This homework uses SQL to answer queries about the Lahman baseball database. If you have not obtained the SQLite Lahman database you must do to prior to answering these questions.

Each question will have too boxes for your answers. After you have a working query copy the query from the SQLite query manager and paste into the box label Query. Execute the query and paste your answer in the Results box. I will give an example of what your output should look like below. Use a snipping tool to copy the query results.

**Example: Query**

Use the Layman salary table to determine the first year that salary data was available for MLB (Major League Baseball) in the Lahman database.

**Query**

|  |
| --- |
| SELECT min( yearID )  FROM Salaries |

**Query Result**

|  |
| --- |
|  |

**Homework Questions**

1. How many teams were there in the year 2014 in major league baseball?
2. Use the **GROUP BY** statement to count the number of Teams in the American and National Leagues in the year 1960?
3. List all the players who played for the Texas Rangers who were not Pitchers during the 2010 season. Use the Batting Table and the Master table to answer this question. You only need to show the batters that show up on the first page.
4. Repeat Question 3 including using the Pitching table instead of the Batting table.
5. Go to the SQLite documentation and look up the UNION statement. Use the UNION statement to get all the players.
6. Look up the SQLite syntax for creating a view. Create a view named Rangers2010

That will contain just the first and last name of the Texas Rangers players for the year 2010.

1. Use the view you just created to produce a list of the Rangers who played in 2010 sorted in order by last name, and then first name. Print the names in the same order.
2. Look up how to concatenate using SQLite. We want the same query results as in Question 7 but we want the output to be concatenated in last name, comma, space, first name order. The column output should be called “Player.” An example for the first few lines of output is shown below.



1. Count the number of years that Babe Ruth played major league baseball. There is only one player in the history of baseball with the last name of Ruth.
2. Compute the total salaries of every major league baseball team in the year 2014. List the team name and the total salaries.
3. Compute the percent of runs scored for the Texas Rangers player Elvin Andres for the 2010 baseball season compared to the total runs scored by the team. Hint you will need to three select statements to do this in a single select your query should have a structure that looks something like this:

SELECT ( SELECT ….. FROM … WHERE … ) / ( SELECT …. FROM …WHERE …. )

1. Consider the Batting statistics tables. Some of the modern statistics have only been kept recently. Amount these statistics are **IBB** (Intentional walks) and HBP (Hit by Pitch). Write a query that will give the first year that both statistics were kept. You will need this year in the next query.
2. Create a view call **BPerformance** from the Batting table for all years equal to or greater than the year you discovered in Question 12. In your view the original Batting columns and include the following new columns.

|  |  |
| --- | --- |
| Column name | Equation |
| BA | **BA = HITS / AB** |
| OBP | **OBP = ( H + BB + IBB +HBP ) / (AB + + BB + IBB +HBP)** |
| SLG | **SLG = ( H + `2B` + 2 \* `3B` + 3 \* HR ) / AB**  Note: that SQL variables that start with numbers must be enclosed in backward quotes. |

Note that you will need to cast (change) one expression to a floating point number to get decimal answers for **BA, OBP**, and **SLG**. We do this with the **SQLite** CAST function. For example, the expression to compute the batting will look like:

**CAST( H AS FLOAT) / ( CAST( AB AS FLOAT) ) AS BA,**

1. Use the view you just created to list the Texas Rangers players for the 2014 season. Return the players last name, **BA, OPB**, **SLG**, and the sum of **OPB** and **SLG** as a variable called **OPS.** Order the players by **OPS** in descending order**.** Only list the players with at least 100 bats.
2. Use a nested query to determine the name of player or players who have the highest salary in MLB for the year 2014. List the first and last names.
3. Use SQL to determine which of those players plaid for the Texas Rangers?
4. Compute the total salary for each MLB team for the 2010 season. Order your list in descending order by team salary. Use the **teamID** to identify the teams.
5. UTA is not in the **Schools** table. Add it to the table. Confirm you add with a suitable query.
6. Hunter Pence played for UTA. Find his **playerID** from the Master table.
7. Hunter Pence played for UTA during the 2003 season. Update the **CollegePlaying** table to reflect this fact.