

# HW SQL

David Gerard

2023-11-13

©David Gerard (2023), not to be posted online.

## Instructions

- Write your solutions in this starter file. You should modify the “author” field in the YAML header.
- Do not modify the paths of any files.
- Only commit R Markdown and HTML files (no PDF or Word files). Make sure you have knitted to HTML for your final submission.
- **Make sure to commit each time you answer a question.** Lack of informative and frequent commits will result in point deductions.
- Only include the necessary code, not any extraneous code, to answer the questions.
- Learning objectives:
  - Basic SQL commands.

## (9 pts) Lahman Data

Consider the Lahman data on baseball statistics that you can read about here: <http://seanlahman.com/download-baseball-database/>

I have placed a DuckDB database in the “data” folder, along with a data dictionary. Use this data dictionary to understand the variables and tables in this database.

Below, answer all questions using SQL code chunks (except connecting/disconnecting to the database).

1. (0.5 pts) Create a connection to the Lahman data in the “data” folder.
2. (1 pts) What tables are in this database?
3. (1 pt) These data end at the end of the 2021 season. Give me a table with all the folks who played in the 2021 season (which started on Apr 1, 2021), but only include their player ID, and their first and last names
4. (2 pts) Display the batting averages (defined as hits divided by at-bats, or  $H/AB$ ), and the associated playerID’s, for the 1998 season in Descending order. Only include individuals with at least 50 at-bats.
5. (2 pts) What’s the batting average for Barry Bonds in 1998? Use just SQL to get Bonds’ playerID and then use that ID to get his AB in 1998. This should only be one SQL call (and you won’t get credit if you manually type Bonds’ playerID).
6. (1 pt) Give me the average batting average by year (total hits divided by total at bats).
7. (1 pt) Write the result of the previous query to the output folder as “ba.csv”.
8. (0.5 pts) Close your connection.