Using the programming language of your choice (SQL, Python, R, Bash, etc...) identify as many data quality issues as you can. We are not expecting a full blown review of all the data provided, but instead want to know how you explore and evaluate data of questionable provenance.

1. **Identify Outlier Transaction Amounts:**

SELECT \*

FROM receipts

WHERE totalSpent = 0 OR totalSpent > 10000;

1. **Locate Invalid Dates**

SELECT \*

FROM receipts

WHERE purchaseDate < '2000-01-01' OR purchaseDate > CURRENT\_DATE;

SELECT \*

FROM users

WHERE createdDate < '2000-01-01' OR createdDate > CURRENT\_DATE;

1. **Locate instances where data is missing:**

SELECT

COUNT(\*) AS total\_rows,

SUM(CASE WHEN brand\_name IS NULL THEN 1 ELSE 0 END) AS missing\_brand\_name,

SUM(CASE WHEN \_id IS NULL THEN 1 ELSE 0 END) AS missing\_brand\_id

FROM brands;

SELECT

COUNT(\*) AS total\_rows,

SUM(CASE WHEN user\_id IS NULL THEN 1 ELSE 0 END) AS missing\_user\_id,

SUM(CASE WHEN brand\_id IS NULL THEN 1 ELSE 0 END) AS missing\_brand\_id,

SUM(CASE WHEN totalSpent IS NULL THEN 1 ELSE 0 END) AS missing\_totalSpent

FROM receipts;

SELECT

COUNT(\*) AS total\_rows,

SUM(CASE WHEN \_id IS NULL THEN 1 ELSE 0 END) AS missing\_user\_id,

SUM(CASE WHEN createdDate IS NULL THEN 1 ELSE 0 END) AS missing\_createdDate

FROM users;