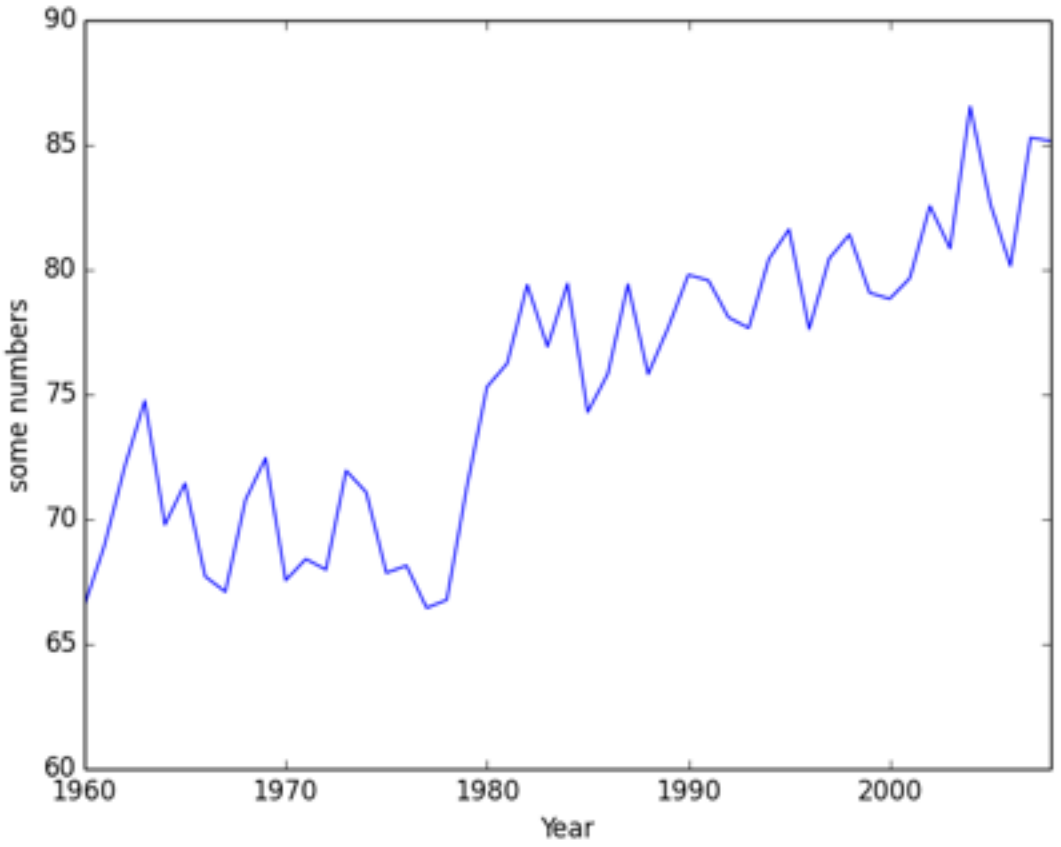
Domain:

The database consists of three tables: Players, Seasons, and Games. The Players table contains the players names and some basic information about them. Each player is given a unique player id. The Seasons table contains statistics for each season a player played in the NFL. Each row is linked to a player through the player id. Each row has a unique combination of player id, year, and team, since some players played for multiple teams in one season. The games table has statistics for each game a player played in starting in 1995 and is linked to the player table through the player id. Each row has a unique combination of player id, year, and week. The data was downloaded from pro-football-reference.com. The original data contained player names in the Seasons table, but these were dropped because they were unnecessary and wasted space. The large sample size of data allows users to look up any specific statistics as well as analyze trends across a long period of time.

Database Design:

Applications:

I wrote a short Python function that queries the database for QB's that threw a minimum number of passes each season and then called a stored procedures on the database to calculate their passer rating. These numbers were averaged together for each year from 1960 to 2008. The Python script is included in the .zip file. I then used a separate Python script to graph the results as seen below.



Future work:

I would like to continue working on programs similar to the Python function above. There is very little information available that can be viewed like the graph that I generated. This would probably work best as a web application.