NBA teams revenue analysis

2023-03-02

Purpose: To investigate the factors that will affect NBA teams revenue using regression analysis.

Factors: team wins, revenue, market size, number of championships

links: 1. NBA official 2. Forbes

```
#import library and create dataset
library(tidyr)
library(jtools)
data <- read.delim("/Users/chingshawn/Desktop/NBA.txt", header=FALSE)
names(data) <- c("team","revenue", "wins", "salary_expenses", "market_size", "num_champ")
```

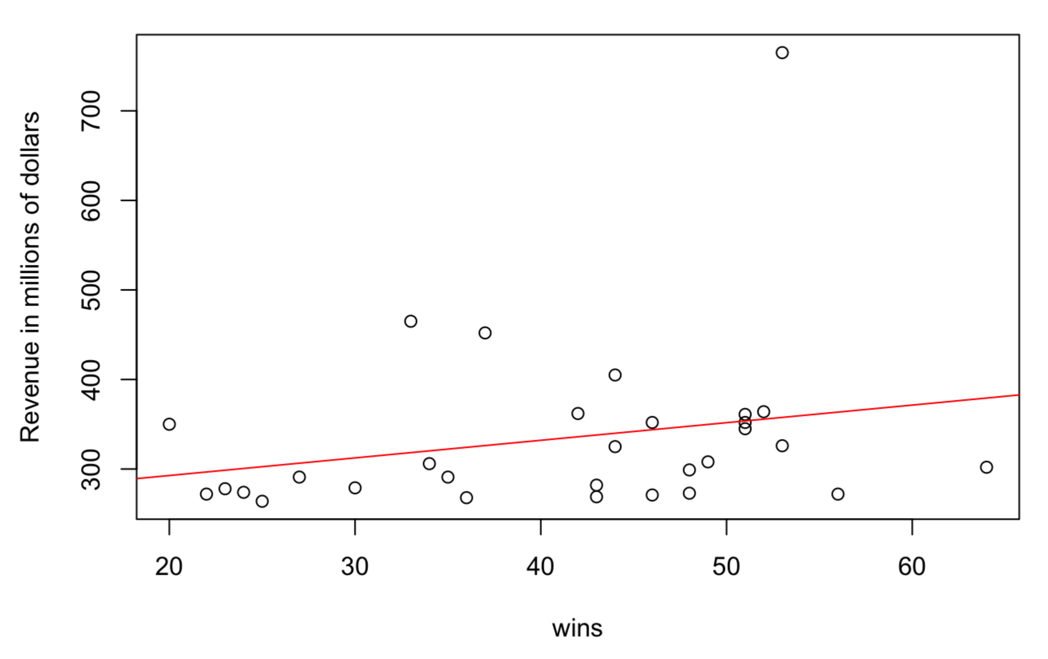
```
# make sure all the values are numeric
data$revenue <- as.numeric(data$revenue)
data$wins <- as.numeric(data$wins)
data$salary_expenses <- as.numeric(data$salary_expenses)
data$market_size <- as.numeric(data$market_size)
data$num_champ <- as.numeric(data$num_champ)
```

```
# remove NA data and show the table
data <- na.omit(data)
data
```

```
##                       team revenue wins salary_expenses market_size num_champ
## 1            Atlanta Hawks     282   43            10.0        4.48         1
## 2           Boston Celtics     361   51            12.6       14.75        17
## 3            Brooklyn Nets     405   44            11.0       11.61         0
## 4        Charlotte Hornets     269   43             7.5        3.19         0
## 5            Chicago Bulls     352   46             9.7       14.37         6
## 6      Cleveland Cavaliers     325   44             8.6        6.52         1
## 7         Dallas Mavericks     364   52            11.4       12.73         1
## 8           Denver Nuggets     273   48            10.7        3.90         0
## 9          Detroit Pistons     278   23             7.0        4.73         3
## 10   Golden State Warriors     765   53            13.6       30.24         7
## 11         Houston Rockets     350   20             3.8       11.05         2
## 12          Indiana Pacers     264   25             7.6        2.96         0
## 13    Los Angeles Clippers     362   42            12.8       13.18         0
## 14      Los Angeles Lakers     465   33            12.0       28.72        17
## 15       Memphis Grizzlies     272   56             8.3        3.57         0
## 16              Miami Heat     326   53             9.9       10.20         3
## 17         Milwaukee Bucks     352   51            12.8        6.80         2
## 18  Minnesota Timberwolves     271   46             9.5        3.51         0
## 19    New Orleans Pelicans     268   36             9.8        2.59         0
## 20         New York Knicks     452   37            10.5       28.21         2
## 21   Oklahoma City Thunder     274   24             6.5        4.08         1
## 22           Orlando Magic     272   22             6.6        4.02         0
## 23      Philadelphia 76ers     345   51            10.0       10.29         3
## 24            Phoenix Suns     302   64            11.7        7.11         0
## 25  Portland Trail Blazers     291   27            10.0        5.09         1
## 26        Sacramento Kings     279   30             8.8        5.35         1
## 27       San Antonio Spurs     306   34             4.9        5.27         5
## 28         Toronto Raptors     299   48             9.6        7.22         1
## 29               Utah Jazz     308   49             6.6        4.96         0
## 30      Washington Wizards     291   35             9.1        6.44         1
```

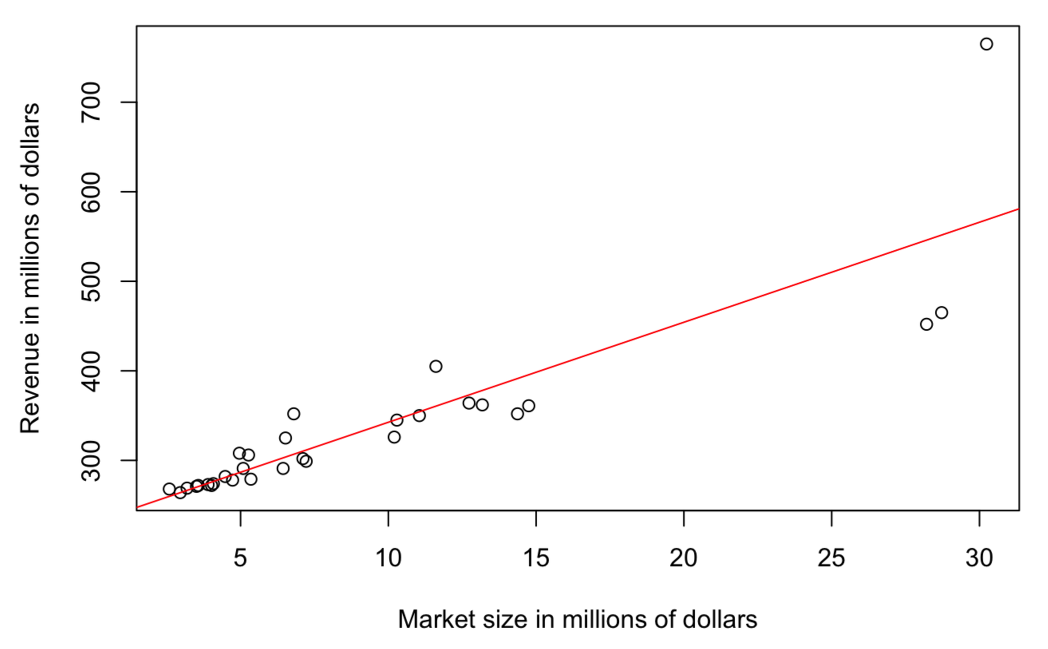
#Regression Analysis: Want to show affect of each factor on revenue

```
#Regression line (wins on Revenue)
reg_wins <- lm(revenue~wins, data=data)
plot(data$wins, data$revenue, xlab = "wins", ylab = "Revenue in millions of dollars")
abline(lm(revenue~wins, data=data),col="red")
```

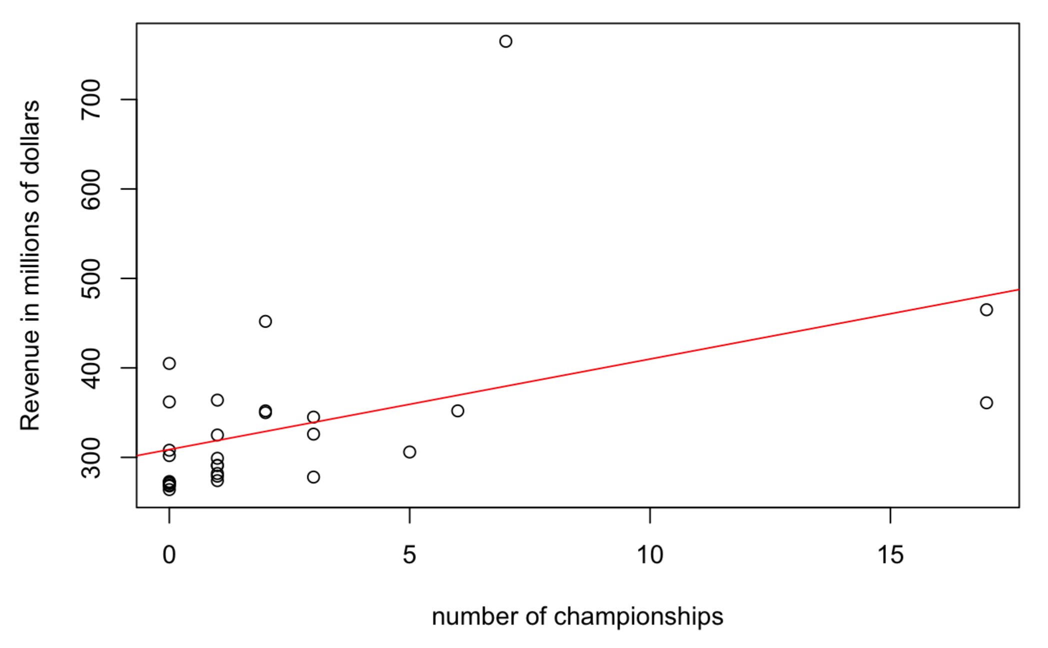



```
#Regression line (salary expenses on revenue)
reg_salary <- lm(revenue~salary_expenses, data=data)
plot(data$salary_expenses, data$revenue, xlab = "salary expense in millions of dollars", ylab = "Revenue in milli
ons of dollars")
abline(lm(revenue~salary_expenses, data=data),col="red")
```


```
#Regression line (market size on revenue)
reg_market <- lm(revenue~market_size, data=data)
plot(data$market_size, data$revenue,xlab="Market size in millions of dollars", ylab="Revenue in millions of dolla
rs")
abline(lm(revenue~market_size, data=data),col="red")
```




```
#Regression line (number of championships on revenue)
reg_pop <- lm(revenue~num_champ, data=data)
plot(data$num_champ, data$revenue, xlab = "number of championships", ylab = "Revenue in millions of dollars")
abline(lm(revenue~num_champ, data=data),col="red")
```




```
reg_all <- lm(revenue~wins+market_size+salary_expenses+num_champ, data=data)
summ(reg_all)
```

```
## MODEL INFO:
## Observations: 30
## Dependent Variable: revenue
## Type: OLS linear regression
##
## MODEL FIT:
## F(4,25) = 23.20, p = 0.00
## R² = 0.79
## Adj. R² = 0.75
##
## Standard errors: OLS
## -----------------------------------------------------
##                           Est.    S.E.   t val.      p
## --------------------- -------- ------- -------- ------
## (Intercept)             192.67   39.29     4.90   0.00
## wins                      0.76    0.99     0.76   0.45
## market_size              11.90    1.66     7.17   0.00
## salary_expenses           0.83    5.55     0.15   0.88
## num_champ                -2.95    2.62    -1.13   0.27
## -----------------------------------------------------
```

```
library(gtsummary)
tbl_regression(reg_all)
```

```
## Registered S3 methods overwritten by 'broom':
##   method            from
##   tidy.glht         jtools
##   tidy.summary.glht jtools
```

Characteristic	Beta	95% CI[1]	p-value
wins	0.76	-1.3, 2.8	0.5
market_size	12	8.5, 15	<0.001
salary_expenses	0.83	-11, 12	0.9
num_champ	-3.0	-8.4, 2.5	0.3

[1] CI = Confidence Interval

```
library(huxtable)
```

```
##
## Attaching package: 'huxtable'
```

```
## The following object is masked from 'package:gtsummary':
##
##     as_flextable
```

```
export_summs(reg_wins, reg_salary, reg_market, reg_pop,  scale = FALSE)
```

	Model 1	Model 2	Model 3	Model 4
(Intercept)	253.41 ***	138.47 *	230.90 ***	308.80 ***
	(65.72)	(64.51)	(13.88)	(18.67)
wins	1.97			
	(1.54)			
salary_expenses		20.75 **		
		(6.64)		
market_size			11.17 ***	
			(1.17)	
num_champ				10.12 *
				(3.77)
N	30	30	30	30
R2	0.05	0.26	0.77	0.20

*** p < 0.001; ** p < 0.01; * p < 0.05.

After computing the regression model, we can see **salary expenses and market size are strongly correlated to revenue**, while the number of champions and wins have a positive correlation but are not as remarkably effective as salary expenses and market size. **Market sales have a beta value of 11.9** which means as market size increase by 11.9 million, revenue will tend to increase by one million. Salary expenses are also statistically effective on revenue, when Salary expenses increase, revenue will also increase. Because teams with high salary expenses will have more talented players in the teams, which will attract more fans to watch the game and the team will also sell more merch products.