

Interaction.py

- Basic Structure:
 - Import functions from SQL_command.py
 - To-do: Import functions from mdb_command.py
 - Prompt: choose dataset or upload dataset
 - Response: overview of chosen dataset
 - Prompt: What query are you interested in? (placeholder b/c not doing NLP yet)
 - Response
 - SQL Query string example
 - Result of executing SQL query string
 - To-do: mdb query string example
 - To-do: result of executing mbd query string
 - Prompt: Anything else?
 - Loop until user says "No"

SQL_command.py

- Setup functions:
 - Connect
 - Upload
 - Overview
- SQL clause functions:
 - Select
 - Aggregate
 - From
 - Where
 - Group By
 - Order By
 - To-do: having, limit, join, update, delete
- Complete query functions:
 - Suggest query
 - Main query generator
 - Puts together SQL clauses based on different arguments
 - To-do: account for string values, values that contain "" single quote character, and columns with spaces
 - Execute query
 - Executes query string
 - Random query (placeholder b/c not doing NLP yet)
 - Currently generates random queries (as if no user input) that include select, from, where, and order by
 - To-do: implement groupby, aggregation, having, limit, join, update, delete

Checkpoint Goal

- Files: interaction.py, mdb_command.py, and SQL_command.py
 - Basic interaction
 - Example queries without NLP

- Writeup
 - “tell us your progress so far and the challenges you might have encountered.”

Questions/Clarifications

1. Plan for integrating NLP:
 - a. Use NLP to analyze user input to determine which arguments to pass through “suggest query” function from SQL_command.py
2. Deal with string values that contain single quote character (‘’) and column names that have spaces:
 - a. Remove all single quotes
 - b. Rename columns that have spaces by using an underscore
3. Implementation of Update and Delete queries:
 - a. Update and Delete will modify dataset
 - b. Reload datasets at the start of program
4. Interaction structure:
 - a. Is the current structure good? (it is slightly different from diagram in project proposal)
 - i. Changes:
 1. Choose dataset
 2. Ask SQL or NoSQL (if not specified, give both)
 3. Ask what kind of query
 4. Generate random query