Difficulties and Interesting Observations:

Difficulties: The main difficulty is handling missing values or incorrect data entries, especially if the dataset includes non-standardized formats or typos.

Interesting Observation: One interesting observation is the distribution of time spent on different platforms across various demographics, revealing patterns such as younger users spending more time on certain platforms compared to older users.

2. Data Fun

Cool Facts Discovery:

To uncover interesting facts, we can use SQL queries to analyze the data. Here are two examples:

**Most Popular Interests:**

SELECT interest, COUNT(\*) AS frequency

FROM social\_media\_users

GROUP BY interest

ORDER BY frequency DESC

LIMIT 5;

This query will show us the top 5 interests among social media users.

**Average Income Based on Profession:**

SELECT profession, AVG(income) AS average\_income

FROM social\_media\_users

WHERE income IS NOT NULL

GROUP BY profession;

This query calculates the average income for each profession, providing insight into salary disparities across different job titles.

Formulated Questions:

**Ask Away**

**What is the distribution of home ownership among urban vs. suburban users? To answer this, we can use a query like:**

SELECT location, COUNT(\*) AS total\_users,

SUM(CASE WHEN isHomeOwner = TRUE THEN 1 ELSE 0 END) AS homeowners

FROM social\_media\_users

WHERE location IN ('Urban', 'Sub\_Urban')

GROUP BY location;

This will give us the count of users and the number of homeowners in both urban and suburban locations.

**How does the platform preference vary between male and female users? We can explore this with:**

SELECT gender, platform, COUNT(\*) AS user\_count

FROM social\_media\_users

GROUP BY gender, platform

ORDER BY user\_count DESC;

This query helps identify which platforms are more popular among males and females.