

# The Best of Baseball Awards

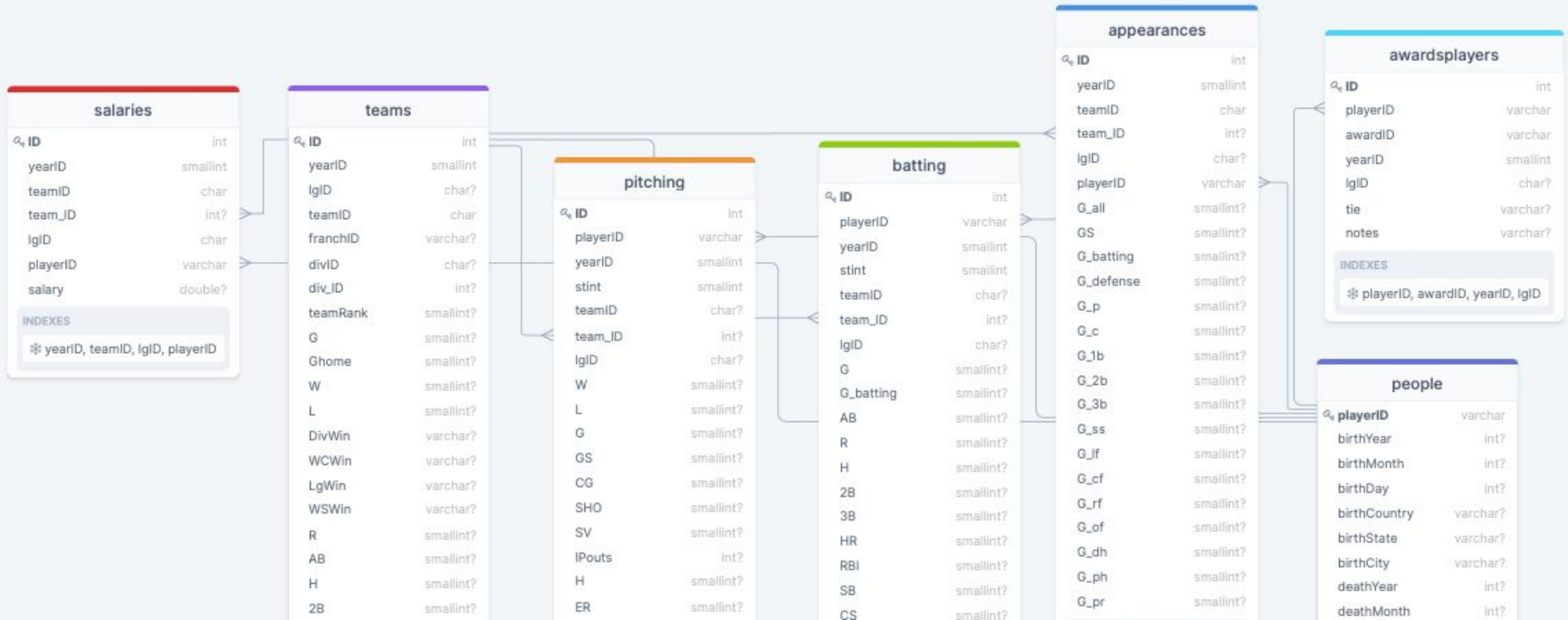


Analyzing data with SQL

**In this project, we will be looking at a massive baseball database containing information about players, teams, managers, salaries, and just about anything you might want to know about baseball. This dataset contains data from 2019 all the way back to 1871. Let's see what interesting facts we can learn from this database!**

---

For this project, I used the following tables related to each other from the entire database.





# Heaviest Hitters

This award goes to the team with the highest average weight of its batters on a given year and to the heaviest player in history.

```
1 SELECT ROUND(AVG(people.weight), 2) AS avg_weight,  
2     teams.name AS team,  
3     batting.yearid AS year  
4 FROM people  
5 INNER JOIN batting  
6     ON people.playerid = batting.playerid  
7 INNER JOIN teams  
8     ON batting.team_id = teams.id  
9 GROUP BY team, year  
10 ORDER BY avg_weight DESC  
11 LIMIT 5;
```

avg_weight	team	year
221.33	Chicago White Sox	2009
220.91	Chicago White Sox	2014
219.74	Chicago White Sox	2008
219.38	Philadelphia Phillies	2015
219.02	Chicago White Sox	2007

The heaviest team was  
**Chicago White Sox**  
with average weight of  
**221.33lb in 2009**

The heaviest player  
title goes to:  
**Walter Young**  
**320lb!**

```
1 SELECT namefirst AS "first name",  
2     namelast AS "last name",  
3     weight  
4 FROM people  
5 WHERE weight = (SELECT MAX(weight)  
6                 FROM people);
```

OR

```
1 SELECT namefirst AS "first name",  
2     namelast AS "last name",  
3     weight  
4 FROM people  
5 WHERE weight IS NOT NULL  
6 ORDER BY 3 DESC  
7 LIMIT 1;
```

first name	last name	weight
Walter	Young	320

# Biggest Spenders

This award goes to the team with the largest total salary of all players in a given year. Additional awards goes to the most expensive player.

**Biggest Spenders Award goes to  
New York Yankees!**  
**They spent almost  
232mln \$ in 2013!**

**On the most expensive players  
rank we have a tie. So the title  
goes to both:**

**Alex Rodriguez who earned  
33mln\$ twice, playing for  
New York Yankees in  
2009 and 2010.**  
**Clayton Kershaw who earned  
33mln\$, playing for  
Los Angeles Dodgers in 2016.**

```
1 WITH top_salary AS (  
2     SELECT *  
3     FROM salaries  
4     ORDER BY salary DESC  
5     LIMIT 5)  
6 SELECT top_salary.salary AS salary,  
7     people.namefirst AS first_name,  
8     people.namelast AS last_name,  
9     teams.name AS team,  
10    top_salary.yearid AS year  
11 FROM top_salary  
12 JOIN people  
13     ON top_salary.playerid = people.playerid  
14 JOIN teams  
15     ON top_salary.team_id = teams.id  
16     AND top_salary.yearid = teams.yearid;
```

salary	first_name	last_name	team	year
33000000	Clayton	Kershaw	Los Angeles Dodgers	2016
33000000	Alex	Rodriguez	New York Yankees	2009
33000000	Alex	Rodriguez	New York Yankees	2010
32571000	Clayton	Kershaw	Los Angeles Dodgers	2015
32000000	Alex	Rodriguez	New York Yankees	2011

```
1 SELECT SUM(salaries.salary) AS total_salary,  
2     teams.name AS team,  
3     salaries.yearid AS year  
4 FROM salaries  
5 INNER JOIN teams  
6     ON salaries.team_id = teams.id  
7     AND salaries.yearid = teams.yearid  
8 GROUP BY team, year  
9 ORDER BY total_salary DESC  
10 LIMIT 5;
```

total_salary	team	year
231978886	New York Yankees	2013
223362196	Los Angeles Dodgers	2013
222997792	New York Yankees	2016
221288380	Los Angeles Dodgers	2016
217014600	Los Angeles Dodgers	2014

# Most Bang For Their Buck In 21st Century

This award goes to the team that had the smallest "cost per win" in 21st century. Cost per win is determined by the total salary of the team divided by the number of wins in a given year.

And the winner is

...

**Florida Marlins**

They won **78** games in **2006** and spent less than **15 mln \$** which gives them around **188k \$** per win.

---

```
1 SELECT ROUND(SUM(salaries.salary) / teams.w) AS cost_per_win,  
2     teams.name AS team,  
3     teams.yearid AS year,  
4     SUM(salaries.salary) AS total_salary,  
5     teams.w AS wins  
6 FROM salaries  
7 INNER JOIN teams  
8   ON salaries.teamid = teams.teamid  
9   AND salaries.yearid = teams.yearid  
10 WHERE teams.yearid >= 2001  
11 GROUP BY team, wins, year  
12 ORDER BY cost_per_win  
13 LIMIT 5;
```



cost_per_win	team	year	total_salary	wins
188096	Florida Marlins	2006	14671500	78
259661	Florida Marlins	2008	21811500	84
283882	Minnesota Twins	2001	24130000	85
311587	Tampa Bay Devil Rays	2003	19630000	63
331478	Oakland Athletics	2001	33810750	102

# Granny Awards



```
1 SELECT extract(year from age(finalgame_date, birth_date)) AS age,
2     namefirst AS first_name,
3     namelast AS last_name,
4     birth_date,
5     finalgame_date,
6     teams.name AS team
7 FROM people
8 INNER JOIN appearances
9     ON people.playerid = appearances.playerid
10    AND extract(year from people.finalgame_date) = appearances.yearid
11 INNER JOIN teams
12     ON appearances.team_id = teams.id
13 WHERE age(finalgame_date, birth_date) IS NOT NULL
14 ORDER BY 1 DESC
15 LIMIT 5;
```

age	first_name	last_name	birth_date	finalgame_date	team
59	Satchel	Paige	1906-07-07	1965-09-25	Kansas City Athletics
58	Charley	O'Leary	1875-10-15	1934-09-30	St. Louis Browns
57	Nick	Altrock	1876-09-15	1933-10-01	Washington Senators
54	Jim	O'Rourke	1850-09-01	1904-09-22	New York Giants
54	Minnie	Minoso	1925-11-29	1980-10-05	Chicago White Sox

This award goes to the oldest player  
in history.

And the winner is:

**Satchel Paige**

who played his last game for

**Kansas City Athletics**

when he was **59** years old.

**Congratulations Grandpa!**



# Awards Factory

This award goes to the city where the players with the most total rewards together were born.

Players who were born in **New York** got the highest number of awards, together they have **134** awards!  
So "**The Awards Factory**" goes to **New York!!!**



```
1 SELECT p.birthcity,  
2     COUNT(a.playerid) AS total_awards  
3 FROM awardsplayers a  
4 JOIN people p  
5     ON a.playerid = p.playerid  
6 GROUP BY 1  
7 ORDER BY 2 DESC  
8 LIMIT 5;
```

birthcity	total_awards
New York	134
Los Angeles	130
Santo Domingo	109
Chicago	95
San Francisco	95



# Congratulations to all winners!



and

# Thank you for checking

# "The Best of Baseball Awards"!