COMP 310 — SQL Homework 2 — Ch 1.4

Develop SQL queries for the follow. Hint: As is the case with all programming, it is very helpful to build and test complex queries from smaller parts. So, do not hesitate to start with quires that return part of what you want and build from there. This can mean excluding columns or artificially restricting rows (i.e. only one year).

We'll focus on selecting a small number of columns for these queries so that you can focus on the join logic. Unless given a more restricted date range, restrict the years to 1985-2005.

1. Let's combine pitching stats with player information. Write a query that gets, for each pitcher and each season in the requested time frame, the first name, last name, and birth year from the people table along with teamId, stint, year, league, wins, loses, and games played from pitching. Sort the data by wins (most to least).

2. Now again with batting, get the same biographical information from people and combine it with team, stint, year, league, hits and at bats from batting. Sort the data by hits (most to least)

3. Let's expand the pitching query above (copy and modify that query) to include team name (rather than id), rank, wins, losses, and games played from teams. (Imagine you want to compare pitcher performance (wins,loses, etc,) to team performance (wins, losses)). Rename columns as needed to avoid confusing player and team stats. (Hint: This query should return the same number of rows as its pitching counterpart above.)

4. Now batting, but do hits and at bats from teams rather than wins, loses, and games played. Don't forget to rename columns as needed. (Hint: This query should return the same number of rows as its batting counterpart above.)

5. Now let's try to get the roster and some stats for an entire time. Restrict this query to the 2001 Chicago Cubs. Write a query that pulls and combines our restricted pitching (wins, loses, games) and batting stats (hits and at bats) along with playerID, teamID, yearID, stint league (from both batting and pitching). Remember, most batters will not have pitching stats but pitcher will have batting stats. The goal is to have one row per player+stint. (Hint: this is an outer join situation).

6. Now expand the above query to get player and team names rather than ids, and while you're at it, include team stats (wins, loses, games, hits, and at bats) like we did previously.

7. Now lets return to our fourth query that lists player and team batting stats for 1985-2005. Let's include post season batting stats (H and AB) as well as post season games for each batter. These stats can be found in the batting_post table. Note, only players with post-season appearances will have stats. (Yes, more outer joining). Sort the data by post season games played (make sure NULLs come after non-NULL values).