

Homework 10: Solutions

36-350, Fall 2011

1. SOLUTION

- (a) The tables in the database can be found with

```
> dbListTables(con)
 [1] "AllstarFull"           "Appearances"           "AwardsManagers"
 [4] "AwardsPlayers"        "AwardsShareManagers"  "AwardsSharePlayers"
 [7] "Batting"              "BattingPost"          "Fielding"
[10] "FieldingOF"           "FieldingPost"         "HallofFame"
[13] "Managers"             "ManagersHalf"         "Master"
[16] "Pitching"             "PitchingPost"         "Salaries"
[19] "Schools"              "SchoolsPlayers"       "SeriesPost"
[22] "Teams"                "TeamsFranchises"      "TeamsHalf"
[25] "sqlite_sequence"     "xref_stats"
```

Now we know that the table we will be working with for the rest of the problem is **Salaries**.

- (b) After performing the query (see *R* file for solutions)

```
> head(team.payrolls)
  teamID  payroll
1   NYA 206333389
2   BOS 162447333
3   CHN 146609000
4   PHI 141928379
5   NYN 134422942
6   DET 122864928
```

The teams with the highest payrolls are the New York Yankees, Boston Red Sox, Chicago Cubs, Philadelphia Phillies, New York Mets, and Detroit Tigers.

- (c) See *R* file for code.

2. SOLUTION

- (a) The figures are plotted in Figure 1.
- (b) The code is included in the additional *R* file and the figure is plotted in Figure 2. There does not seem to be a strong relationship between win percentage and payroll.
- (c) The plot is included in Figure 3. The plot suggests that the relationship between winning percentage and inflation adjusted team salary has been becoming stronger over time.

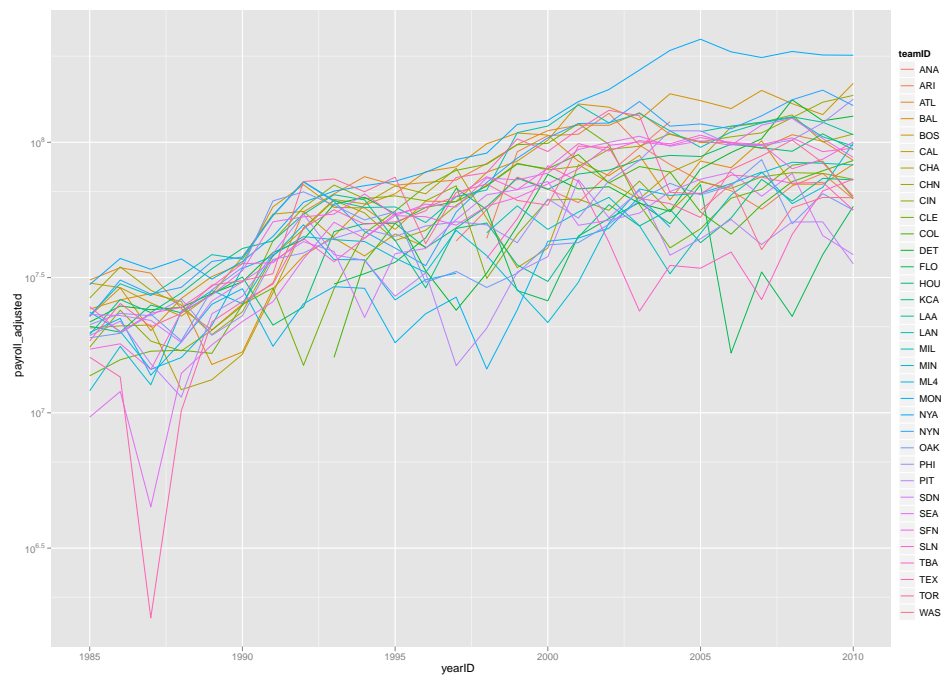
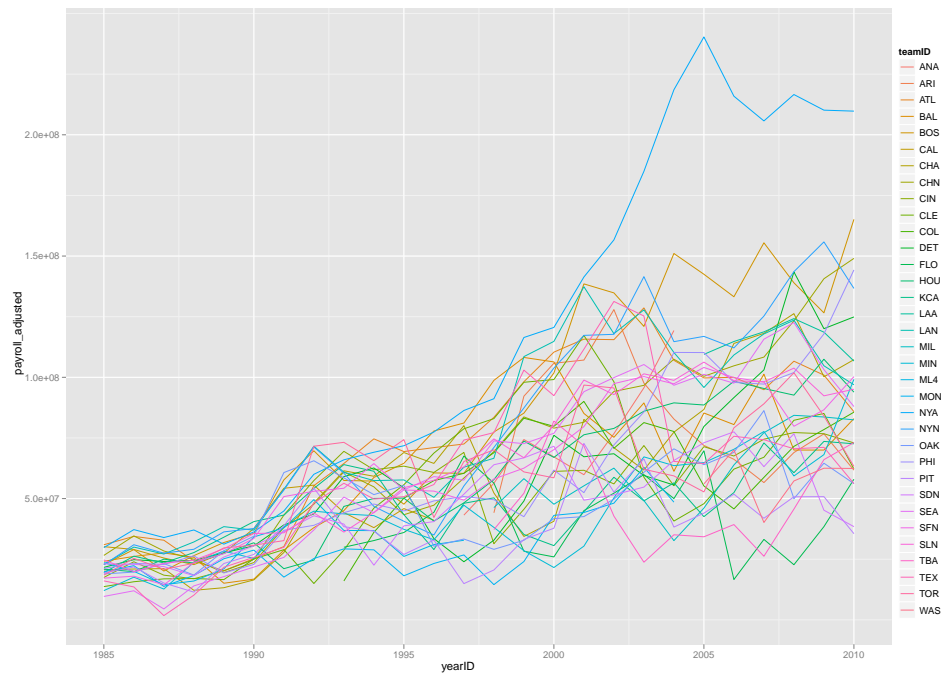


Figure 1: Top: inflation adjusted team payrolls over time. Bottom: log of inflation adjusted team payrolls.

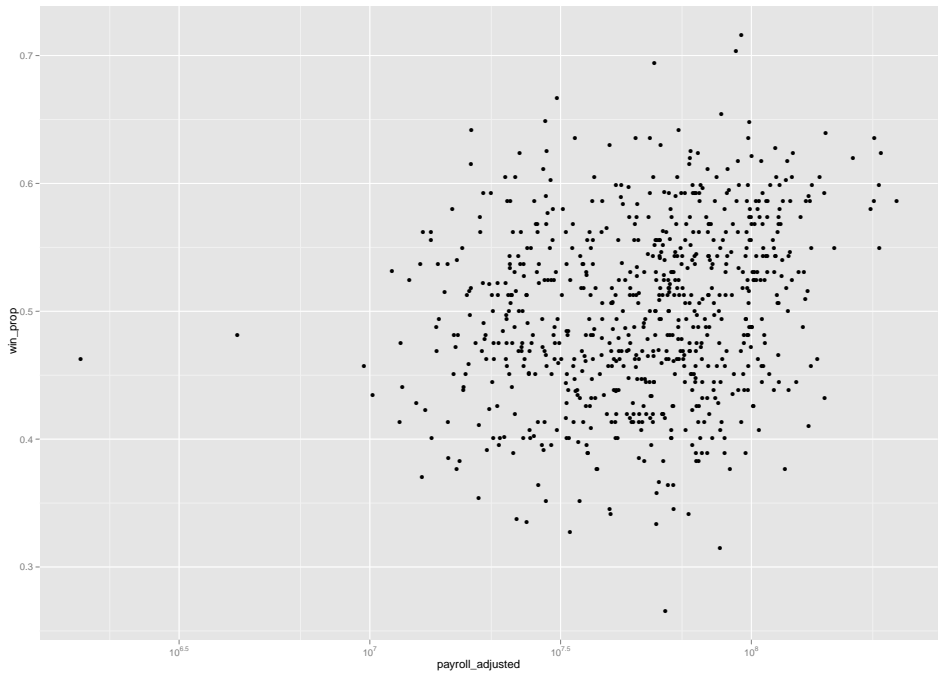


Figure 2: Scatterplot of inflation adjusted payrolls versus proportion of games won

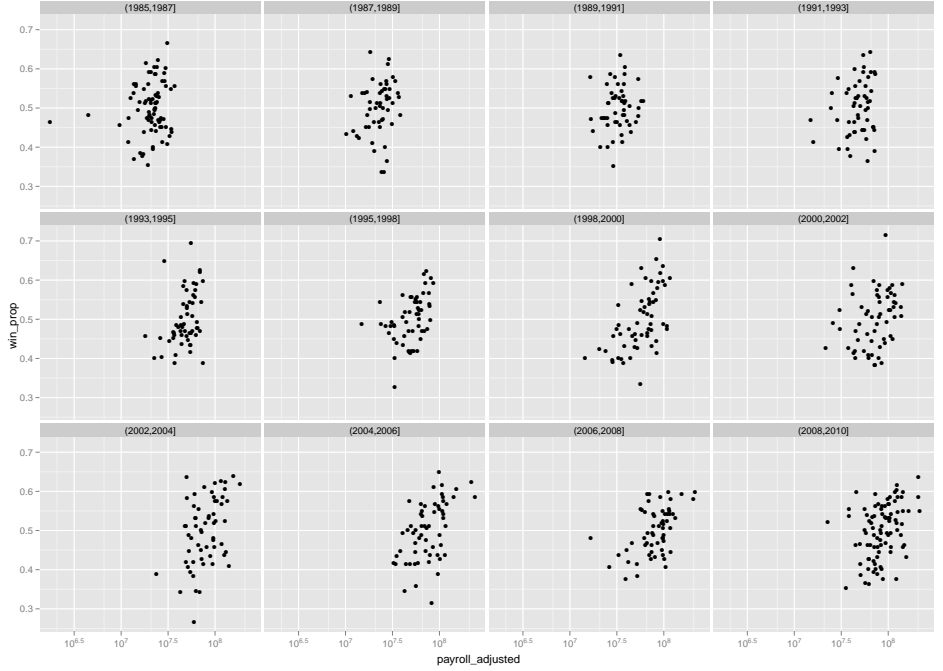


Figure 3: Scatterplot of inflation adjusted payrolls versus proportion of games won