



DATA 514 Section 2 Worksheet

Lecture 4

Pework

1. From this website,
2. Open a command line interface (CLI) and using the change directory command, navigate to the file directory where you stored the database. For example `cd "~/Desktop/Data Science/data"`. Notice the double quotes surrounding the entire path term. Use these when you have spaces. Alternatively you can run `cd ~/Desktop/'Data Science'/data` to get the same output. Note the single quotes around the term 'Data Science'. On windows the commands would be `cd "C:\Users\[insert user]\Desktop\Data Science\data"` or `cd C:\Users\[insert user]\Desktop\'Data Science\'data`.
3. In the CLI, in the folder containing the csv file with extension .csv, run the command `sqlite3 employees.db` to open an empty database to load the csv file into.
4. Run the following commands:

```
CREATE TABLE salaries (id int PRIMARY KEY, first_name varchar(30),  
last_name varchar(30), salary int);  
CREATE TABLE titles (id int PRIMARY KEY, first_name varchar(30),  
last_name varchar(30), title varchar(30));  
.mode csv  
.import [file name] salaries  
.import [file name] titles
```

Question 1

What are the two most common titles?

Question 2

What is the average pay for software engineers?

Question 3

How many employees in each role have a salary less than \$7000 for the job title average?

Lecture 5

Question 1 (5 Points)

Given tables created with these commands:

```
CREATE TABLE A (a int);
```

```
CREATE TABLE B (b int);
```

```
INSERT INTO A VALUES (1), (2), (3), (4);
```

```
INSERT INTO B VALUES (3), (4), (5), (6);
```

Part 1 (1 point)

What's the output for the following:

```
SELECT *  
  FROM A INNER JOIN B  
    ON A.a=B.b;
```

Part 2 (1 point)

What's the output for the following:

```
SELECT *  
  FROM A RIGHT OUTER JOIN B  
    ON A.a=B.b;
```

Part 3 (1 point)

What's the output for the following:

```
SELECT *  
  FROM A LEFT OUTER JOIN B  
    ON A.a=B.b;
```

Part 4 (1 point)

What's the output for the following:

```
SELECT *  
  FROM A FULL OUTER JOIN B  
    ON A.a=B.b;
```

Part 5 (1 point)

What type of join does the following code produce?

```
SELECT * FROM A INNER JOIN B;
```

Question 2 (3 points)

Consider the following over simplified Employee table:

```
CREATE TABLE Employees (id int, bossId int);
```

Suppose all employees have an id which is not null. How would we find all distinct pairs of employees with the same boss? Provide your sql query below.

Question 3 (9 points)

For the following tables:

```
CREATE TABLE Movies (id int PRIMARY KEY, name varchar(30),  
budget int, gross int, rating int, year int);  
CREATE TABLE Actors (id int PRIMARY KEY, name varchar(30),  
age int);  
CREATE TABLE ActsIn (mid int REFERENCES Movies(id), aid int  
REFERENCES Actors(id));
```

Part 1 (3 points)

What is the number of movies, and the average rating of all movies that the actor "Patrick Stewart" has appeared in?

Part 2 (3 points)

What is the minimum age of an actor who has appeared in a movie where the gross of the movie has been over \$1,000,000,000?

Part 3 (3 points)

What is the name and budget of each movie released in 2017 whose oldest actor is less than 30?