Kramer Johnson

CSPC 5021 02 HW4

11/2/20

1. We were able to get our team database into the AWS cloud with no issues.

a.

/\* Display top 20 shortstops by slugging percentage \*/

SELECT

nameFirst AS 'First Name',

nameLast AS 'Last Name',

Slugging

FROM (

SELECT

p.playerID,

nameFirst,

nameLast,

POS,

AB,

b.yearId,

((H - 2B - 3B - HR) + (2B \* 2) + (3B \* 3) + (HR \* 4)) / AB AS Slugging # Calculate slugging pct

FROM people p

JOIN batting b ON p.playerID = b.playerID

JOIN fielding f ON p.playerID = f.playerID

WHERE

POS = 'SS' AND # Only players who played as shortstop

b.yearID = '2019' AND # Only data from 2019

f.G > 120 # Only players who played as SS for more than 75% of games in 2019

GROUP BY p.playerID

) t1

WHERE

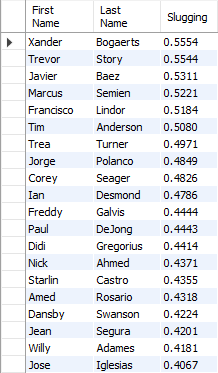
AB > 300 AND # Players with at least 300 at bats

Slugging > .400 # Slugging pct over .400

ORDER BY Slugging DESC

LIMIT 20;

1. The above SQL statement accurately retrieves the top shortstops by slugging percentage from 2019 (the last year we have data). The query ended up being more difficult than I expected and I may have overcomplicated it with the subquery. The squery is used to retrieve data about only players who played at least 120 games (3/4 of the season) as a shortstop in the year 2019. From this subquery, I pull the at bats (AB) and make sure they are over 300. I pull the slugging attribute and only get batters who had a percentage over .400. I order the results from the greatest slugging to lowest and display only the top 20.



c.

/\* Display home run leading shortstops for last 10 years \*/

SELECT

p.playerID,

p.nameFirst AS 'First Name',

p.nameLast AS 'Last Name',

SUM(b.HR) AS 'Home Runs',

ROUND(AVG(b.HR)) AS 'Average HR'

FROM people p

JOIN batting b ON p.playerID = b.playerID

JOIN (

# Get table of shortstops in past 10 years

SELECT playerID, POS

FROM fielding

WHERE

yearID >= '2009' AND # only last 10 years

POS = 'SS' AND # only shortstops

G > 120 # only players who played 75% of the season as shortstop

GROUP BY playerID

) f ON p.playerID = f.playerID

WHERE

b.yearID >= 2009

GROUP BY playerID

ORDER BY SUM(b.HR) DESC

LIMIT 20;

EX.

/\* Display home run leading shortstops over the past 10 years that are under 30 \*/

SELECT

p.playerID,

p.nameFirst AS 'First Name',

p.nameLast AS 'Last Name',

YEAR(CURDATE()) - p.birthYear AS Age,

SUM(b.HR) AS 'Total HR',

ROUND(AVG(b.HR)) AS 'Avg HR'

FROM people p

JOIN batting b ON p.playerID = b.playerID

JOIN (

# Get table of shortstops in past 10 years

SELECT playerID, POS

FROM fielding

WHERE

yearID >= '2009' AND # only last 10 years

POS = 'SS' AND # only shortstops

G > 120 # only players who played 75% of the season as shortstop

GROUP BY playerID

) f ON p.playerID = f.playerID

WHERE

b.yearID >= 2009 AND # Show results for past 10 years

(YEAR(CURDATE()) - p.birthYear) < 30 # Only players under 30

GROUP BY playerID

ORDER BY SUM(b.HR) DESC

LIMIT 20;

/\*

\* Stored procedure to get home run leading players since a

\* given year who are currently under a specified age

\*

\* @param maxAge INT the maximum age player to search for

\* @param sinceYear INT the beginning year you want data for

\* @param position CHAR(2) the position player you want to search for

\*/

DELIMITER //

CREATE PROCEDURE filterPlayers(

IN maxAge INT,

IN sinceYear INT,

IN position CHAR(2)

)

BEGIN

SELECT

p.playerID,

p.nameFirst AS 'First Name',

p.nameLast AS 'Last Name',

YEAR(CURDATE()) - p.birthYear AS Age,

SUM(b.HR) AS 'Total HR',

ROUND(AVG(b.HR)) AS 'Avg HR'

FROM people p

JOIN batting b ON p.playerID = b.playerID

JOIN (

# Get table of shortstops in past 10 years

SELECT playerID, POS

FROM fielding

WHERE

yearID >= sinceYear AND # only since sinceYear

POS = position AND # only players who played postion

G > 120 # only players who played 75% of the season as shortstop

GROUP BY playerID

) f ON p.playerID = f.playerID

WHERE

b.yearID >= sinceYear AND # show results since sinceYear

(YEAR(CURDATE()) - p.birthYear) < maxAge # only players under the max age

GROUP BY playerID

ORDER BY SUM(b.HR) DESC

LIMIT 20;

END //

DELIMITER ;