BBLeagues
Basketball Database System

Cory Hamelehle, Adam Jenkins, Kok Siew

CSU, Chico CINS 370

2016-04-23

outline

ER Diagram

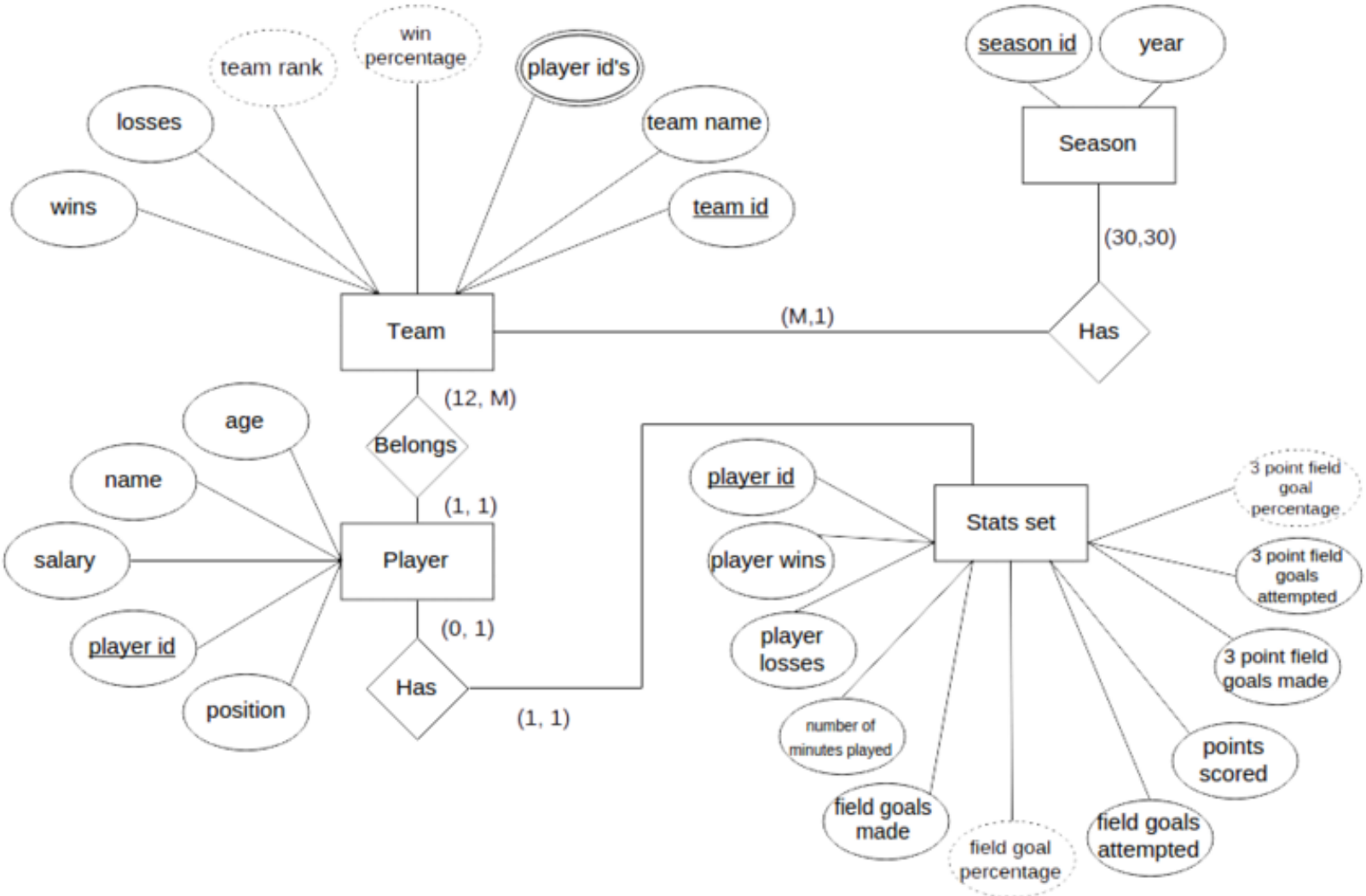
SQL Tables

Making .csv files

Plots

References

ER Diagram


win percentage
team rank
player id's
losses
team name
wins
team id
season id
year
Season
(30,30)
Team
(M,1)
Has
(12, M)
Belongs
age
name
salary
player id
position
(1, 1)
Player
(0, 1)
Has
(1, 1)
player id
player wins
player losses
number of minutes played
field goals made
field goal percentage
Stats set
3 point field goal percentage
3 point field goals attempted
3 point field goals made
points scored
field goals attempted

outline

ER Diagram

SQL Tables

Making .csv files

Plots

References

SQL Tables

Table definitions for our four tables

```
SQL> describe season
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------
 SEASON_ID                                 NOT NULL NUMBER(3)
 YEAR                                               CHAR(7)

SQL>
```

SQL Tables

```
SQL> describe player;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------
 PLAYER_ID                                 NOT NULL NUMBER(4)
 NAME                                      NOT NULL CHAR(30)
 AGE                                                NUMBER(2)
 SALARY                                             VARCHAR2(11)
 POSITION                                  NOT NULL CHAR(9)

SQL>
```

SQL Tables

```
SQL> describe team
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------
 TEAM_ID                                   NOT NULL NUMBER(2)
 TEAM_NAME                                 NOT NULL VARCHAR2(30)
 RECORD                                    NOT NULL CHAR(6)
 WINS                                               NUMBER(5)
 LOSSES                                             NUMBER(5)
 WIN_PERCENTAGE                                     NUMBER(3,2)

SQL>
```

SQL Tables

```
SQL> describe stats_set
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------
 PLAYER_ID                                 NOT NULL NUMBER(4)
 PLAYER                                    NOT NULL CHAR(30)
 POSITION                                  NOT NULL CHAR(9)
 AGE                                               NUMBER(3)
 GAMES                                             NUMBER(3)
 MINUTES_PLAYED                                    NUMBER(5)
 PER                                               NUMBER(20,10)
 TS_PCT                                            NUMBER(20,10)
 THREE_PAR                                         NUMBER(20,10)
 FTR                                               NUMBER(20,10)

SQL>
```

outline

ER Diagram

SQL Tables

Making .csv files

Plots

References

Making .csv files

In order to make our own .csv files from the data, we used a combination of spool and PL/SQL. Spool is a tool in Oracle that writes everything printed to standard out to the specified file and the PL/SQL code below allowed us to print the ages of players and their salaries separated by comma.

```
declare
age number(3);
salary varchar(11);
begin
        for i in 1..492 loop
                select age into age from player where player_id = i;
                select salary into salary from player where player_id = i;
                dbms_output.put_line(age||', '||salary);
        end loop;
end;
/
```

Running spool and the PL/SQL script

```
SQL> spool age_salary.csv
SQL> @age_salary.sql
```

```
23, 4400000
25,
28,
28,
32, 2100000
28,
28,
24,
23,
27, 5464000
29,
21,
19, 1749840
29, 5219169
26,
22, 4204200
25, 2616975

PL/SQL procedure successfully completed.

SQL>
```

View of .csv file created

```
 1 Age, Salary
 2 24,
 3 20, 1404600
 4 21,
 5 28,
 6 29,
 7 26, 4389607
 8 23,
 9 26, 1100602
10 29, 19689000
11 25,
12 33, 5168539
13 24, 8042895
14 32,
15 36, 5000000
16 32, 4000000
17 21, 1142880
18 26, 8500000
19 20, 1953960
20 30, 22875000
@
"age_salary.csv" 501L, 38607C written
```

outline

ER Diagram

SQL Tables

Making .csv files

Plots

References

Plots

Now that the .csv files have been created, we can then plot them using R

```
nba <- read.csv("/home/adam/Documents/age_salary.csv")
nba <- na.omit(nba)
library(ggplot2)
p <- qplot(Age, Salary, data = nba,
    main = "Age v Salary")
```

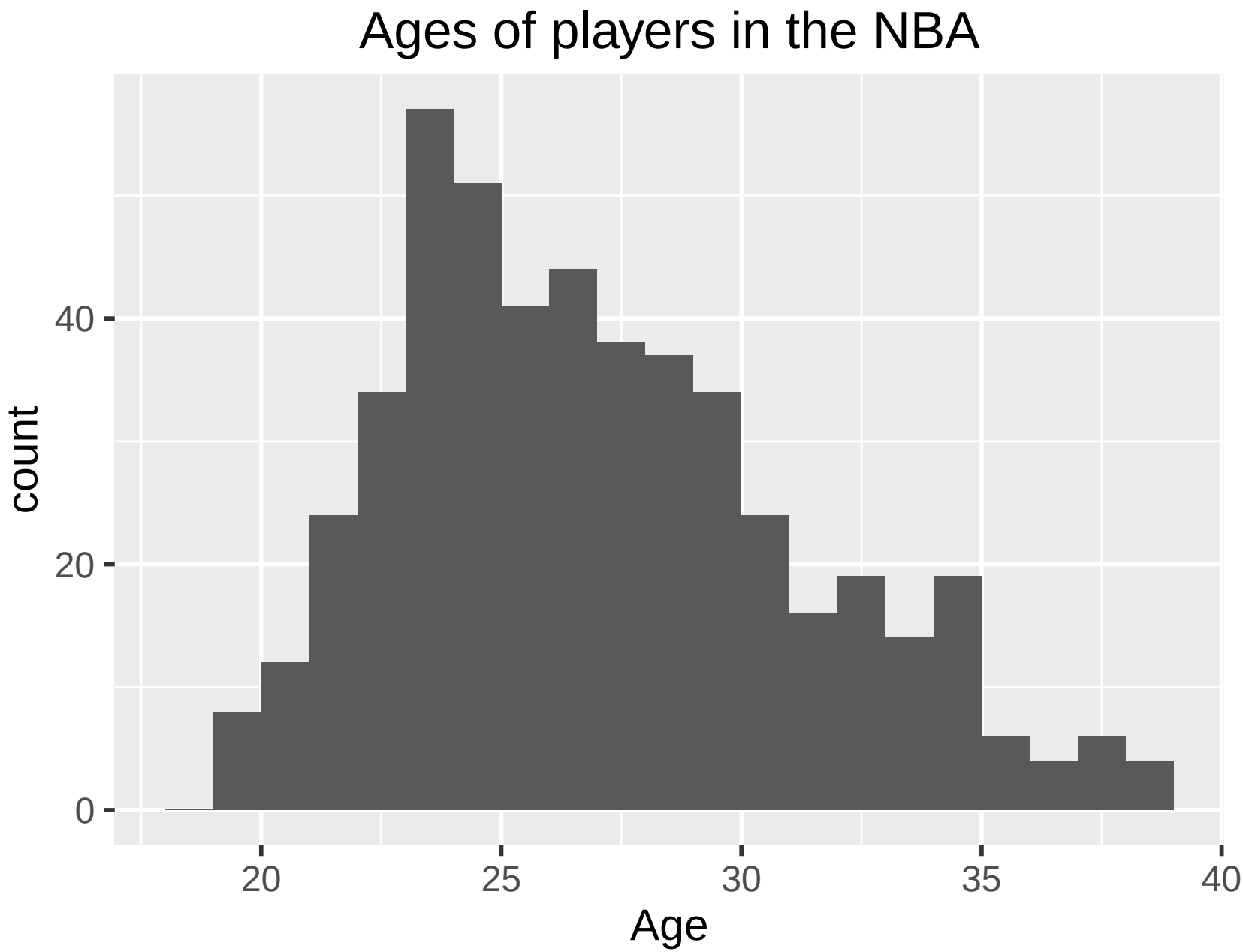
Plots


Age v Salary
Salary
2.5e+07
2.0e+07
1.5e+07
1.0e+07
5.0e+06
0.0e+00
20
25
30
35
Age

Plots

```
nba <- read.csv("/home/adam/Documents/age.csv")
p <- qplot(AGE, data = nba, geom = "histogram",
    main = "Ages of players in the NBA",
    xlab = "Age", binwidth = 1)
```

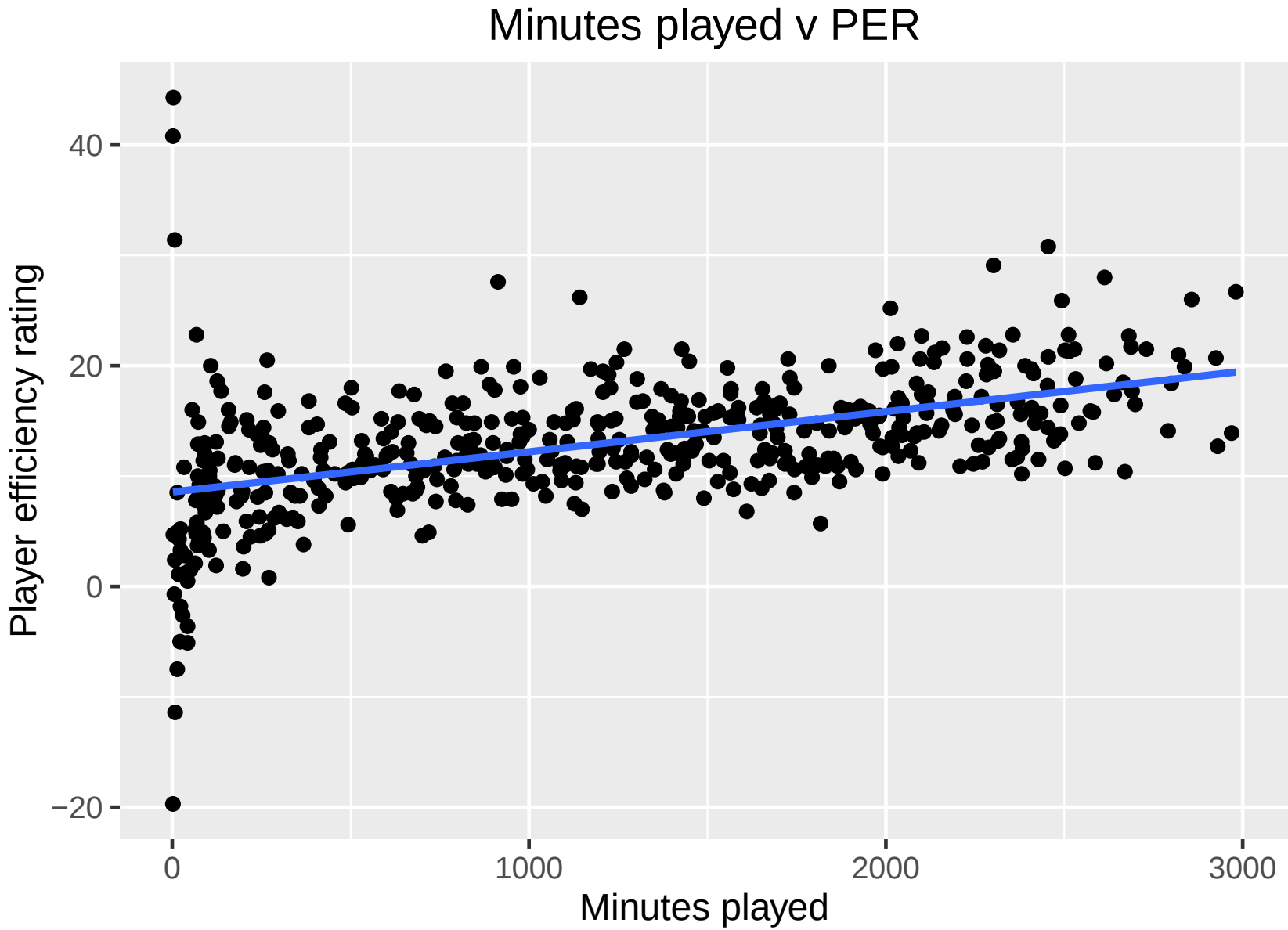
Plots



Plots

```
nba <- read.csv("/home/adam/Documents/min_per.csv")
p <- qplot(min, per, data = nba,
    main = "Minutes played v PER",
    xlab = "Minutes played",
    ylab = "Player efficiency rating") +
    geom_smooth(method = "lm", se = FALSE)
```

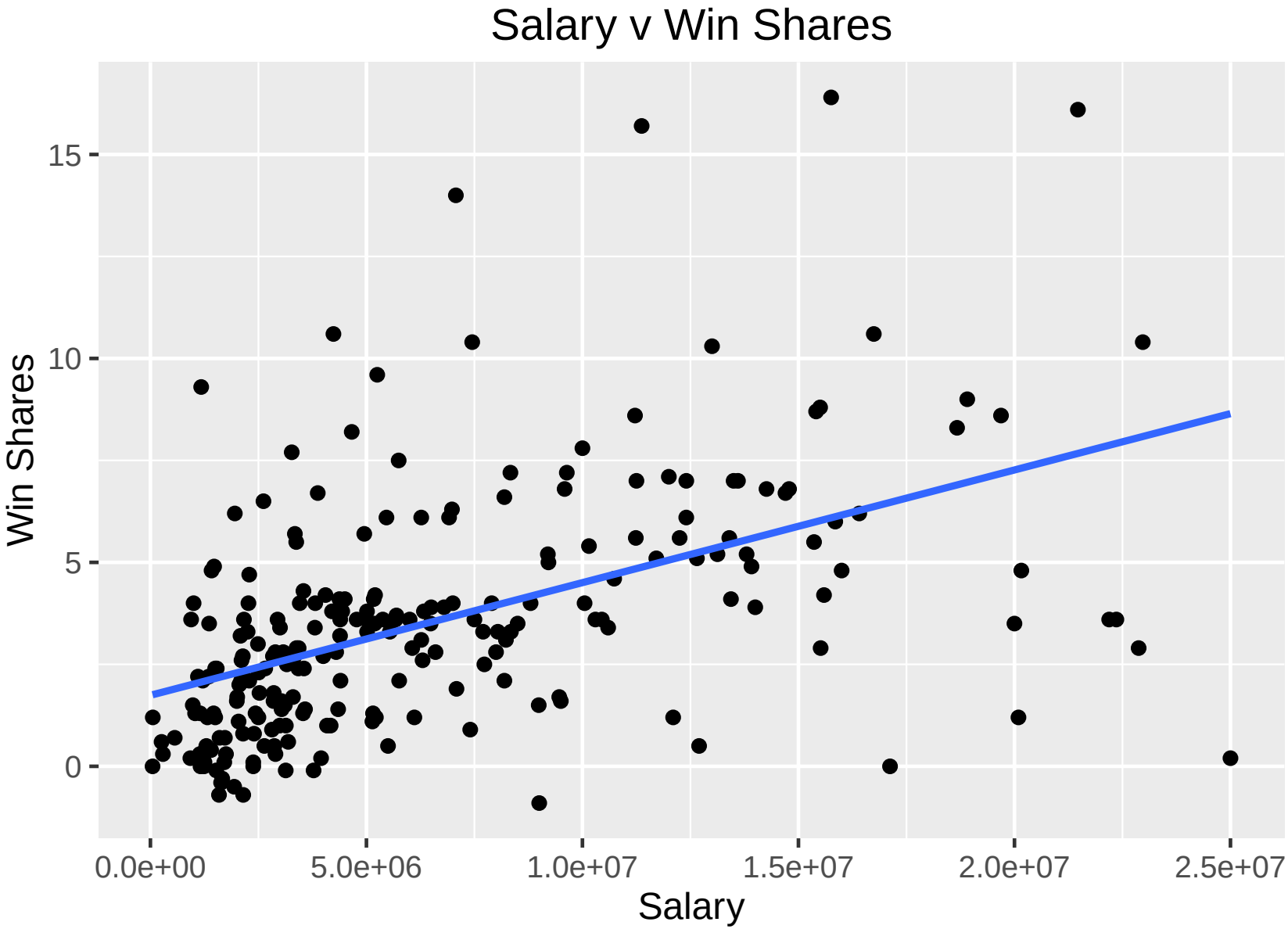
Plots


Minutes played v PER
Player efficiency rating
40
20
0
−20
0
1000
2000
3000
Minutes played

Plots

```
nba <- read.csv("/home/adam/Documents/salary_ws.csv")
nba <- na.omit(nba)
p <- qplot(salary, ws, data = nba,
    main = "Salary v Win Shares",
    xlab = "Salary", ylab = "Win Shares") +
    geom_smooth(method = "lm", se = FALSE)
```

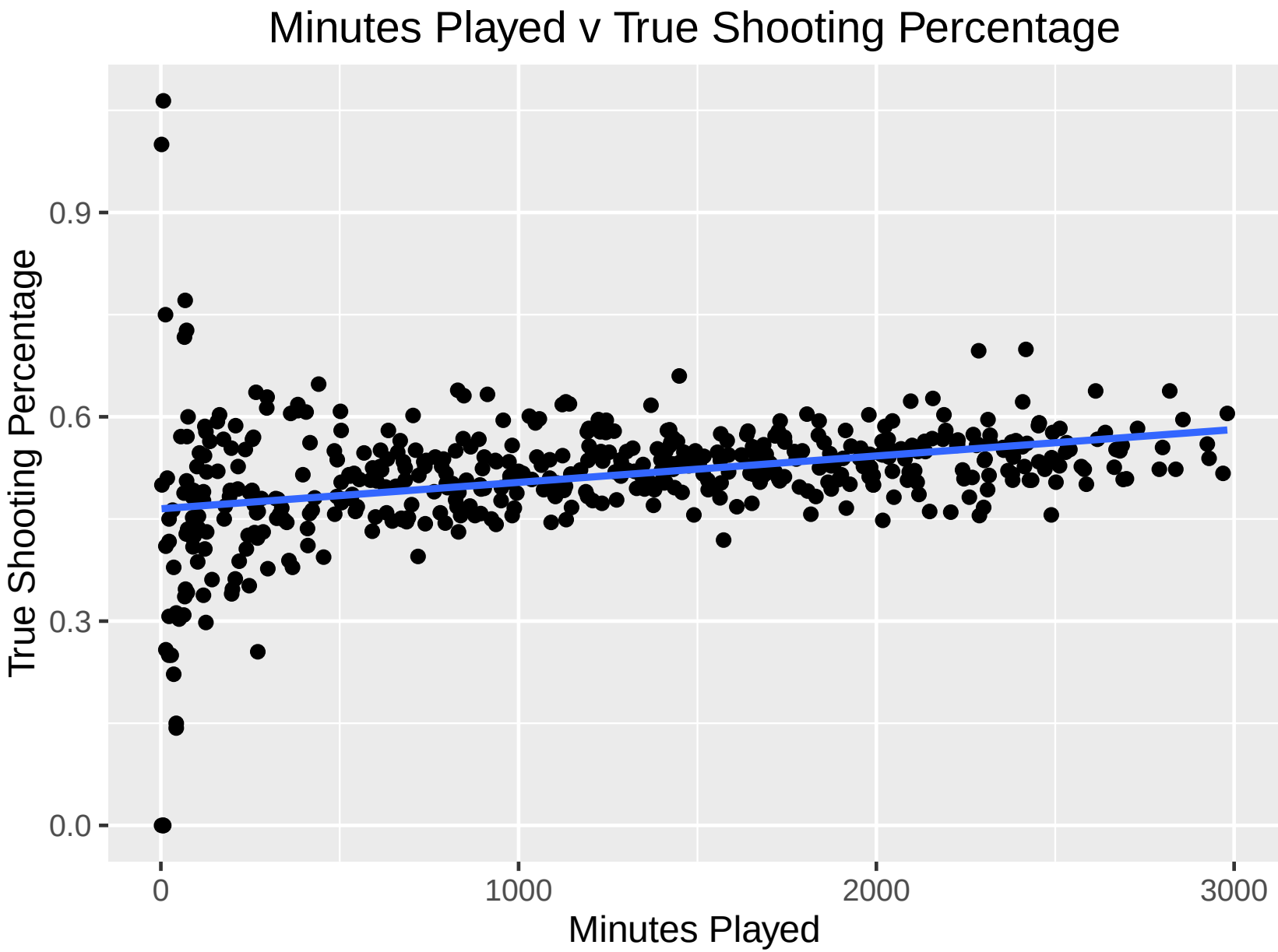
Plots


Salary v Win Shares
Win Shares
15
10
5
0
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07
2.5e+07
Salary

Plots

```
nba <- read.csv("/home/adam/Documents/min_ts_pct.csv")
nba <- na.omit(nba)
p <- qplot(min, ts_pct, data = nba,
    main = "Minutes Played v True Shooting Percentage",
    xlab = "Minutes Played",
    ylab = "True Shooting Percentage") +
    geom_smooth(method = "lm", se = FALSE)
```

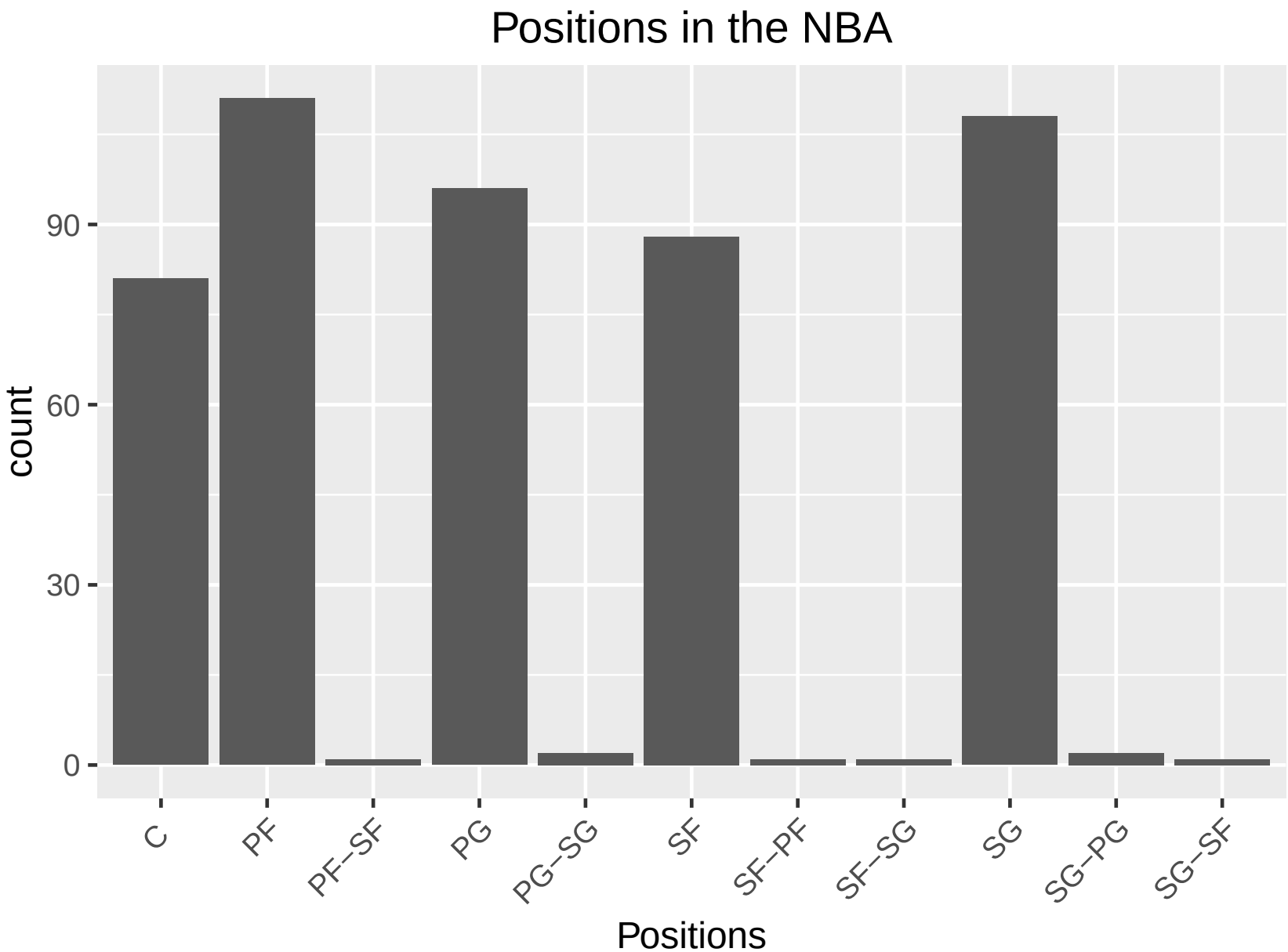
Plots


Minutes Played v True Shooting Percentage
True Shooting Percentage
0.0
0.3
0.6
0.9
0
1000
2000
3000
Minutes Played

Plots

```
nba <- read.csv("/home/adam/Documents/positions.csv")
p <- qplot(nba$Position, geom = "bar",
    main = "Positions in the NBA",
    xlab = "Positions") + theme(axis.text.x =
    element_text(angle = 45, hjust = 1))
```

Plots


Positions in the NBA
count
90
60
30
0
C
PF
PF–SF
PG
PG–SG
SF
SF–PF
SF–SG
SG
SG–PG
SG–SF
Positions

outline

ER Diagram

SQL Tables

Making .csv files

Plots

References

References

All data gathered from www.basketball-reference.com.
Database built using Oracle 11.2.0.2.0.