must be committed to a git repository and the link must be provided to the recruiter. Emailed files will not be reviewed.

Tips:

- Communicate your thoughts in each section. Code should be commented. Findings and conclusions should be written down and shared.
- Share assumptions you are making.

Exercise:

First: explore the data

Review the unstructured csv files and answer the following questions with code that supports your conclusions:

- Are there any data quality issues present?
- Are there any fields that are challenging to understand?

We recommend using SQL or python and data visualization to examine the data.

Second: provide SQL queries

Answer three of the following questions with at least one question coming from the closed-ended and one from the open-ended question set. Each question should be answered using one query.

Closed-ended questions:

- What are the top 5 brands by receipts scanned among users 21 and over?
- What are the top 5 brands by sales among users that have had their account for at least six months?
- What is the percentage of sales in the Health & Wellness category by generation?

Open-ended questions: for these, make assumptions and clearly state them when answering the question.

- Who are Fetch's power users?
- Which is the leading brand in the Dips & Salsa category?
- At what percent has Fetch grown year over year?
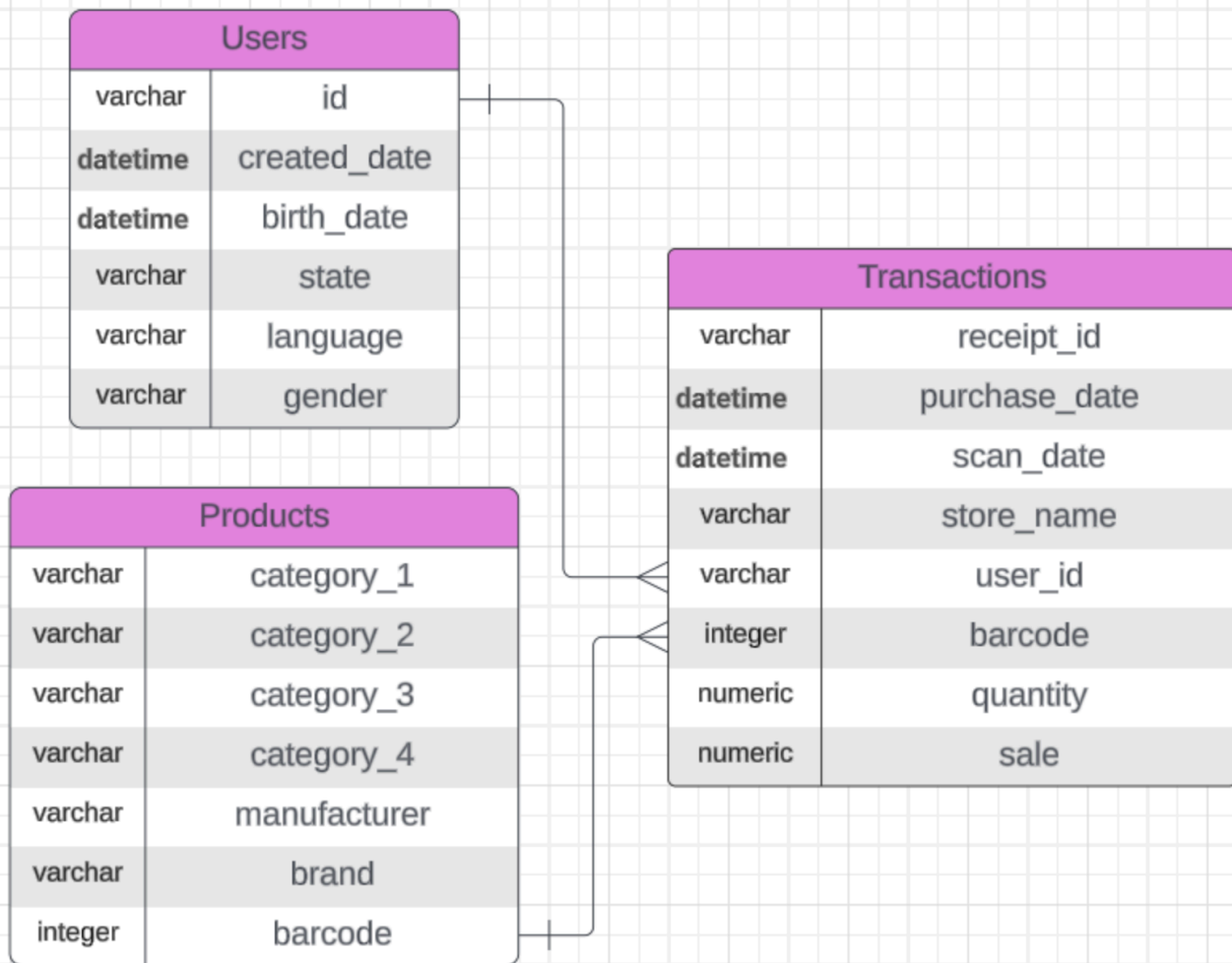
Third: communicate with stakeholders

Construct an email or slack message that is understandable to a product or business leader who is not familiar with your day-to-day work. Summarize the results of your investigation. Include:

- Key data quality issues and outstanding questions about the data
- One interesting trend in the data
  - Use a finding from part 2 or come up with a new insight
- Request for action: explain what additional help, info, etc. you need to make sense of the data and resolve any outstanding issues

Data

- Users: DOWNLOAD CSV
- Transactions: DOWNLOAD CSV
- Products: DOWNLOAD CSV

Entity Relationship Model

FAQs

How do I submit my exercise?

Provide a link to a public repository (i.e., GitHub, Bitbucket) to your recruiter.

How will this exercise be evaluated?

A team member will review the code and documentation you submit and evaluate based on your ability to make sense of the data, write SQL, provide a reasonable visualization, and explain your findings.

I have questions about the problem statement.

For any requirements not specified above, use your best judgement to determine the expected result. You are encouraged to explain your assumptions.