CS 121: Relational Databases — Winter 2026

Project 00: SQL Setup

Installing SQLite

- **MacOS**: SQLite comes pre-installed on recent versions of MacOS.
- **Windows**: Visit the SQLite Download Page. Find the right pre-compiled binary depending on your system. Extract the contents to a folder and add to system PATH.
- **Linux**: Follow instructions to install `sqlite3` for your distribution's package manager.

Accessing the Database

Download and extract the database from the course website and add it to the repository you cloned from Gitlab.

SQLite provides a command-line shell.

Run SQLite and view the database `lahman.db` using the command line.

```
sqlite3 -header lahman.db
```

Once inside, run the following commands within your project to familiarize yourself with the environment.

```
# Inspect the 'people' schema
.schema people

# Additional SQL Queries
SELECT playerid, namefirst, namelast FROM people;
SELECT COUNT(*) FROM fielding;
```

Schema

The database is comprised of the following main tables:

- **People**: player names, DOB, and biographical info
- **Batting**: batting statistics
- **Pitching**: pitching statistics
- **Fielding**: fielding statistics

It is supplemented by auxiliary tables such as `AllStarFull`, `HallofFame`, `Managers`, `Teams`, `Salaries`, and `Schools`. For more detailed information, refer to the documentation linked on the course website.

Instructions

The project repository contains a separate `.sql` file for each question below (e.g., `q1i.sql`, `q2ii.sql`). Write the SQL query for each problem in its corresponding file.

Each file also additionally contains a comment displaying the **first few lines** of the expected output as well as the **total number of expected lines (including the header row)** which you can use to compare against your output. *Note: depending on how you write your query, your header line may look different.*

Example

Consider the following question: **What is the highest ERA (earned run average) recorded in baseball history?**

The file `q0.sql` already contains a correct SQL query that answers this question. To validate the answer, run the following command in your terminal:

```
sqlite3 -header lahman.db < q0.sql
```

This command feeds the contents of `q0.sql` into SQLite and executes it against `lahman.db`. The `-header` flag ensures that column names are displayed in the output. **You should use this format when running your own queries later on.**

You will see the output:

```
max(era)
189.0
```

This matches the expected output and expected line count indicated by the comment in `q0.sql`:

```
/*
max(era)
189.0

Expected lines: 2
*/
```