DATA 514 Section 1 Worksheet 1

Use the same command line interface session for all three Lecture's questions.

Lecture 1

Question 1

1. For windows, open a command line interface (CLI), such as command prompt, and navigate to a folder using the command cd (change directory).
    - Navigate to "C:\Users[insert your username]\AppData\Local\Google\Chrome\User Data\Profile [fill in a number]".
    - To find the profile number to look at you can navigate to the same file directory using File Explorer.
    - `cd path` where path is the one listed above with your information filled in.

1. For mac, open a command line interface (CLI), such as terminal, and navigate to a folder using the command cd (change directory).
    - Navigate to "~/Library/Application Support/Google/Chrome/Profile [fill in a number]/".
    - To find the profile number to look at you can navigate to the same file directory using Finder.
    - `cd path` where path is the one listed above with your information filled in.

2. Make sure all Chrome instances with this profile are closed out, with no windows or tabs open for your selected profile. You cannot access the database while the profile is in use.
3. Once you have navigated to the folder for your selected profile, run the command `sqlite3 History`.
4. Run the following sql query `select * from urls limit 3;`. You will use the output to answer questions 1-3.

What is your output?

Question 2

In the space below provide an example of an element, also known as a "row".

Question 3

In the space below provide an example of an attribute, also known as a "field".

Question 4

Run the following sql query `.schema urls`.
What information does this provide you about the table called urls? Does it have to do with data type or domain or both? Why?

Question 5

Use both of the previous outputs to answer this question. Which attributes are a key? Why are keys crucial to identify in SQL?

Lecture 2

Question 1

Run the following sql query `.tables`.
Construct a for-each loop to aquire a unique output using one of the tables, a different one from the example below.

```
foreach row in urls:
    if row.visit_count > 10:
        output(row.id, row.title)
```

Question 2

Using the for-each loop you wrote above, what is the equivalent sql query?

Question 3

How many websites have you visited more than 10 times? 100 times? 1000 times? Use the `COUNT` command in your select statement.

```
SELECT COUNT(visit_count)
FROM urls
WHERE visit_count > 10;
```

Question 4

Choose two tables that have at least one overlapping attribute (IE visit_count in urls and visit_count in segment_usage). Which attribute from which table is the key? Which attribute from which table is the foregin reference?

Lecture 3

Question 1

In Lecture 2, Question 4, what type of join would combining these tables create? Why?

Question 2

Execute this join to discover something about your search history. Be sure to include some form of aggregration by using a WHERE clause, a COUNT clause, or a AND/OR statement. Please provide your sql query below.

What did you learn about your search history?