

# PSTAT 126 final

Justin Lee

6/7/2020

```
library(readr)
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.0      v dplyr   0.8.5
## v tibble  3.0.1      v stringr 1.4.0
## v tidyr   1.0.2      v forcats 0.5.0
## v purrr   0.3.4

## Warning: package 'ggplot2' was built under R version 3.6.3

## Warning: package 'tibble' was built under R version 3.6.3

## Warning: package 'tidyr' was built under R version 3.6.3

## Warning: package 'purrr' was built under R version 3.6.3

## Warning: package 'dplyr' was built under R version 3.6.3

## Warning: package 'forcats' was built under R version 3.6.3

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(alr4)
```

```
## Warning: package 'alr4' was built under R version 3.6.3

## Loading required package: car

## Warning: package 'car' was built under R version 3.6.3

## Loading required package: carData

##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':  
##  
##   recode
```

```
## The following object is masked from 'package:purrr':  
##  
##   some
```

```
## Loading required package: effects
```

```
## Warning: package 'effects' was built under R version 3.6.3
```

```
## Registered S3 methods overwritten by 'lme4':  
##   method                      from  
##   cooks.distance.influence.merMod car  
##   influence.merMod             car  
##   dfbeta.influence.merMod      car  
##   dfbetas.influence.merMod     car
```

```
## lattice theme set by effectsTheme()  
## See ?effectsTheme for details.
```

```
library(lindia)
```

```
## Warning: package 'lindia' was built under R version 3.6.3
```

```
nba <- read_csv("~/school/3. Spring 2020/PSTAT 126 (real)/final project/nba/nba 2018-2019.csv")
```

```
## Parsed with column specification:  
## cols(  
##   .default = col_double(),  
##   Team = col_character()  
## )
```

```
## See spec(...) for full column specifications.
```

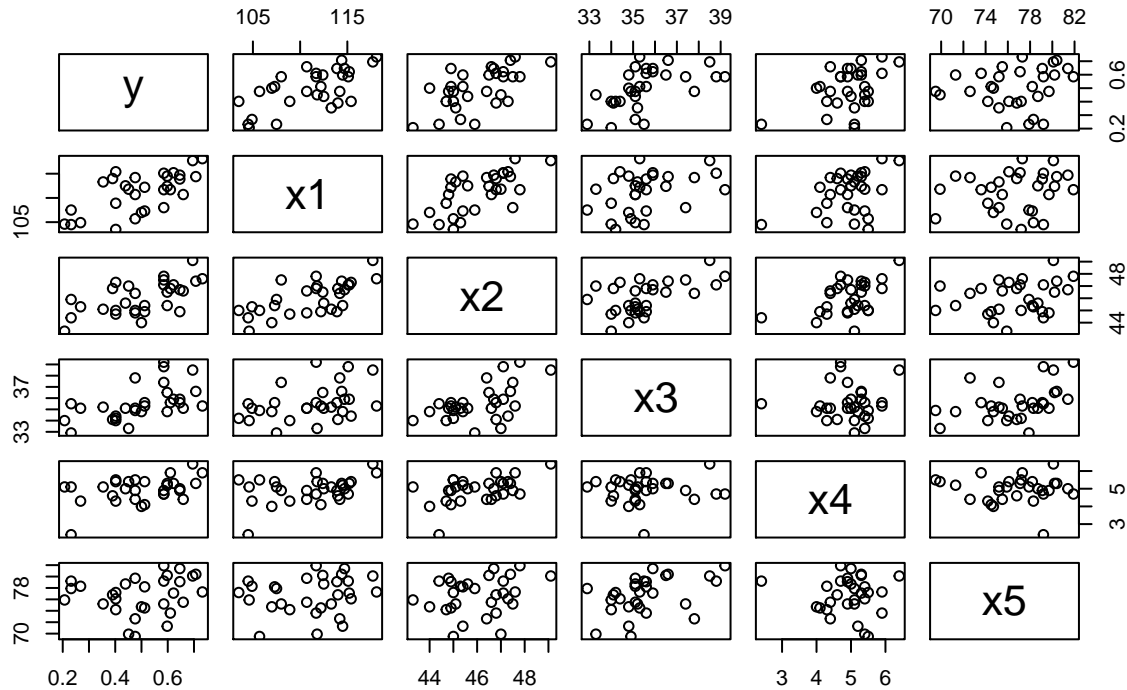
```
View(nba)
```

```
y = nba$`W/L`  
x1 = nba$PTS  
x2 = nba$`FG%`  
x3 = nba$`3P%`  
x4 = nba$TOV  
x5 = nba$`FT%`  
  
y
```

```
## [1] 0.732 0.707 0.695 0.659 0.646 0.646 0.622 0.610 0.598 0.598 0.585 0.585  
## [13] 0.585 0.512 0.512 0.500 0.476 0.476 0.476 0.451 0.439 0.402 0.402 0.402  
## [25] 0.390 0.354 0.268 0.232 0.232 0.207
```

```
pairs(~y+x1+x2+x3+x4+x5,data = nba, main = "Scatterplot Matrix")
```

## Scatterplot Matrix



```
mod0 <- lm(y~1)
mod.all<- lm(y~x1+x2+x3+x4+x5)
step(mod0, scope = list(lower = mod0, upper= mod.all))
```

```
## Start:  AIC=-114.18
## y ~ 1
##
##          Df Sum of Sq    RSS    AIC
## + x1      1  0.272538 0.35163 -129.39
## + x2      1  0.232651 0.39151 -126.17
## + x3      1  0.183083 0.44108 -122.59
## + x4      1  0.118108 0.50606 -118.47
## <none>                    0.62416 -114.18
## + x5      1  0.016084 0.60808 -112.96
##
## Step:  AIC=-129.39
## y ~ x1
##
##          Df Sum of Sq    RSS    AIC
## + x3      1  0.060554 0.29107 -133.06
## + x2      1  0.031622 0.32000 -130.22
## <none>                    0.35163 -129.39
## + x4      1  0.021438 0.33019 -129.28
```

```
## + x5      1  0.002017 0.34961 -127.56
## - x1      1  0.272538 0.62416 -114.18
##
## Step: AIC=-133.06
## y ~ x1 + x3
##
##          Df Sum of Sq      RSS      AIC
## + x4      1  0.032167 0.25891 -134.57
## <none>                                0.29107 -133.06
## + x2      1  0.006768 0.28430 -131.77
## + x5      1  0.003158 0.28792 -131.39
## - x3      1  0.060554 0.35163 -129.39
## - x1      1  0.150009 0.44108 -122.59
##
## Step: AIC=-134.57
## y ~ x1 + x3 + x4
##
##          Df Sum of Sq      RSS      AIC
## <none>                                0.25891 -134.57
## - x4      1  0.032167 0.29107 -133.06
## + x5      1  0.001685 0.25722 -132.77
## + x2      1  0.000103 0.25880 -132.59
## - x3      1  0.071282 0.33019 -129.28
## - x1      1  0.077052 0.33596 -128.76

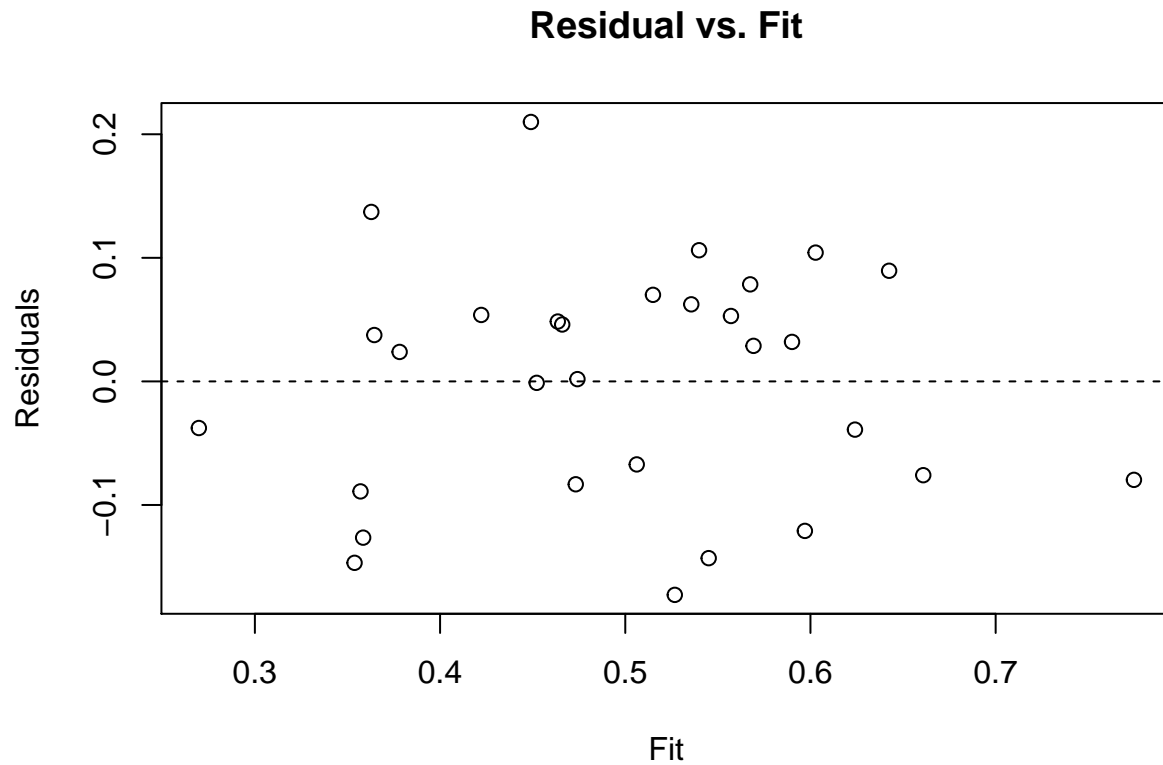
##
## Call:
## lm(formula = y ~ x1 + x3 + x4)
##
## Coefficients:
## (Intercept)          x1          x3          x4
##    -2.67318      0.01500      0.03534      0.05022
```

```
fit <- lm(y~x1+x3+x4)
summary(fit)
```

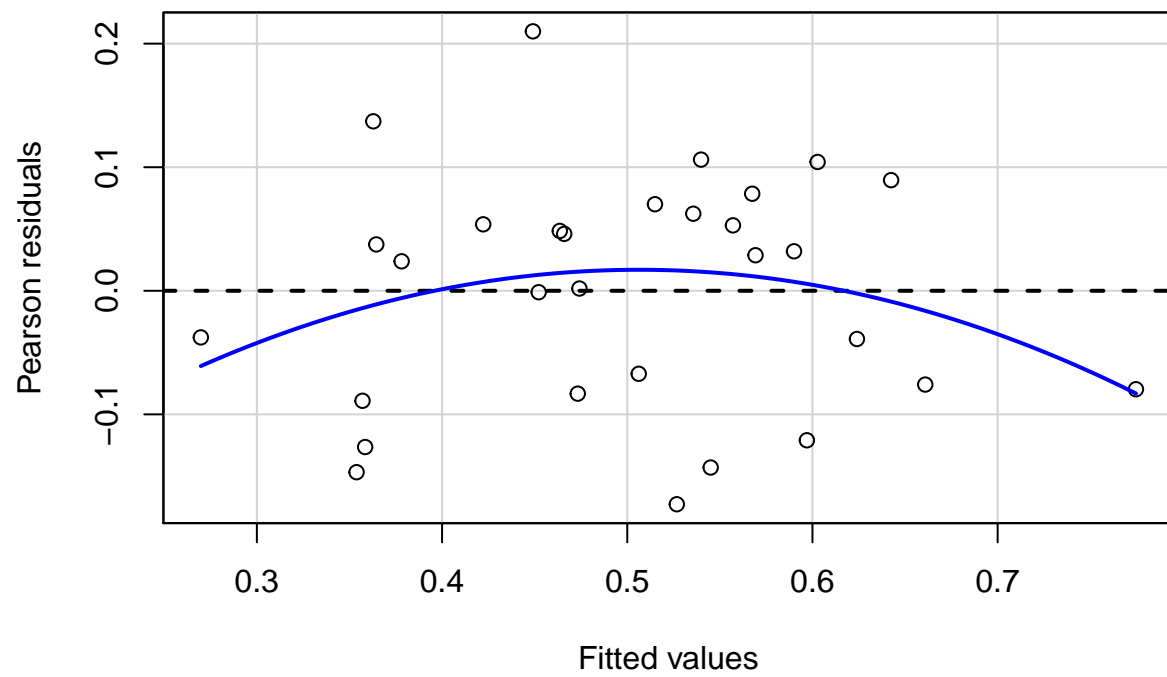
```
##
## Call:
## lm(formula = y ~ x1 + x3 + x4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17275 -0.07874  0.02630  0.06022  0.20994
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.673175   0.567434  -4.711 7.21e-05 ***
## x1           0.015002   0.005393   2.782 0.00993 **
## x3           0.035344   0.013210   2.676 0.01274 *
## x4           0.050224   0.027944   1.797 0.08391 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.09979 on 26 degrees of freedom
## Multiple R-squared:  0.5852, Adjusted R-squared:  0.5373
## F-statistic: 12.23 on 3 and 26 DF,  p-value: 3.531e-05
```

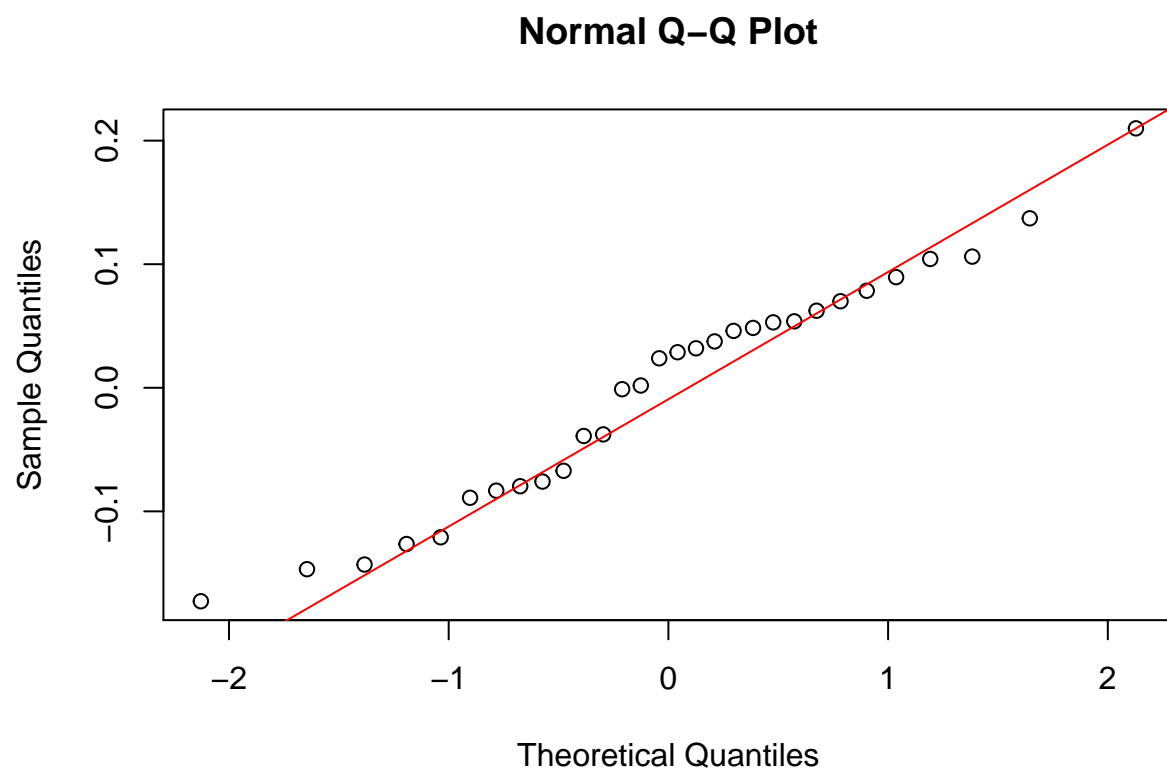
```
res = resid(fit)
y_hat = fitted(fit)
plot(y_hat, res, main = "Residual vs. Fit", ylab = "Residuals", xlab = "Fit")
abline(h = 0, lty = 2)
```



```
residualPlot(fit)
```



```
qqnorm(res)
qqline(res, col = 2)
```



yes

no

#yesa

```
anova(fit)
```

```
## Analysis of Variance Table
##
## Response: y
##          Df    Sum Sq  Mean Sq F value    Pr(>F)
## x1         1  0.272538  0.272538  27.3690 1.826e-05 ***
## x3         1  0.060554  0.060554   6.0810  0.02058 *
## x4         1  0.032167  0.032167   3.2303  0.08391 .
## Residuals 26  0.258906  0.009958
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(fit)
```

```
##
```

```
## Call:
## lm(formula = y ~ x1 + x3 + x4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17275 -0.07874  0.02630  0.06022  0.20994
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.673175   0.567434  -4.711 7.21e-05 ***
## x1           0.015002   0.005393   2.782 0.00993 **
## x3           0.035344   0.013210   2.676 0.01274 *
## x4           0.050224   0.027944   1.797 0.08391 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09979 on 26 degrees of freedom
## Multiple R-squared:  0.5852, Adjusted R-squared:  0.5373
## F-statistic: 12.23 on 3 and 26 DF,  p-value: 3.531e-05
```