



ECE 467 TERM PROJECT REPORT



Jacob Ahrens and Simon Jaimes

1. Description:

Our database is a database of baseball statistics from professional baseball for each year that there are publicly available and verified statistics for. Our database is based off the data collected for the Lahman Database through the 2023 MLB season. We have three main entities being players, teams, and franchises with our other tables coming from relationships between teams and players as representations for the statistical data.

The entities represent the basis of the databases and tie the statistics to specific teams, players, and years. The player entity represents a player with at least one recorded statistic at the highest level of professional baseball at the time, and contains data about the player such as weight, height, name, and professional debut. The franchise entity represents the organization responsible for fielding teams across multiple years and has minimal data, only the franchise name, the national association of the franchise, and if the franchise is currently active as of 2023. The team entity represents the team fielded by a franchise in a given year and the data tied with that specific team, such as record, league and division, or attendance numbers.

The relations represent the main bulk of the statistical data as a relationship between a player and a team as the player recording the specific category of statistic for the team. The first relationship is batting statistics which represent all simple statistics a player records throughout a season for a team as a hitter such as, at bats, home runs, and walks. The second relationship is pitching statistics which represent all simple statistics a player records throughout a season for a team as a pitcher such as, innings pitched, earned runs, and strikeouts. The third relationship is fielding statistics which represent all simple statistics a player records throughout a season for a team as a fielder such as, put outs, assists, and errors. The fourth relationship is appearances which represent the number of times a player has appeared for a team with total appearances as well as appearances per position.

2. Relation Sizes:

PLAYER: 21,010 tuples with indexes at: playerId

TEAM: 3,045 tuples with indexes at: teamID, year

FRANCHISE: 120 tuples with indexes at: franchID

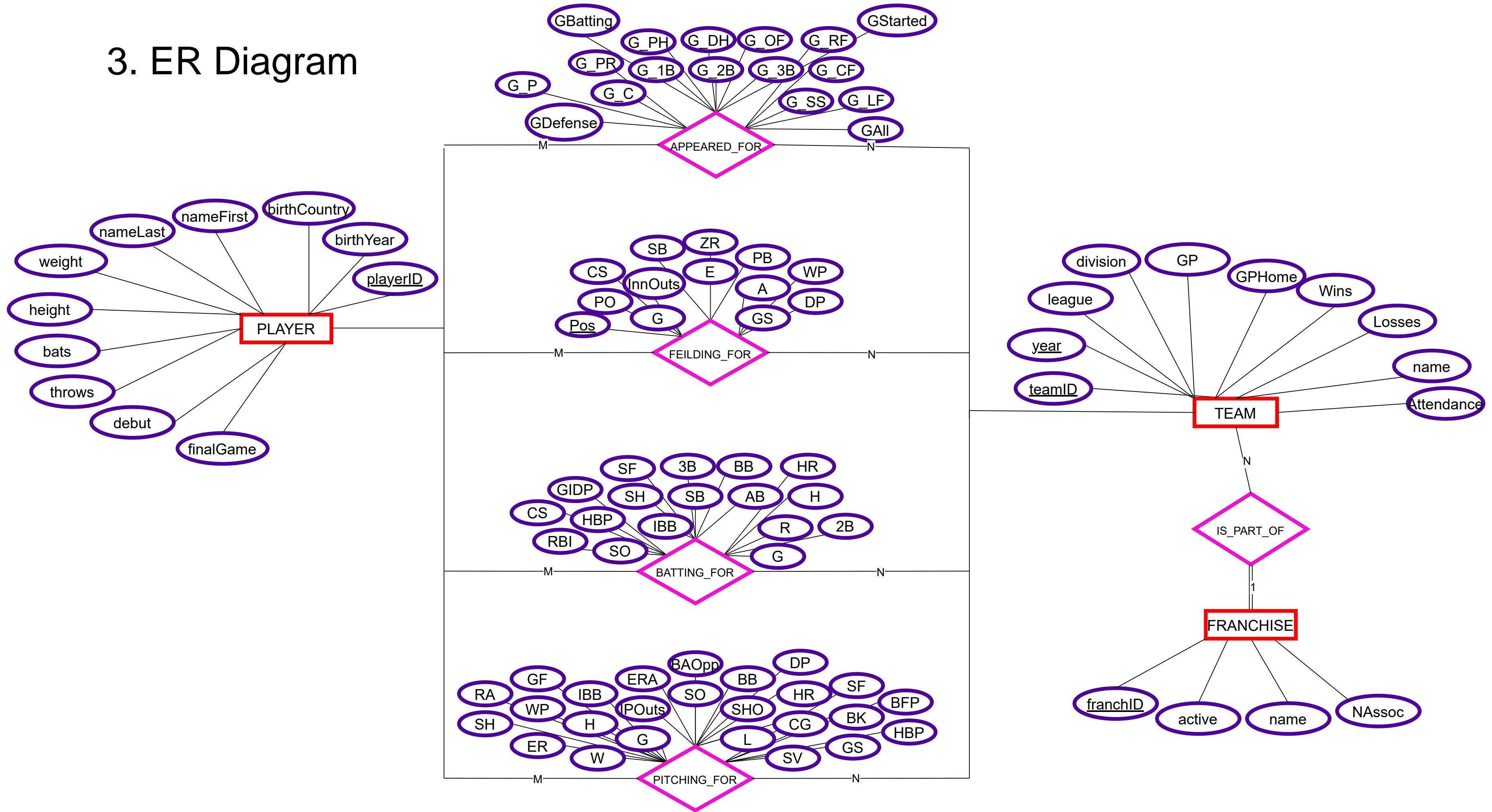
BATTING: 113, 721 tuples with indexes at: playerId, teamID, year

PITCHING: 51,347 tuples with indexes at: playerId, teamID, year

FIELDING: 151,414 tuples with indexes at: playerId, teamID, year, Pos

APPEARANCES: 113,653 tuples with indexes at: playerId, teamID, year

3. ER Diagram



4. Relational Schema

a. PLAYER

The PLAYER table represents the players who the statistics are about and are the focal point of the entire database. The key is the playerID of which each player has a unique ID and there are no foreign keys. The attributes include birthYear and birthCountry which represent the year and place a player was born, nameFirst and name Last for the player's name, weight and height which are self-explanatory, bats and throws which represent the handedness of the player when hitting or throwing respectively, and debut and finalGame which represent the first and last game a player had played in.

b. TEAM

The TEAM table represents the team fielded for a given franchise each year. The keys are a combination of teamID and year as teams of the same name will have the same teamID but different attributes for different years and the foreign key is the franchID which references the FRANCHISE table. The attributes include league and division which are organizations of teams within the entire major league, GP and GPHome which are the number of games the team played in total and at home respectively, Wins and Losses which are the number of each result for the games played, name which is what the team was known as colloquially, and attendance which is the number of spectators that attended the teams home games.

c. FRANCHISE

The FRANCHISE table represents the organization that fields multiple teams across multiple years. The key is the franchID which is a unique ID for each franchise and there are no foreign keys. The attributes are active which represent if the franchise is still actively fielding new teams each year and NAssoc which is the national association that the franchise is affiliated with.

d. BATTING

The BATTING table represents the hitting statistics of a player for a team in a year. The keys are playerID, teamID, and year, with playerID referencing the PLAYER table and teamID and year referencing the TEAM table as the foreign keys. The attributes are all individual batting statistics seen here:

G – Games | AB - At Bats | R – Runs | H – Hits | 2B – Doubles | 3B – Triples | HR – Homeruns |
RBI - Runs Batted In | SB - Stolen Bases | CS - Caught Stealing | BB - Base on Balls |
SO – Strikeouts | IBB - Intentional walks | HBP - Hit by pitch | SH - Sacrifice hits |
SF - Sacrifice flies | GIDP - Grounded into double plays

e. PITCHING

The PITCHING table represents the pitching statistics of a player for a team in a year. The keys are playerId, teamID, and year, with playerId referencing the PLAYER table and teamID and year referencing the TEAM table as the foreign keys. The attributes are all individual pitching statistics seen here:

W - Wins | L - Losses | G - Games | GS - Games Started | CG - Complete Games |
SHO - Shutouts | SV - Saves | IPOuts - Outs Pitched (innings pitched x 3) | H - Hits |
ER - Earned Runs | HR - Homeruns | BB - Walks | SO - Strikeouts | BAOpp - Opponent's
Batting Average | ERA - Earned Run Average | IBB - Intentional Walks | WP - Wild Pitches |
HBP - Batters Hit By Pitch | BK - Balks | BFP - Batters faced by Pitcher | GF - Games Finished
| RA - Runs Allowed | SH - Sacrifices by opposing batters | SF - Sacrifice flies by opposing
batters | DP - Grounded into double plays by opposing batter

f. FIELDING

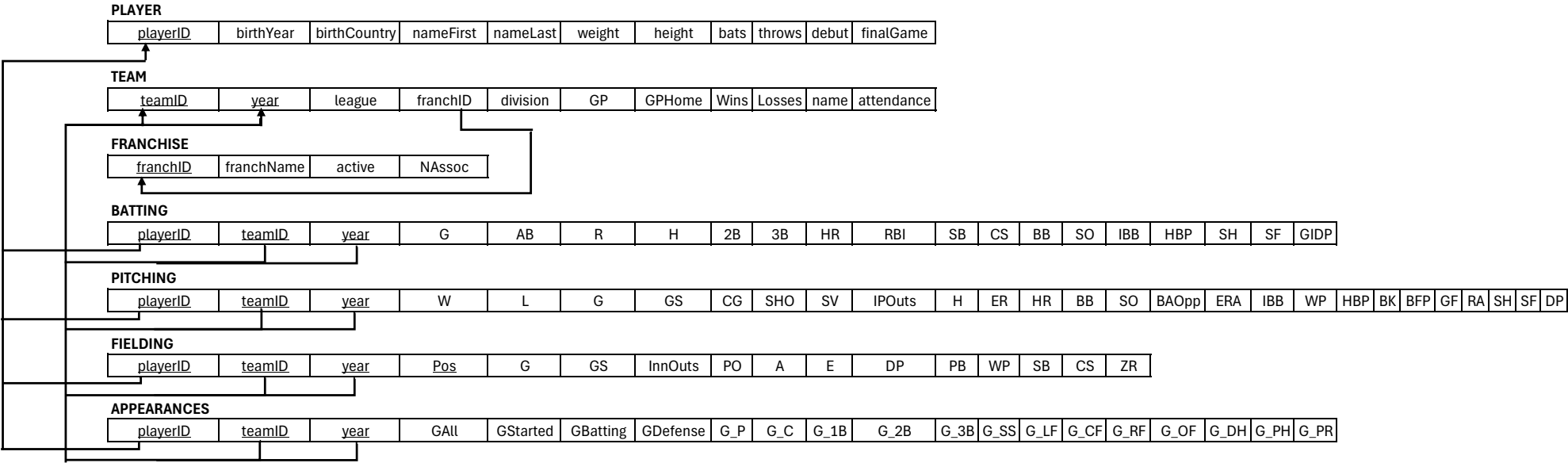
The FIELDING table represents the fielding statistics of a player for a team in a year at a position. The keys are playerId, teamID, year, and pos with playerId referencing the PLAYER table and teamID and year referencing the TEAM table as the foreign keys. The attributes are all individual fielding statistics seen here:

Pos - Position | G - Games | GS - Games Started | InnOuts - Time played in the field
expressed as outs | PO - Putouts | A - Assists | E - Errors | DP - Double Plays | PB - Passed
Balls (by catchers) | WP - Wild Pitches (by catchers) | SB - Opponent Stolen Bases (by
catchers) | CS - Opponents Caught Stealing (by catchers) | ZR Zone Rating

g. APPEARANCES

The APPEARANCES table represents the appearances by a specific player for a specific team in a specific year. The keys are playerId, teamID, and year, with playerId referencing the PLAYER table and teamID and year referencing the TEAM table as the foreign keys. The attributes are all individual appearance statistics seen here:

GALL - Total games played | GStarted - Games started | GBatting - Games in which player
batted | GDefense - Games in which player appeared on defense | G_p - Games as pitcher |
G_c - Games as catcher | G_1b - Games as first baseman | G_2b - Games as second
baseman | G_3b - Games as third baseman | G_ss - Games as shortstop | G_lf - Games as
left fielder | G_cf - Games as center fielder | G_rf - Games as right fielder | G_of - Games as
outfielder | G_dh - Games as designated hitter | G_ph - Games as pinch hitter | G_pr -
Games as pinch runner



5. Views:

a. ActivePlayers

The ActivePlayers view selects the players that are still playing as of the most recent update of the data set. These players are selected from the PLAYER table on the condition that the year of their finalGame attribute is the year of the most recent data. For our data set the year was 2023 however only players who made their debut in 2023 had updated finalGame attributes and all other active players had a finalGame in 2022 so that was taken into account.

The view is interesting because it allows for queries to be done only with players that are still currently playing, which can provide insights into an upcoming season. Some insights into the upcoming season can be predictions for performance, tracking of current players compared to specific records, and comparison of past performance of players playing head-to-head during the current season. Another interesting use of the view can be for a user that may only want to look for players that they are familiar with which are most likely to be current or active players.

b. QualifiedBatters

The QualifiedBatters view selects players that met the MLB's minimum requirements to be eligible for awards in a season, also known as "Qualified Hitters". The MLB requirement to be a "Qualified Hitter" is for a player to have over 3.1 plate appearances per game and play in over half their team's games in a season. Plate Appearances are not a stat that is in our data set but can be derived from the stats we have, as a Plate Appearance (PA) is considered any time that a player sees at least one pitch as a hitter. PA can then be calculated through the sum of At Bats (AB), Sacrifice Hits (SH), Sacrifice Flies (SF), and Walks (BB) which are all possible outcomes of a Plate Appearance and are all statistics in our data set.

The view is interesting because it provides a simple way to filter out junk batting data that may be based on a small sample size and does so based on the official requirement from the league itself. The filtering of "qualified" players allows for certain queries that may be based on average performance such as Batting Average, On Base Percentage, or Slugging to be used to provide a more accurate insight into performance as there is not a skew due to the sample size.

c. Career Stats

Career Stats is not a specific view in our database but rather a collection of four views that are all extremely similar just with different statistics. Each Career view takes the sum of all

statistics for the same playerId through different years and teams and provides a table with the playerId and the sum of all their statistics. There is a view for each of BATTING, PITCHING, FIELDING, and APPEARANCES that is the table name with Career in front of it. Because the queries for the views are so similar in both execution and use, I am treating them as one for the purposes of the report.

The views are interesting because they allow for insight into a player's performance over their entire career which is often used to compare players between each other as consistency throughout multiple years is more impressive than a small number of good years. Career statistics allow for another angle on insight of the data that is not provided through the data set itself.

6. Queries:

```
--#1
--List all players who have played for 10 or more franchises
SELECT
    P.nameFirst || ' ' || P.nameLast AS Name,
    COUNT(DISTINCT T.franchID) AS Franchises
FROM PLAYER P, TEAM T, FRANCHISE F, APPEARANCES A
WHERE P.playerID = A.playerID AND --Joins the player to franchise as "played
for"
    A.teamID = T.teamID AND
    T.franchID = F.franchID
GROUP BY P.playerID --Groups all instance of player into one row
HAVING Franchises > 10 --Selects the players that have 10+ different
franchises
ORDER BY Franchises DESC;
```

Name	Franchises
Edwin Jackson	15
Ron Villone	13
Matt Stairs	13
Rich Hill	13
Octavio Dotel	13
Gus Weyhing	12
Julian Tavarez	12
Fernando Rodney	12
Mike Morgan	12
Deacon McGuire	12
Royce Clayton	12
Miguel Batista	12
Todd Zeile	11
Rick White	11
Pop Smith	11
Orator Shafer	11
Rene Rivera	11
Dennys Reyes	11
Terry Mulholland	11
Kenny Lofton	11
LaTroy Hawkins	11
Jose Guillen	11
Joe Gerhardt	11
Jack Doyle	11
Ross Detwiler	11
Bartolo Colon	11
Tyler Clippard	11
Bruce Chen	11
Marlon Byrd	11
George Bradley	11
Henry Blanco	11
Paul Bako	11

```
--List pitchers with most wins for all teams to win 100 or more games in a
season
```

```
P.nameFirst || ' ' || P.nameLast AS Name,
PI.W AS "Pitcher Wins",
T.name AS Team,
T.year AS Year,
T.Wins || '-' || T.Losses AS "Team Record"
```

WHERE P.playerID = PI.playerID AND --Join pitchers and teams

PI.year = T.year AND --Have to join team to pitchers by specific years

PI.W = (--Compares the pitcher to the most wins for a team

WHERE teamID = T.teamID AND

given team

team

Name	Pitcher Wins	Team	Year	Team Record																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
------	--------------	------	------	-------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```
--#3
/*List all players to have more than 50 innings pitched (150 IPOuts)
and more than 100 hits in the same season ordered by batting average */
SELECT
    P.nameFirst || ' ' || P.nameLast AS Name,
    B.year AS year,
    ROUND(PI.IPOuts / 3.0, 1) AS InningsPitched,
    B.H AS Hits,
    ROUND(1.0 * B.H / B.AB, 3) AS AVG
FROM PLAYER P, PITCHING PI, BATTING B
WHERE P.playerID = PI.playerID AND --Joins the player to the batting and
pitching stats
    P.playerID = B.playerID AND
    B.year = PI.year AND --Makes sure stats are from same year
    PI.IPOuts > 150 AND B.H > 100 --Filters players by requirements
ORDER BY AVG DESC;
```

Name	year	InningsPitched	Hits	AVG
Bob Caruthers	1887	341.0	130	0.357
Dave Foutz	1887	339.3	151	0.357
Cal McVey	1876	59.3	107	0.347
Guy Hecker	1886	420.7	117	0.341
George Van Haltren	1890	223.0	126	0.335
Bob Caruthers	1886	387.3	106	0.334
Al Spalding	1874	617.3	119	0.329
Al Spalding	1873	496.7	106	0.328
Babe Ruth	1919	133.3	139	0.322
Guy Hecker	1887	285.3	118	0.319
Al Spalding	1875	570.7	107	0.312
Jesse Burkett	1890	118.0	124	0.309
Kid Gleason	1895	50.3	130	0.309
Dick Burns	1884	329.7	107	0.306
Shohei Ohtani	2023	132.0	151	0.304
John Coleman	1885	60.3	119	0.299
Win Mercer	1899	186.0	112	0.299
Adonis Terry	1887	318.0	103	0.293
Jimmy Callahan	1897	189.7	105	0.292
John Ward	1879	587.0	104	0.286
Old Hoss Radbourn	1883	632.3	108	0.283
Jim Whitney	1883	514.0	115	0.281
Dave Foutz	1886	504.0	116	0.28
Charlie Hickman	1901	65.0	113	0.278
Adonis Terry	1890	370.0	101	0.278
Bob Caruthers	1892	101.7	142	0.277
Dave Foutz	1888	176.0	156	0.277
Dave Foutz	1889	59.7	152	0.275
Elmer Smith	1892	134.0	140	0.274
Shohei Ohtani	2022	166.0	160	0.273
Blondie Purcell	1883	80.0	114	0.268
William Robinson	1884	75.0	111	0.267
Dave Foutz	1891	52.0	134	0.257
Shohei Ohtani	2021	130.3	138	0.257
John Ward	1884	60.7	122	0.253
Ed Daily	1888	73.7	102	0.225

```
--#4
/* List all teams with multiple players to have less than 10 errors in the
field
and more than 10 home runs in the same season */
SELECT
    T.name AS Team,
    T.year AS Year,
    COUNT(DISTINCT F.playerID) AS numPlayers
FROM PLAYER P, FIELDING F, BATTING B, TEAM T
WHERE P.playerID = F.playerID AND --Joins the players to the fielding and
batting stats
    P.playerID = B.playerID AND
    F.year = B.year AND --Makes sure stats are from the same season
    F.teamID = B.teamID AND --Ensures the stats are for the same team
    T.teamID = F.teamID AND
    T.year = F.year AND --Ensures the team's year matches the stats
    F.E <= 3 AND B.HR > 25 --Filters players by requirements
GROUP BY T.teamID, T.name, F.year --Groups by team and year to get the number
of players for each team
HAVING numPlayers >= 3 --Ensures the "multiple players" requirement is met
ORDER BY numPlayers DESC;
```

Team	Year	numPlayers
Houston Astros	2019	5
Los Angeles Dodgers	2019	4
Minnesota Twins	2009	4
Oakland Athletics	1999	4
Atlanta Braves	1966	3
Atlanta Braves	2004	3
Baltimore Orioles	1996	3
Baltimore Orioles	2015	3
Baltimore Orioles	2016	3
Boston Red Sox	2009	3
Boston Red Sox	2011	3
Boston Red Sox	2016	3
Boston Red Sox	2019	3
Chicago Cubs	2017	3
Chicago Cubs	2019	3
Cincinnati Reds	2017	3
Colorado Rockies	1995	3
Los Angeles Angels of Anaheim	2015	3
Los Angeles Dodgers	2023	3
Milwaukee Brewers	2018	3
Minnesota Twins	2019	3
Milwaukee Braves	1962	3
New York Yankees	2010	3
New York Yankees	2015	3
New York Yankees	2018	3
New York Yankees	2019	3
San Francisco Giants	1963	3
San Francisco Giants	1964	3
St. Louis Cardinals	2008	3
Texas Rangers	2003	3
Toronto Blue Jays	2017	3

```
--#5
--List all players to steal more than 15 bases without being caught stealing
ordered from highest to lowest
SELECT
    P.nameFirst || ' ' || P.nameLast AS Name,
    B.year,
    B.SB
FROM PLAYER P, BATTING B
WHERE P.playerID = B.playerID AND --Joins the players to the batting stats
    B.year >= 1951 AND --First year CS was an official stat to ensure players
we not caught stealing
    B.SB > 15 AND B.CS = 0 --Filters players by requirements
GROUP BY P.playerID --Groups by player to ensure only one row per player
ORDER BY B.SB DESC;
```

Name	year	SB
Trea Turner	2023	30
Carlos Beltran	2004	28
Chase Utley	2009	23
Alcides Escobar	2013	22
Kevin McReynolds	1988	21
Quintin Berry	2012	21
Paul Molitor	1994	20
Craig Gentry	2011	18
Tim Locastro	2019	17
Gary Thurman	1989	16
Jimmy Sexton	1982	16


```
--#6
/*List the season with the most appearances for all players to make over 25
appearances
in both the outfield and pitcher in the same season ordered by batting
averages in those seasons */
SELECT
    P.nameFirst || ' ' || P.nameLast AS Name,
    A.year,
    A.G_P AS GamesPitched,
    A.G_OF AS GamesOF,
    ROUND(1.0* B.H / B.AB, 3) AS AVG
FROM PLAYER P, APPEARANCES A, BATTING B
WHERE P.playerID = A.playerID AND --Joins the players to the batting and
appearances stats
    P.playerID = B.playerID AND
    B.year = A.year AND --Ensures the batting and appearances are from the
same year
    A.G_P > 25 AND A.G_OF > 25 --Filters players by requirements
GROUP BY P.playerID
ORDER BY AVG DESC;
```

Name	year	GamesPitched	GamesOF	AVG
George Van Haltren	1890	28	67	0.335
Dick Burns	1884	40	44	0.306
Walter Thornton	1898	28	34	0.295
Curry Foley	1880	36	35	0.292
Jim Whitney	1883	62	40	0.281
Cy Seymour	1898	45	35	0.276
Jack Manning	1875	27	65	0.27
Scott Stratton	1888	33	38	0.257
Jack Coombs	1908	26	50	0.255
John Ward	1883	34	56	0.255
Otto Hess	1905	26	28	0.254
Charlie Ferguson	1886	48	27	0.253
Ed Daily	1890	27	64	0.239
Old Hoss Radbourn	1882	54	32	0.239
Charlie Buffinton	1883	43	51	0.238
Adonis Terry	1886	34	32	0.237
John Coleman	1883	65	31	0.234
Mickey Welch	1883	54	38	0.234
Bill Wise	1884	50	43	0.233
Bob Caruthers	1888	44	51	0.23
Tony Mullane	1886	63	27	0.225
Tommy Bond	1880	63	26	0.22
Michael Lorenzen	2019	73	29	0.208
Charlie Sweeney	1885	35	39	0.206
Henry Boyle	1885	42	31	0.202
Jocko Flynn	1886	32	28	0.2
Dave Foutz	1892	27	29	0.186
George Weidman	1883	52	35	0.185

```
--#7
/*List all current players born outside of the US with a batting average
over .300 in a single season multiple times
and what their best season AVG was */

SELECT
    AP.nameFirst || ' ' || AP.nameLast AS Name,
    AP.birthCountry,
    ROUND(1.0* B.H / B.AB, 3) AS BestAVG,
    COUNT(*) AS Instances --Number of seasons they hit over .300
FROM ActivePlayers AP, BATTING B --Uses the current players veiw
WHERE AP.playerID = B.playerID AND --Joins the players to the batting stats
from the "ActivePlayers" view
    NOT(AP.birthCountry = 'USA') AND --Filters players by country of birth
outside of USA
    ROUND(1.0* B.H / B.AB, 3) > 0.3 --AVG
GROUP BY AP.playerID --Groups players who have multiple seasons hitting
over .300
HAVING INSTANCES > 1
ORDER BY INSTANCES DESC, BestAvg DESC; --Ordered by number of times then Best
Season AVG
```

Name	birthCountry	BestAVG	Instances
Miguel Cabrera	Venezuela	0.324	11
Robinson Cano	D.R.	0.342	10
Albert Pujols	D.R.	0.329	10
Joey Votto	CAN	0.321	9
Jose Altuve	Venezuela	0.341	6
Nelson Cruz	D.R.	0.311	5
Yadier Molina	P.R.	0.304	5
Jose Iglesias	Cuba	0.373	4
Luis Arraez	Venezuela	0.354	4
Jose Abreu	Cuba	0.317	4
Starling Marte	D.R.	0.311	4
Jose Ramirez	D.R.	0.333	3
Salvador Perez	Venezuela	0.331	3
Xander Bogaerts	Aruba	0.32	3
Yandy Diaz	Cuba	0.312	3
Avisail Garcia	Venezuela	0.304	3
Kenley Jansen	Curacao	1.0	2
Alcides Escobar	Venezuela	0.5	2
Dennis Santana	D.R.	0.5	2
Juan Soto	D.R.	0.351	2
Magneuris Sierra	D.R.	0.35	2
Marcell Ozuna	D.R.	0.338	2
Jhoulys Chacin	Venezuela	0.333	2
Yu Darvish	Japan	0.333	2
Donovan Solano	Colombia	0.33	2
Ketel Marte	D.R.	0.329	2
Raimel Tapia	D.R.	0.321	2
Jean Segura	D.R.	0.319	2
Yordan Alvarez	Cuba	0.313	2
Francisco Lindor	P.R.	0.313	2
Rob Refsnyder	South Korea	0.307	2
Ender Inciarte	Venezuela	0.303	2

```
--#8
--Select all players the are either top 10 in BA as a hitter, top 10 in SO
as a pitcher, or top 10 in Errors in the feild all time in their career while
meeting minimum attempts
SELECT
    P.nameFirst || ' ' || P.nameLast AS Name,
    ROUND(1.0* CB.H / CB.AB, 3) AS AVG,
    CP.SO AS Strikeouts,
    CF.E AS Errors
FROM PLAYER P, CareerBatting CB, CareerFielding CF, CareerPitching CP --Uses
the career views to get all time career stats
WHERE
    (P.playerID IN ( --Top 10 in career BA
        SELECT playerID FROM CareerBatting
        WHERE AB > 250 --Sets minimum at bats requirement
        ORDER BY ROUND(1.0* H / AB, 3) DESC, AB DESC --Gives top players that
met requirements with priority to those with more at bats
        LIMIT 10 )
    OR P.playerID IN ( --Top 10 in career Errors
        SELECT playerID FROM CareerFielding
        WHERE InnOuts > 300 --Sets minimum innings requirement
        ORDER BY E ASC, InnOuts DESC --Gives top players that met
requirements with priority to those with more innings
        LIMIT 10 )
    OR P.playerID IN ( --Top 10 in career SO
        SELECT playerID FROM CareerPitching
        WHERE IPOuts > 300 --Sets minimum innings requirement
        ORDER BY SO DESC, IPOuts ASC --Gives top players that met
requirements with priority to those with more innings
        LIMIT 10 ) ) AND
    P.playerID = CB.playerID AND --Joins the players to the batting,
pitching, and fielding stats
    P.playerID = CF.playerID AND
    P.playerID = CP.playerID
ORDER BY P.nameLast, P.nameFirst; --Ordered by name since ordering by stats
would be too messy
```

Name	AVG	Strikeouts	Errors
Ross Barnes	0.36	0	459
Bert Blyleven	0.131	3701	30
Jim Bullinger	0.188	392	0
Steve Carlton	0.201	4136	42
Roger Clemens	0.173	4672	25
Ty Cobb	0.366	0	278
Josh Collmenter	0.123	494	0
Randy Johnson	0.125	4875	57
Walter Johnson	0.235	3509	53
Greg Maddux	0.171	3371	53
Cal McVey	0.346	37	320
Levi Meyerle	0.356	0	342
Paul Mitchell		277	0
Lefty O'Doul	0.349	19	62
Gaylord Perry	0.131	3534	38
Nolan Ryan	0.11	5714	90
Tom Seaver	0.154	3640	42
Tris Speaker	0.345	0	224
Don Sutton	0.144	3574	31
Paul Wagner	0.166	452	0


```
--#9
--List the height and weight of players who have over 20 HRs, 100 H, 20 SBs
and 150 Assists in a season
--"Five tool players"
SELECT
    P.nameFirst || ' ' || P.nameLast AS Name,
    (P.height / 12) || ' ' || (P.height % 12) || ' ' AS Height, --Turns
hieght data into a more human friendly format
    P.weight || 'lbs' AS Weight,
    Pthrows,
    P.bats,
    B.HR,
    B.H,
    B.SB,
    F.A
FROM PLAYER P, FIELDING F, BATTING B
WHERE P.playerID = F.playerID AND --Joins the players to the fielding and
batting stats
    P.playerID = B.playerID AND
    B.year = F.year AND --Ensures the fielding and batting stats are from the
same year
    B.HR > 20 AND B.H > 150 AND B.SB > 20 AND F.A > 150 --Filters players by
requirements
GROUP BY P.playerID --Limits to one row per player
ORDER BY P.height, P.weight, P.nameLast; --Orders players by size then name
if same size
```

Name	Height	Weight	throws	bats	HR	H	SB	A
Jose Altuve	5' 6"	166lbs	R	R	24	216	30	361
Joe Morgan	5' 7"	160lbs	R	L	26	167	67	440
Jimmy Rollins	5' 7"	175lbs	R	B	25	191	36	446
Tommy Harper	5' 9"	165lbs	R	R	31	179	38	275
Davey Lopes	5' 9"	170lbs	R	R	28	154	44	384
Dustin Pedroia	5' 9"	170lbs	R	R	21	195	26	425
Jose Ramirez	5' 9"	190lbs	R	B	39	156	34	205
Bert Campaneris	5' 10"	160lbs	R	R	22	168	42	414
Juan Samuel	5' 11"	170lbs	R	R	28	178	35	434
Pedro Guerrero	5' 11"	176lbs	R	R	32	174	23	305
Howard Johnson	5' 11"	178lbs	R	B	36	164	41	180
Craig Biggio	5' 11"	185lbs	R	R	22	167	33	419
Chris Sabo	5' 11"	185lbs	R	R	25	153	25	273
Fernando Tatis	5' 11"	185lbs	R	R	34	160	21	267
Francisco Lindor	5' 11"	190lbs	R	B	38	183	25	383
Nomar Garciaparra	6' 0"	165lbs	R	R	30	209	22	450
Alan Trammell	6' 0"	165lbs	R	R	21	159	25	445
Roberto Alomar	6' 0"	184lbs	R	B	24	182	37	466
Barry Larkin	6' 0"	185lbs	R	R	33	154	36	426
Javier Baez	6' 0"	190lbs	R	R	34	176	21	155
Bo Bichette	6' 0"	190lbs	R	R	29	191	25	364
Marcus Semien	6' 0"	195lbs	R	R	26	163	25	398
Ian Kinsler	6' 0"	200lbs	R	R	32	158	30	398
David Wright	6' 0"	205lbs	R	R	30	196	34	324
Brandon Phillips	6' 0"	211lbs	R	R	30	187	32	433
Jonathan Villar	6' 0"	233lbs	R	B	24	176	40	200
Ryne Sandberg	6' 1"	175lbs	R	R	26	186	54	500
Alfonso Soriano	6' 1"	195lbs	R	R	39	209	41	402
Chase Utley	6' 1"	195lbs	R	L	31	161	23	408
Bobby Witt	6' 1"	200lbs	R	R	30	177	49	344
Trea Turner	6' 2"	185lbs	R	R	21	194	27	334
Mike Schmidt	6' 2"	195lbs	R	R	36	160	23	404
Trevor Story	6' 2"	213lbs	R	R	37	174	27	416
Hanley Ramirez	6' 2"	235lbs	R	R	29	212	51	392
Derek Jeter	6' 3"	195lbs	R	R	21	191	27	343
Corey Koskie	6' 3"	215lbs	R	L	26	155	27	306
Ian Desmond	6' 3"	217lbs	R	R	24	151	24	416
Alex Rodriguez	6' 3"	230lbs	R	R	36	172	28	262
Chipper Jones	6' 4"	210lbs	R	B	45	181	25	238

--#10

--List all players to ever hit 25 or more doubles, 5 or more triples, and 25 or more home runs in multiple seasons ordered by their length of carrer
SELECT

```
P.nameFirst || ' ' || P.nameLast AS Name,  
P.debut,  
P.finalGame,  
(P.finalGame - P.debut) || ' Years' AS CareerLength,  
CB."2B" AS Career2B,  
CB."3B" AS Career3B,  
CB."HR" AS CareerHR
```

FROM PLAYER P, BATTING B, CareerBatting CB --Uses season stats as criteria
and career stats for output

WHERE P.playerID = B.playerID AND --Joins the players to the batting stats

```
P.playerID = CB.playerID AND
```

```
B."2B" > 25 AND B."3B" > 5 AND B.HR > 25 --Filters players by  
requirements
```

GROUP BY P.playerID --Groups players who have multiple seasons meeting the
requirements

```
HAVING COUNT(*) > 1
```

```
ORDER BY (P.finalGame - P.debut) ASC, P.debut DESC; --Orders by length of  
career then debut date to get the longest careers first
```

Name	debut	finalGame	CareerLength	Career2B	Career3B	CareerHR
Shohei Ohtani	2018-03-29	2022-10-05	4 Years	129	29	171
Nolan Arenado	2013-04-28	2022-10-05	9 Years	364	33	325
Wally Berger	1930-04-15	1940-07-04	10 Years	299	59	242
Mike Trout	2011-07-08	2022-10-05	11 Years	310	52	368
Charlie Blackmon	2011-06-07	2022-09-27	11 Years	310	63	215
Hack Wilson	1923-09-29	1934-08-25	11 Years	266	67	244
George Bell	1981-04-09	1993-10-02	12 Years	308	34	265
Joe Gordon	1938-04-18	1950-09-30	12 Years	264	52	253
Dolph Camilli	1933-09-09	1945-09-23	12 Years	261	86	239
Bob Johnson	1933-04-12	1945-09-23	12 Years	396	95	288
Earl Averill	1929-04-16	1941-04-25	12 Years	401	128	238
Ryan Braun	2007-05-25	2020-09-27	13 Years	408	49	352
Nomar Garciaparra	1996-08-31	2009-10-04	13 Years	370	52	229
Hal Trosky	1933-09-11	1946-09-27	13 Years	331	58	228
Chick Hafey	1924-08-28	1937-09-30	13 Years	341	67	164
Brian Giles	1995-09-16	2009-06-18	14 Years	411	55	287
Will Clark	1986-04-08	2000-10-01	14 Years	440	47	284
Bobby Thomson	1946-09-09	1960-07-17	14 Years	267	74	264
Ken Williams	1915-07-14	1929-08-10	14 Years	285	77	196
Vladimir Guerrero	1996-09-19	2011-09-28	15 Years	477	46	449
Bernie Williams	1991-07-07	2006-10-01	15 Years	449	55	287
Joe Carter	1983-07-30	1998-09-28	15 Years	432	53	396
Jim Rice	1974-08-19	1989-08-03	15 Years	373	79	382
Johnny Callison	1958-09-09	1973-08-17	15 Years	321	89	226
Joe DiMaggio	1936-05-03	1951-09-30	15 Years	389	131	361
Jim Bottomley	1922-08-18	1937-09-16	15 Years	465	151	219
Andruw Jones	1996-08-15	2012-10-03	16 Years	383	36	434
Jeff Kent	1992-04-12	2008-09-27	16 Years	560	47	377
Ryne Sandberg	1981-09-02	1997-09-28	16 Years	403	76	282
Eddie Mathews	1952-04-15	1968-09-27	16 Years	354	72	512
Chuck Klein	1928-07-30	1944-06-11	16 Years	398	74	300
Lou Gehrig	1923-06-15	1939-04-30	16 Years	534	163	493
Dale Murphy	1976-09-13	1993-05-21	17 Years	350	39	398
Mike Schmidt	1972-09-12	1989-05-28	17 Years	408	59	548
Billy Williams	1959-08-06	1976-10-02	17 Years	434	88	426
Duke Snider	1947-04-17	1964-10-03	17 Years	358	85	407
Bill Nicholson	1936-06-13	1953-09-19	17 Years	272	60	235
Johnny Mize	1936-04-16	1953-09-26	17 Years	367	83	359
Hank Greenberg	1930-09-14	1947-09-18	17 Years	379	71	331
Steve Finley	1989-04-03	2007-06-03	18 Years	449	124	304
Ernie Banks	1953-09-17	1971-09-26	18 Years	407	90	512
Dwight Evans	1972-09-16	1991-10-06	19 Years	483	73	385
Frank Robinson	1956-04-17	1976-09-18	20 Years	528	72	586
Jimmie Foxx	1925-05-01	1945-09-23	20 Years	458	125	534
Al Simmons	1924-04-15	1944-07-01	20 Years	539	149	307
Barry Bonds	1986-05-30	2007-09-26	21 Years	601	77	762
Ted Williams	1939-04-20	1960-09-28	21 Years	525	71	521
Mel Ott	1926-04-27	1947-07-11	21 Years	488	72	511
Babe Ruth	1914-07-11	1935-05-30	21 Years	506	136	714
Dave Winfield	1973-06-19	1995-10-01	22 Years	540	88	465
Tony Perez	1964-07-26	1986-10-05	22 Years	505	79	379
Hank Aaron	1954-04-13	1976-10-03	22 Years	624	98	755
Willie Mays	1951-05-25	1973-09-09	22 Years	523	140	660
Stan Musial	1941-09-17	1963-09-29	22 Years	725	177	475
Rogers Hornsby	1915-09-10	1937-07-20	22 Years	541	169	301