Requirements

If you’re unfamiliar with baseball, you can find a quick summary of the objectives of the game [here](https://www.baseball-rules.com/basicbb.htm), though you don’t need to understand much about baseball to complete this project. You need to write a Python or R script that does the following:

* Connects to an SQL database file and queries for all players who have played at least 50 games and are still active.  Use the “finalGame” field from the “People” table to determine if a player is active. Retrieve weight, throws, bats, throws, all birth-related and all name-related columns from the “People” table and retrieve all columns from the “Batting” table.
* Converts this data into either an R data frame or a pandas data frame.
* Adds a calculated column with the player’s age and a calculated column with each player’s first and last name concatenated.
* Once the calculated columns are added, drops the other columns related to birth date and name.
* Deletes any rows with missing values
* Answers the following questions:  
  + Which active player had the most runs batted in (“RBI” from the Batting table) from 2015-2018?
  + How many double plays did Albert Pujols ground into (“GIDP” from Batting table) in 2016?
* Creates the following plots:  
  + A histogram of triples (3B) per year.
  + Create a scatter plot relating triples (3B) and steals (SB).
* Comes up at least three additional questions about the data and answers them. At least one should be a question about the relationship between two variables, e.g., triples and steals, as above.