# Chapter 3 SQL Assignment – Part 2

The following are the queries that are to be written using the BaseBall database you created for the course. Each question is worth 3 points. Points will be taken off for incorrect formatting (dollar amounts should be in $xxx,xxx.00 format, percentages in XX.XX%, etc…). Each of the questions states what the query should do and provides a limited sample of the results set you should get from your query. Note that due to differences in the databases, you result numbers may not exactly match the examples.

The questions follow the Chapter 3 PowerPoint in terms of the SQL commands used. The PowerPoint can be used as a guide. For some of the formatting or other requirements you may need to use Google for assistance. Using **TSQL** (this is the nickname for Microsoft’s SQL formats) and the words for what you are trying to do often works. Adding the word **EXAMPLE** will find solutions that include examples of the SQL. Results from **StackOverflow** and **Microsoft** often give the best information.

1. Using the Salaries table and the appropriate aggregate functions, calculate the average salary by lgid, teamid and team name sorted by lgid and teamid for 2015

**lgid teamid name Average\_Salary**

AL ARI Arizona Diamondbacks $4,875,000.00

AL BAL Baltimore Orioles $4,108,744.04

AL BOS Boston Red Sox $5,659,481.25

1. Using the Master and Salaries tables along with the appropriate aggregate functions list the playerid, names (in nameGiven (nameFirst) nameLast format) and average salary for all players who had a salary greater than $400,000. You must use a HAVING clause to handle the salary condition and your answer must be sorted by playerid.

**playerid Name Average\_Salary**

aardsda01 David Allan (David) Aardsma $1,322,821.43

aasedo01 Donald William (Don) Aase $575,000.00

abadfe01 Fernando Antonio (Fernando) Abad $753,280.00

1. Use a nested subquery and an IN statement in the WHERE Clause to find all players who played in 2010 and 2015. Your results should show the playerid, full name, teamid and team name. **HINT:** The subquery should use the Appearances table to find out who played in any given year. Use that query with an IN clause in the WHERE clause for your query. The main query will use the Master, Appearances and Teams tables.

**playerid Player\_Name teamid name**

**aardsda01 David Allan (David) Aardsma SEA Seattle Mariners**

**abadfe01 Fernando Antonio (Fernando) Abad HOU Houston Astros**

**affelje01 Jeremy David (Jeremy) Affeldt SFN San Francisco Giants**

1. Using the Salaries table, a nested subquery similar to that in #3 and a SOME clause to find the list of players on the NY Yankees and their salary that have a salary higher than at least one other player on the team in 2014

**playerid teamid yearid Salary**

**abbotji01 NYA 1993 $2,350,000.00**

**abbotji01 NYA 1994 $2,775,000.00**

**abreubo01 NYA 2007 $15,000,000.00**

1. Using the Salaries table, select the team name, minimum, average and maximum salary for all teams in the AL league in 2015

**lgid name Minimum Salary Average Salary Maximum Salary**

AL Baltimore Orioles $514,000.00 $4,108,744.04 $13,333,333.00

AL Boston Red Sox $508,500.00 $5,659,481.25 $19,750,000.00

AL Chicago White Sox $510,000.00 $4,161,988.89 $15,750,000.00

1. Using the Salaries table, a HAVING clause, and a subquery similar to that used in the question #4, find the teamids and average salary of the teams with and average salary greater than the average salary of the NY Mets in 2010 HINT: Use teamid = NYN to find the Mets information.

**teamid Average Salary**

BOS $5,601,632.17

CHN $5,429,962.96

NYA $8,253,335.56

1. Using the Master, Teams and Appearances tables, list the play name and team name of all players playing on the same team in 2010 and 2015. You must use a nested query in your answer. **HINT:** Use a subquery using the Appearances table as part of your from statement along with the appropriate comparisons in the WHERE statement to identify the teams and playerids in the 2010 to 2015 timeframe.

**Player name Team Name**

Affeldt, Jeremy San Francisco Giants

Alvarez, Pedro Pittsburgh Pirates

Andrus, Elvis Texas Rangers

1. Using an All clause and a subquery using the Salaries table, find the names of all players who played in 2016 and had a salary higher than all the players who played in 2013. You will need to use the Master, Teams and Salaries tables in the main query to get the required information

**Player Name Team Name**

Greinke, Zack Arizona Diamondbacks

Price, David Boston Red Sox

Kershaw, Clayton Los Angeles Dodgers

**NOTE: for the queries from this point on, there will need to be GO statements in the script to make sure the script executes correctly. Remember, I should be able to click the Execute button and have all 15 queries run multiple times without any intervention on my part. Five (5) points will be taken off if your script is not able to do this.**

1. Using a WITH clause to calculate the average salary by team for 2012, find the names of all the player, their salaries and the difference between their salary and the average salary of their team who played in 2012. Use the subquery created by the WITH clause along with the Master, Salaries and Teams tables in the FROM clause of the main query.

**Team Name Player Name Salary Difference**

Arizona Diamondbacks Bell, Heath $10,000,000.00 $7,266,487.67

Arizona Diamondbacks Bloomquist, Willie $1,900,000.00 ($833,512.33)

Arizona Diamondbacks Cahill, Trevor $5,700,000.00 $2,966,487.67

1. Using WITH statements to create 3 subqueries (one each for the Manager Win %, Team Win % and Team Names, find the managers who’s win percent (total of W/total of G) for all the years they managed is greater that the win percentage of all the years for that team. You will need to use the Managers table along with the 3 temporary tables created using the WITH clause in the main query. Only show managers who’s win percent is higher than the team’s and sort your answer by the team name.

**Team Name Manager Manager Percent TeamPercent Per\_Difference**

Anaheim Angels Maddon, Joe 65.52% 51.23% 14.28%

Anaheim Angels Scioscia, Mike 52.47% 51.23% 1.23%

Arizona Diamondbacks Showalter, Buck 51.44% 48.83% 2.61%

1. Using a scalar query using the Appearances table and the Master table, find the number of teams each player played for.

**Player # of Teams**

Abbott, Fred 2

Abbott, Glenn 3

Abercrombie, Reggie 2

1. Add a column to the MASTER table named NJITID\_Avg\_Salary (NJITID is your NJIT ID such as JM234\_Avg\_Salary) and write a query that will calculate the player’s average salary and update the new column with that information. Your SQL should use an IF and GO statements to check and see if the new column exists before adding it. After updating the information in the new column, write a query that shows the player name and the value in the new column you created. Exclude the players where the salary information is null. You will need to use the Master and Salaries tables for this problem.

**Player name Total Salary**

Aardsma, David $9,259,750.00

Aase, Don $2,300,000.00

Abad, Andy $327,000.00

1. The player’s new contract says that each team must contribute to a player’s 401K retirement account. The contract says 10% of the salary must be put in the 401K for players making under $2 million dollars and 5% for players making $2 million and above. Create a column (with the appropriate IF and GO statement) in the Salaries table named NJITID\_401K (NJITID is your NJIT ID such as JM234\_401K) and write a query to populate the data from all past years. Use a CASE clause to update the column using the correct percentage. Also write a query to show the playerid, salary and 401K amount for each year. You will need to use only the Salaries table for this problem.

**playerid yearid Salary Amount 401K**

durhale01 1985 $800,000.00 $80,000.00

dwyerji01 1985 $375,000.00 $37,500.00

dwyerji01 1986 $400,000.00 $40,000.00

1. Using the Master and Salaries tables, show the playerID, full names as formatted below and birthdates (properly formatted) and salary for any Yankee(NYA) playing in 1990 whose salary is greater than the salary of any Boston Red Sox (BOS) using the salaries and mater tables sort highest to lowest salary

**playerID Full Name Birth Day NYA Salary**

mattido01 Donald Arthur (Don) Mattingly 4/20/1961 $2,500,000.00

winfida01 David Mark (Dave) Winfield 10/3/1951 $1,958,652.00

righeda01 David Allan (Dave) Righetti 11/28/1958 $1,550,000.00

1. Using only the Master table for this problem. Players are inducted into the Hall of Fame on July 7th each year. Calculate the exact age (using days, months and years) of the players when they were inducted into the Hall of Fame. You will need to use a CONVERT function to properly combine the 3 date columns to get a derived column you can use in the required calculation. To get the exact age, you will need to use a DATEDIFF function using the month and then convert the months to years to get the exact age. Your results should have the playerID, Full Name, Birth Year, Inducted Year, Full Birth Date (in the proper format), Full Induction Date (in the proper format) and the age in years. **Note:** subtracting the 2 years will give an incorrect answer.

**Hint:** I used the DATEDIFF function in the main query and used subqueries to calculate the 2 dates needed for the DATEDIFF. I calculated 1 date in each subquery to break up the logic and make the query easier to debug. You can easily manually check to see if the age is correct.

**playerid Player Name birth yearID cate- calcdate inductdate age**

**year gory**

aaronha01 Hank Aaron 1934 1982 Player 2 -5 -1934 07-25-1982 48

alexape01 Pete Alexander 1887 1938 Player 2 -26-1887 07-25-1938 51