

Lab 2 Part 2

- Gary Miller

Querying the Database

- Write a query that will return the count of elements in the Entry columns of the Alzh table.

In order to return the number of elements in the Entry column for the table Alzh, the following query was entered in the terminal:

```
SELECT COUNT(id) FROM Alzh;
```

The query output that 3,384 elements exist in the id column in the table Alzh.

- Write a query that will return the distinct count elements in the Entry column of the Alzh table.

In order to find the distinct number of elements in the entry column the previous query was updated and entered. This new query that was entered looked like:

```
SELECT COUNT(DISTINCT id) FROM Alzh;
```

The query output that 3,384 distinct elements exist in the id column in the table Alzh.

- Discuss: From the above two queries, is this column a good primary key for the Alzh table? why or why not? (if not, then what column would you recommend, instead?)

This column is a good column to be used as the primary key, because there are no duplicate entries in the column and any more entries that would be entered in the future will also be unique. Further more, all entries in the database must have a value for the id column.

- Write a query that will return the number of records associated with the organism “Zea mays (Maize)” in the Alz and Park tables. The following query was entered in order to find the number of elements associated with the organism “Zea mays (Maize)”:

```
SELECT COUNT(organism) FROM Alzh WHERE organism == "Zea mays (Maize)";
```

The output of the query showed that 0 elements are associated with the organism Zea mays (Maize) in the table Alzh.

- Write a query that will report how many organisms were listed in each of the three tables. The following queries were entered to determine how many organisms were listed in the tables Apop, Park, and Alzh:

```
SELECT COUNT(Apop.organism) FROM Apop;
SELECT COUNT(Park.organism) FROM Park;
SELECT COUNT(Alzh.organism) FROM Alzh;
```

the output these queries showed that there were 101,912, 126,535, and 3,384 entries in Apop, Park and Alzh respectively.

- Write a query that will return the number of organisms which are common to both the Parkinson's and Alzheimer's tables (i.e., This is the intersection of proteins of both tables). The following query was entered in order to find the number of organisms common to both the Parkinson's and Alzheimer's tables:

```
SELECT COUNT(Park.organism), COUNT(Alzh.organism) FROM Park, Alzh WHERE Park.organism =
```

The output of the query showed that there were 79,304 organisms common to the Park and Alzh tables.

- Write a query that will return the number of proteins which are common to Apoptosis and Parkinson's, which are associated to the Bothrops brazili organism. The following query was entered to find the number of protein names that were associated with the organism Bothrops brazili and were common to both the Apop and Park tables:

```
SELECT COUNT(Park.protein_name), COUNT(Apop.protein_name) FROM Park, Apop WHERE Park.or
```

The output of this query showed that there are no common proteins that are associated with this organism in the Apop and Park tables.

- Create a query to determine how many genes are in common in the Alz and Park tables, for all organisms (i.e., the intersection of all information about genes across all organisms). The following query was entered to find the number of genes that were common to the Alzh and Park tables for all organisms:

```
SELECT COUNT(Alzh.gene_name), COUNT(Park.gene_name) FROM Alzh, Park WHERE Alzh.gene_nam
```

The output of this query showed that 1,015,664 genes were common to the tables Alzh and Park.

- Create another query to determine the names of first ten of these genes which are at the intersection of Alz and Apop tables, across all organisms. The following query was entered to find the names of the first ten genes that were common to both the Alzh and Apop table for all organisms:

```
SELECT Alzh.gene_name, Apop.gene_name FROM Alzh, Apop WHERE Alzh.gene_name == Apop.gene
```

The output of the query was:

```
Gene names  Gene names
APP A4 AD1  APP A4 AD1
App App
App App
```

App App
App App
PSEN1 AD3 PS1 PSNL1 PSEN1 AD3 PS1 PSNL1
MAPT MAPTL MTBT1 TAU MAPT MAPTL MTBT1 TAU
APP APP
APP APP