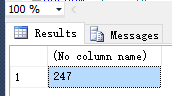
————————————questions——————————————————————

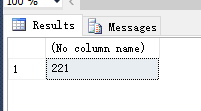
**1 how many players are inducted into the hall of fame?**



SELECT COUNT(playerID) FROM HallOfFame

WHERE inducted='Y'AND category='Player'

**b) how many distinct players in the batting data set are in the hall of fame?**



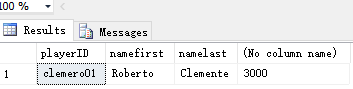
SELECT COUNT(DISTINCT Batting.playerID) FROM Batting

WHERE Batting.playerID IN

(SELECT HallOfFame.playerID FROM HallOfFame

WHERE category='Player' AND inducted='Y')

**2 a) what player has exactly 3000 hits total for his career?**



SELECT Batting.playerID,namefirst,namelast,SUM(Hits) FROM Batting,Players

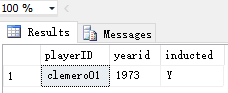
WHERE Batting.playerID=Players.playerid

--use playerID to connect 2 tables

GROUP BY Batting.playerID,namefirst,namelast

Having SUM(Hits)=3000

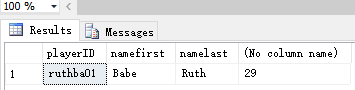
**b) what year was he inducted into the hall of fame?**



SELECT playerID,yearid,inducted FROM HallOfFame

WHERE playerID='clemero01'

**3 what player had the highest number of homeruns in 1919?**



SELECT TOP 1 Batting.playerID,namefirst,namelast,SUM(Homeruns) FROM Batting,Players

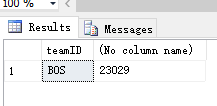
--use top 1 to choose the first row from result

WHERE yearid=1919 AND Batting.playerID=Players.playerid

GROUP BY Batting.playerID,namefirst,namelast

ORDER BY SUM(Homeruns) DESC

**4 what team has had the most hits after the year 2000, how many?**



SELECT TOP 1 teamID,SUM(Hits) FROM Batting

WHERE yearID>2000

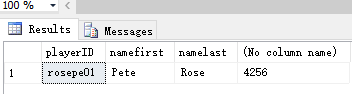
--choose teams' sum of hits after year 2000

GROUP BY teamID

ORDER BY SUM(Hits) DESC

**5 a )what player has the most cumulative hits?**

**b) how many does he have?**



SELECT TOP 1 Batting.playerID,namefirst,namelast,SUM(Hits) FROM Batting,Players

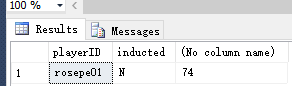
WHERE Batting.playerID=Players.playerid

Group by Batting.playerID,namefirst,namelast

Order by SUM(Hits) DESC

**c) is he in the hall of fame?**

**d) how many votes has he received total?**



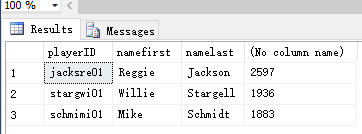
SELECT playerID, inducted,SUM(votes) FROM HallOfFame

WHERE playerID='rosepe01'

Group by playerID

**6 a) what 3 players have the highest number of cumulative strikeouts over their career that have been inducted into the hall of fame?**

**b)how many total strikeouts do they each have?**



SELECT TOP 3 HallOfFame.playerID,namefirst,namelast,SUM(strikeOuts) FROM Batting,HallOfFame,Players

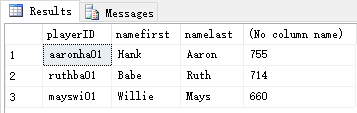
WHERE HallOfFame.playerID=Batting.playerID AND HallOfFame.playerID=Players.playerid AND inducted='Y'

--connect 3 tables with playerID and set inducted value as Y

Group by HallOfFame.playerID,namefirst,namelast

ORDER BY SUM(strikeOuts) DESC

**7 a) what 3 players have the highest cumulative total number of homeruns and have been inducted into the hall of fame?**



SELECT TOP 3 HallOfFame.playerID,namefirst,namelast,SUM(Homeruns) FROM Batting,HallOfFame,Players

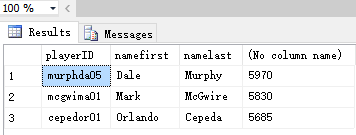
WHERE HallOfFame.playerID=Batting.playerID AND HallOfFame.playerID=Players.playerid AND inducted='Y'

--connect 3 tables with playerID and set inducted value as Y

Group by HallOfFame.playerID,namefirst,namelast

ORDER BY SUM(Homeruns) DESC

**b) what 3 players have the highest cumulative total number of homeruns and have NOT been inducted into the hall of fame?**



SELECT TOP 3 HallOfFame.playerID,namefirst,namelast,SUM(Homeruns) FROM Batting,HallOfFame,Players

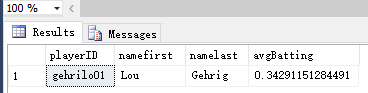
WHERE HallOfFame.playerID=Batting.playerID AND HallOfFame.playerID=Players.playerid AND inducted='N'

--connect 3 tables with playerID and set inducted value as N

Group by HallOfFame.playerID,namefirst,namelast

ORDER BY SUM(Homeruns) DESC

**8 what player had the highest batting average from 1930 to 1940 with a cumulative total of hits over 1800 and has NOT been elected into the hall of fame? (batting average = hits/atbats)**



SELECT Batting.playerID,SUM(Hits) as SUMHits,SUM(Atbats) as SUMAt into BatAVG

--Caculate the sum of Hits and sum of Atbats and save as sh and sab in table BatAVG

FROM Batting

WHERE yearID between 1930 AND 1940

GROUP BY playerID

--Caculate the sum of Hits and sum of Atbats by players

SELECT TOP 1 BatAVG.playerID,namefirst,namelast,CAST(SUMHits as float)/SUMAt as avgBatting FROM BatAVG,Players

--use cast to get the average of Hits/Atbats

--use float to preserve decimals

WHERE BatAVG.playerID=Players.playerid AND SUMAt>0 AND SUMHits>1800 AND BatAVG.playerID IN

--Get the average of Hits/Atbats only if the sum of Atbats>0 and Hits over 1800

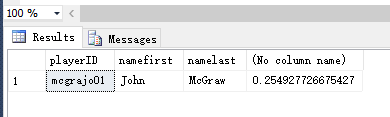
(SELECT HallOfFame.playerID FROM HallOfFame

--choose the players not inducted in the hall of fame,nesting query

WHERE inducted='N')

ORDER BY avgBatting DESC

**9 what player averaged the highest number of walks from 1900-1920 with over 500 atbats?**



SELECT playerID,SUM(Walks) AS SUMWalks,SUM(Atbats) AS SUMAt

INTO WalkAVG

--Create a table called WalkAVG to save the sum of walks and sum of Atbats records

FROM Batting

WHERE yearID BETWEEN 1900 AND 1920

GROUP BY playerID

--Caculate the sum of Walks and sum of Atbats by players

SELECT TOP 1 WalkAVG.playerID,namefirst,namelast,CAST(SUMWalks AS float)/SUMAt

FROM WalkAVG,Players

WHERE WalkAVG.playerID=Players.playerid AND SUMAt>500

ORDER BY CAST(SUMWalks AS float)/SUMAt DESC

**10 write a query that uses a non equi join and explain what you are doing.**

Q: want to know whether every player-year in table HallOfFame has records in table Batting

A: 3936 rows

![C:\Users\lili8\AppData\Roaming\Tencent\Users\328900451\QQ\WinTemp\RichOle\](J)Q}T)XPX788B4H5TJB4W.png](data:image/png;base64,)

SELECT HallOfFame.yearid, HallOfFame.playerID, HallOfFame.inducted

INTO answ10

FROM Batting, HallOfFame

WHERE HallOfFame.category = 'Player' AND HallOfFame.playerID = Batting.playerID

AND HallOfFame.inducted ='Y' AND HallOfFame.yearid != Batting.yearID;

--3936 records without Batting info

SELECT COUNT(DISTINCT playerID)

FROM answ10

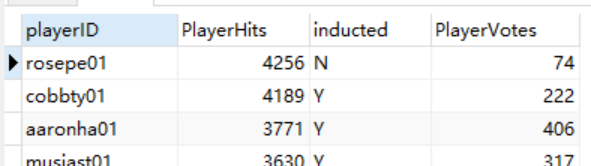
--impacts 221 players

**11 write a query that uses an OUTER Join and explain what you are doing.**

Q:Another answer version to question

5. a) what player has the most cumulative hits? b) how many does he have? c) is he in the hall of fame? d) how many votes has he received total?

A:



SELECT Batting.playerID, SUM(Batting.Hits) as PlayerHits

INTO sumHits

FROM Batting

GROUP BY Batting.playerID;

--create a table contains accumulative hits of every player in Batting

SELECT playerID, inducted, sum(votes) as PlayerVotes

INTO sumVotes

FROM HallOfFame

GROUP BY playerID, inducted;

-- create a table contains accumulative votes of every player in table HallOfFame

SELECT sumHits.playerID, sumHits.PlayerHits, sumVotes.inducted, sumVotes.PlayerVotes

FROM sumVotes

FULL OUTER JOIN sumHits

ON sumVotes.playerID = sumHits.playerID

ORDER BY PlayerHits desc;