

Data Report for Braves Trainee Position

Isaac Zelcer

SQL Questions:

1.

Bonds, Barry
Rodriguez, Alex
Pujols, Albert
Edmonds, Jim
Rolen, Scott
Beltran, Carlos
Sheffield, Gary
Helton, Todd
Berkman, Lance
Thome, Jim
Giambi, Jason
Jones, Andruw
Giles, Brian
Boone, Bret
Ramirez, Manny
Cameron, Mike
Abreu, Bobby
Soriano, Alfonso
Garciaparra, Nomar
Suzuki, Ichiro
Renteria, Edgar
Posada, Jorge
Guerrero, Vladimir
Koskie, Corey
Kent, Jeff
Jones, Chipper
Chavez, Eric
Delgado, Carlos
Jeter, Derek
Ordonez, Magglio
Kotsay, Mark
Gonzalez, Luis
Anderson, Garret
Walker, Larry
Bagwell, Jeff
Winn, Randy
Mora, Melvin
Tejada, Miguel
Sosa, Sammy
Rodriguez, Ivan
Nixon, Trot
Hunter, Torii

Finley, Steve
Vidro, Jose
Giles, Marcus
Valentin, Jose
Castillo, Luis
Lee, Derrek
Palmeiro, Rafael
Beltre, Adrian
Lopez, Javy
Mueller, Bill
Damon, Johnny
Polanco, Placido
Lowell, Mike
Jones, Jacque
Green, Shawn
Furcal, Rafael
Lofton, Kenny
Olerud, John
Durham, Ray
Alfonzo, Edgardo
Sexson, Richie
Kendall, Jason
Kennedy, Adam
Hidalgo, Richard
Stewart, Shannon
Lieberthal, Mike
Piazza, Mike
Hinske, Eric
Sweeney, Mike
Sanders, Reggie
Pierre, Juan
Klesko, Ryan
Kearns, Austin
Bradley, Milton
Mientkiewicz, Doug
Floyd, Cliff
Cabrera, Orlando
Gonzalez, Alex
Wilkerson, Brad
Pierzynski, A.J.
Lee, Carlos
Eckstein, David
Varitek, Jason
Dunn, Adam
Thomas, Frank
Salmon, Tim
Cruz, Jose
Wilson, Preston
Wells, Vernon

Loretta, Mark
Ventura, Robin
Huff, Aubrey
Erstad, Darin
Ensberg, Morgan
Williams, Bernie
Alou, Moises
Matthews, Gary
Burrell, Pat
Payton, Jay
Blalock, Hank
Spivey, Junior
Ortiz, David
Vizquel, Omar
Hernandez, Ramon
Glaus, Troy
Lee, Travis
Jenkins, Geoff
DaVanon, Jeff
Burks, Ellis
Bellhorn, Mark
Byrd, Marlon
Hatteberg, Scott
Podsednik, Scott
Gerut, Jody
Bell, David
Matos, Luis
Hernandez, Jose
Mondesi, Raul
Guillen, Jose

2.

Moylan, Peter|0|0|0|-0.4
Sanchez, Anibal|1|1|0|2.4
Socolovich, Miguel|0|0|0|0.1
Venters, Jonny|0|0|0|0.3
Ravin, Josh|0|0|0|-0.1
Culberson, Charlie|0|0|0|0
Vizcaino, Arodys|0|0|0|0.3
Teheran, Julio|0|0|0|0.7
Ramirez, Jose|0|0|0|-0.1
Biddle, Jesse|0|0|0|0.2
Foltynewicz, Mike|1|1|1|3.9
Sims, Lucas|0|0|0|-0.2
Jackson, Luke|0|0|0|0.3
Wisler, Matt|0|0|0|0
Fried, Max|0|0|0|0.4
Toussaint, Touki|0|0|0|0.4
Gohara, Luiz|0|0|0|0
Wright, Kyle|0|0|0|-0.2
Allard, Kolby|0|0|0|-0.3
Soroka, Mike|0|0|0|0.6
Wilson, Bryse|0|0|0|0.1

McCarthy, Brandon|0|0|0|0.2
Freeman, Sam|0|0|0|0.3
Brach, Brad|0|0|0|0.3
Brothers, Rex|0|0|0|0
Gausman, Kevin|0|0|0|0.9
Winkler, Dan|1|0|0|1.2
Whitley, Chase|0|0|0|-0.1
McCreery, Adam|0|0|0|0
Phillips, Evan|0|0|0|-0.2
Minter, A.J.|1|0|0|1.4
Parsons, Wes|0|0|0|-0.1
Newcomb, Sean|1|0|0|1.9
Carle, Shane|0|0|0|0.6
Sobotka, Chad|0|0|0|0

3.

Luke Jackson data:

44|16|24|18

Data Wrangling Question:

Differences between SQL database and the API:

The key differences that I found between the PitchByPitch database and the StatsAPI were that the API contained the total stats for each player on the braves in 2018. However, the database only contained 3 pitchers (Jonny Venters, Luke Jackson, and Shane Carle). The database and the API also had some differences in total values. For example, when I queried data to find the number of doubles Jonny Venters gave up, I found the database has the value of 6 doubles while the API claimed his total was only 3. These differences allow us to come up with some ideas for the auditing system. We create queries to check total stat lines for each player to ensure that the API and database match.

Potential Audit System:

First, I would use the API to check to make sure that the names in the PitchByPitch table are in the players in the Stats API. Next, we would aggregate the data in the PitchByPitch Table to replicate the API data. From here we would be able to compare the data and find any issues/errors in the system.