# Analysis of MBA SALARIES

# NAME: <ASHUTOSH SHAW>

# EMAIL: <ashu.shaw09@gmail.com>

# COLLEGE: <DTU>

**OUTPUTS**

summary(mba.df)

age sex gmat\_tot gmat\_qpc gmat\_vpc gmat\_tpc s\_avg f\_avg quarter

Min. :22.00 Min. :1.000 Min. :450.0 Min. :28.00 Min. :16.00 Min. : 0.0 Min. :2.000 Min. :0.000 Min. :1.000

1st Qu.:25.00 1st Qu.:1.000 1st Qu.:580.0 1st Qu.:72.00 1st Qu.:71.00 1st Qu.:78.0 1st Qu.:2.708 1st Qu.:2.750 1st Qu.:1.250

Median :27.00 Median :1.000 Median :620.0 Median :83.00 Median :81.00 Median :87.0 Median :3.000 Median :3.000 Median :2.000

Mean :27.36 Mean :1.248 Mean :619.5 Mean :80.64 Mean :78.32 Mean :84.2 Mean :3.025 Mean :3.062 Mean :2.478

3rd Qu.:29.00 3rd Qu.:1.000 3rd Qu.:660.0 3rd Qu.:93.00 3rd Qu.:91.00 3rd Qu.:94.0 3rd Qu.:3.300 3rd Qu.:3.250 3rd Qu.:3.000

Max. :48.00 Max. :2.000 Max. :790.0 Max. :99.00 Max. :99.00 Max. :99.0 Max. :4.000 Max. :4.000 Max. :4.000

work\_yrs frstlang salary satis

Min. : 0.000 Min. :1.000 Min. : 0 Min. : 1.0

1st Qu.: 2.000 1st Qu.:1.000 1st Qu.: 0 1st Qu.: 5.0

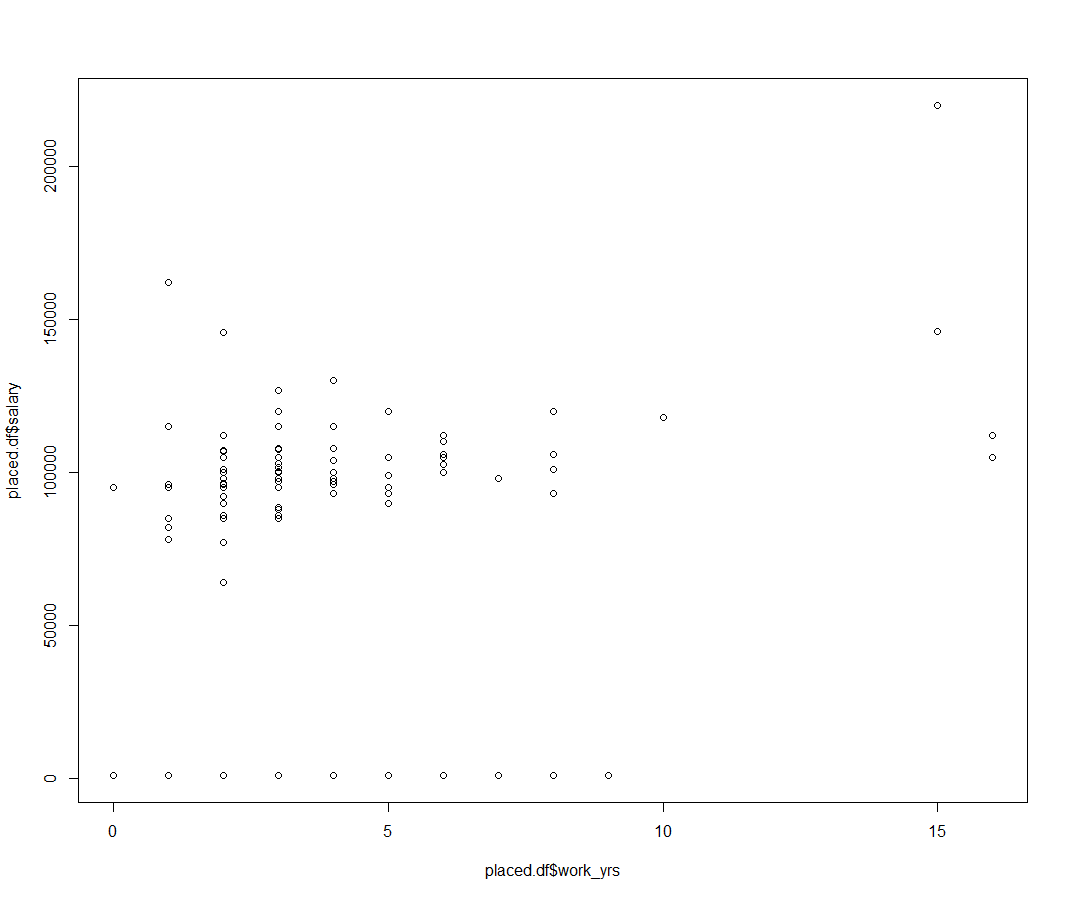
Median : 3.000 Median :1.000 Median : 999 Median : 6.0

Mean : 3.872 Mean :1.117 Mean : 39026 Mean :172.2

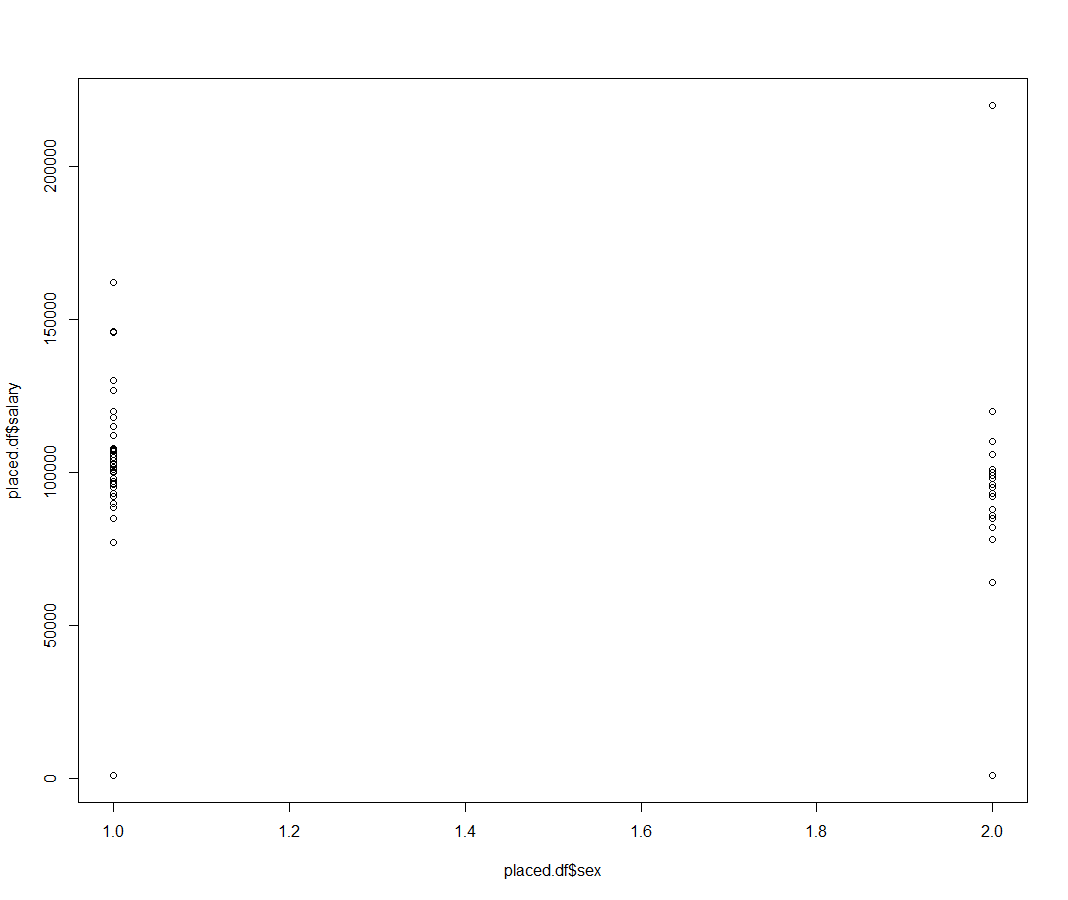
3rd Qu.: 4.000 3rd Qu.:1.000 3rd Qu.: 97000 3rd Qu.: 7.0

Max. :22.000 Max. :2.000 Max. :220000 Max. :998.0

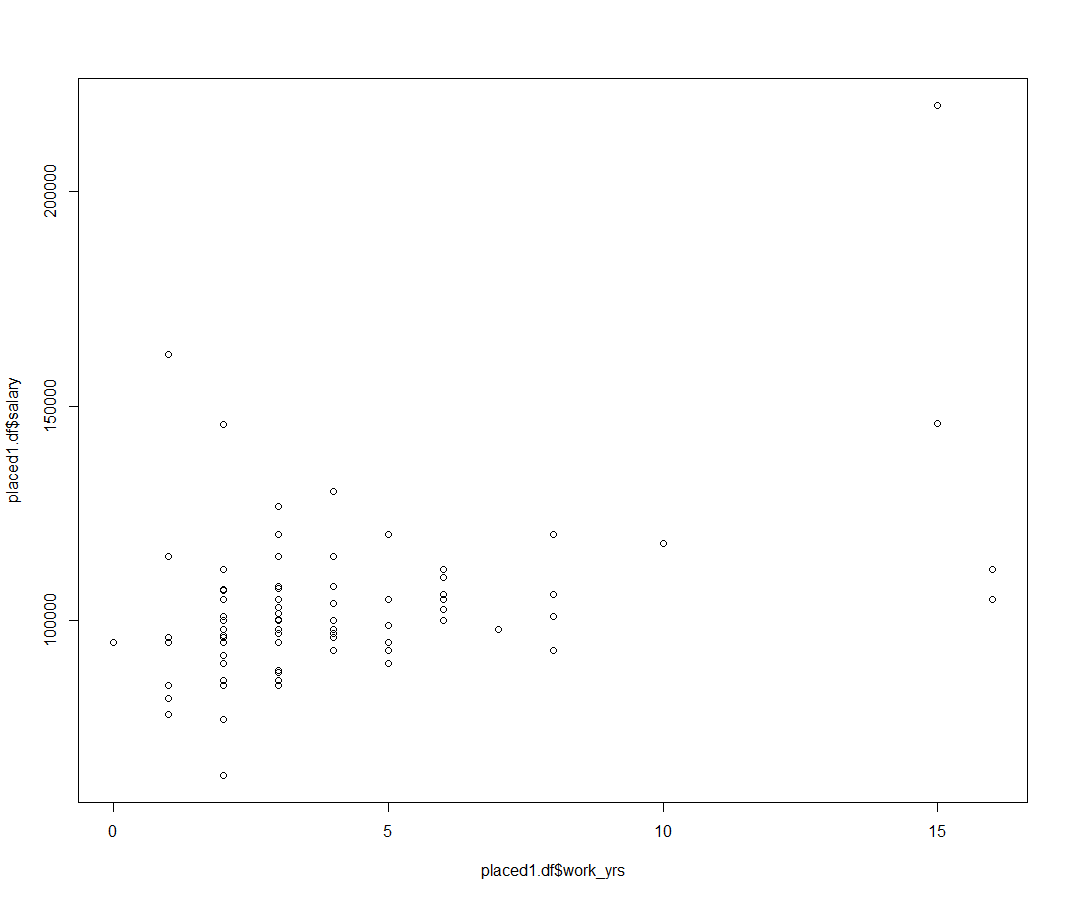
plot(placed.df$work\_yrs,placed.df$salary)



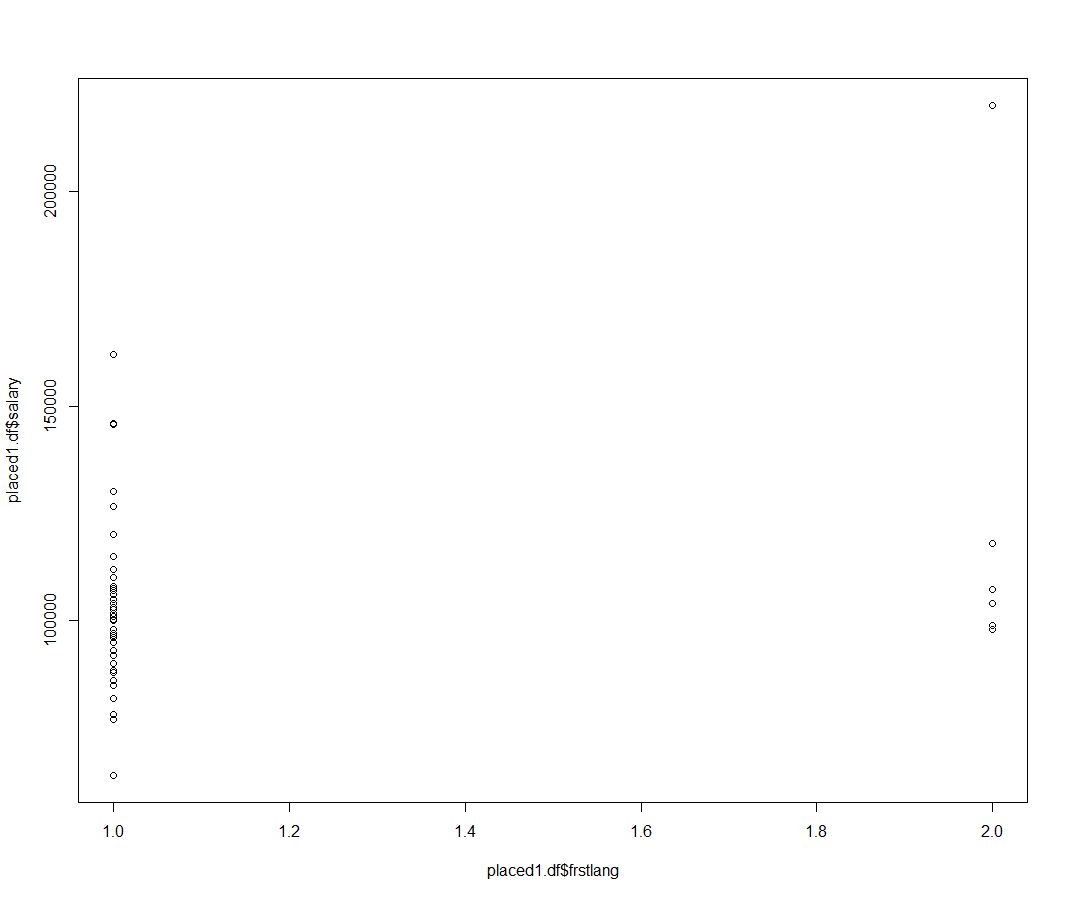
plot(placed.df$sex,placed.df$salary)



plot(placed1.df$work\_yrs,placed1.df$salary)



plot(placed1.df$frstlang,placed1.df$salary)



cor.test(placed1.df$frstlang,placed1.df$salary)

Pearson's product-moment correlation

data: placed1.df$frstlang and placed1.df$salary

t = 2.7846, df = 101, p-value = 0.0064

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

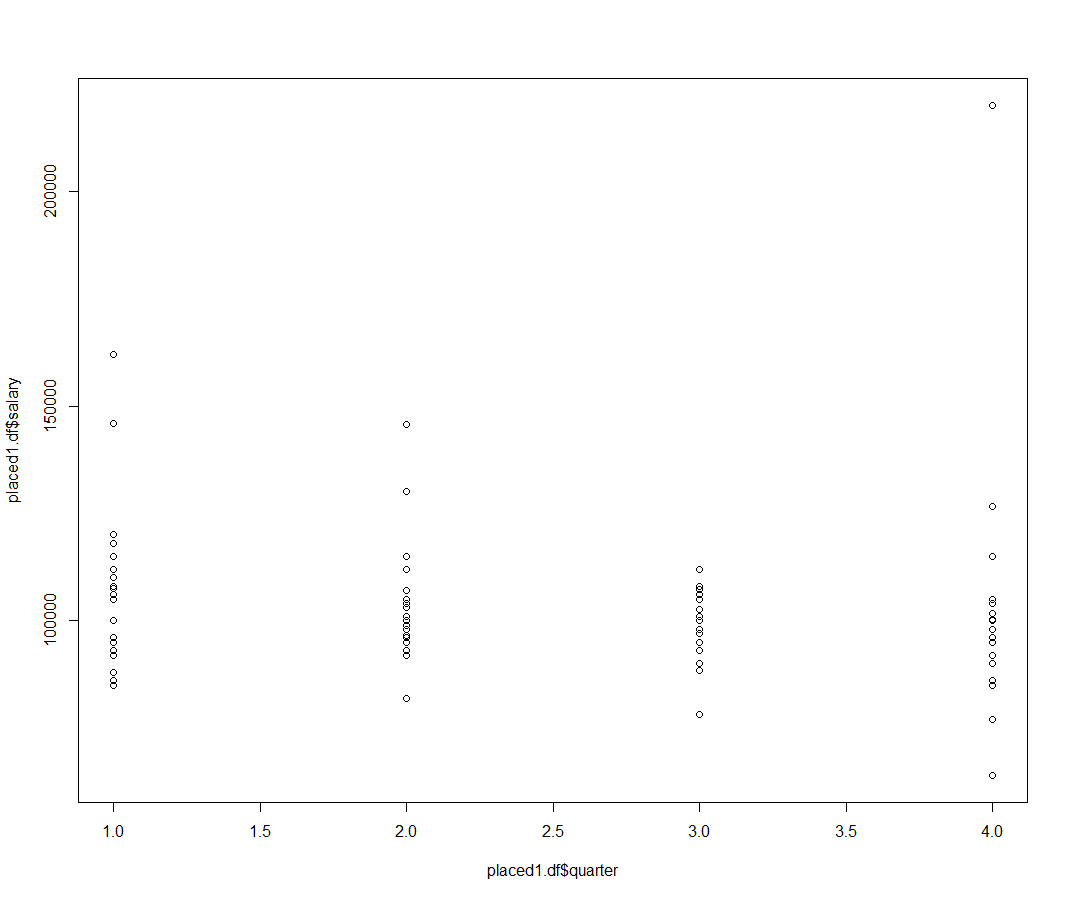
0.07749965 0.43791500

sample estimates:

cor

0.267019

plot(placed1.df$quarter,placed1.df$salary)



cor.test(placed1.df$quarter,placed1.df$salary)

Pearson's product-moment correlation

data: placed1.df$quarter and placed1.df$salary

t = -1.3021, df = 101, p-value = 0.1959

alternative hypothesis: true correlation is not equal to 0

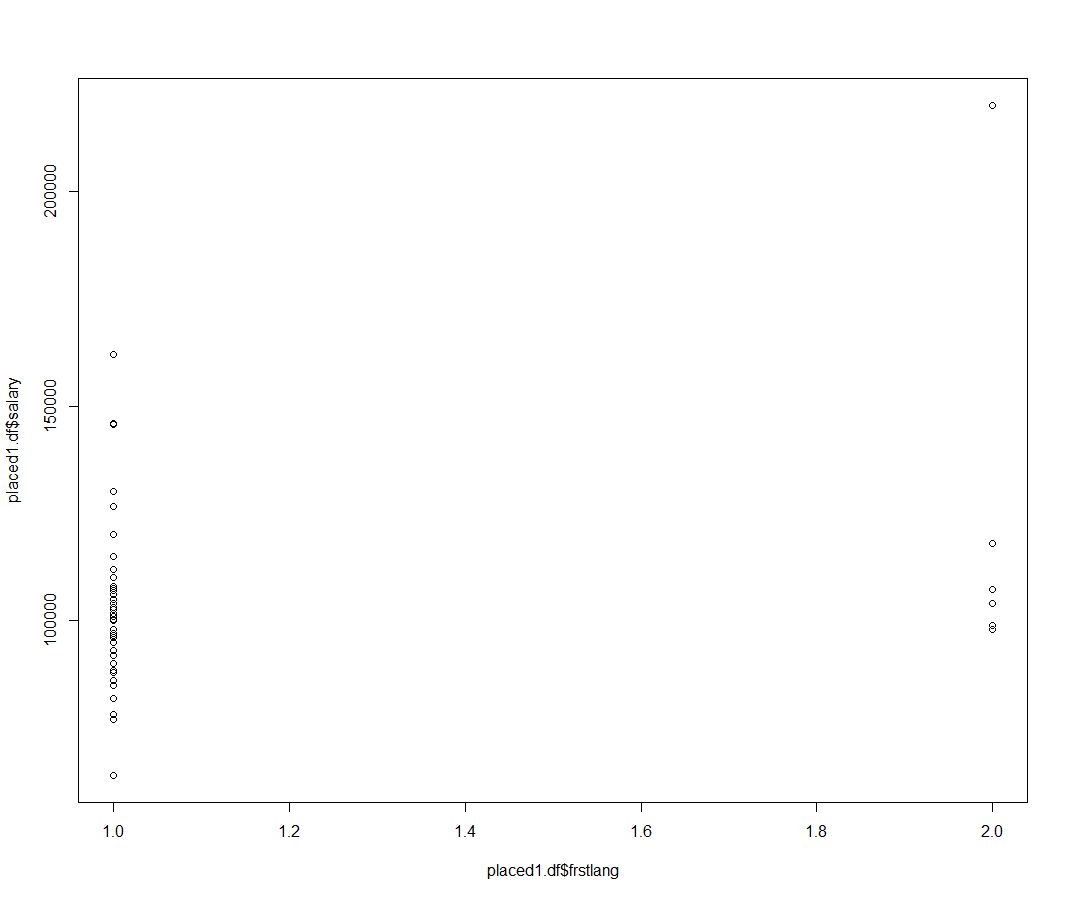
95 percent confidence interval:

-0.31419739 0.06669785

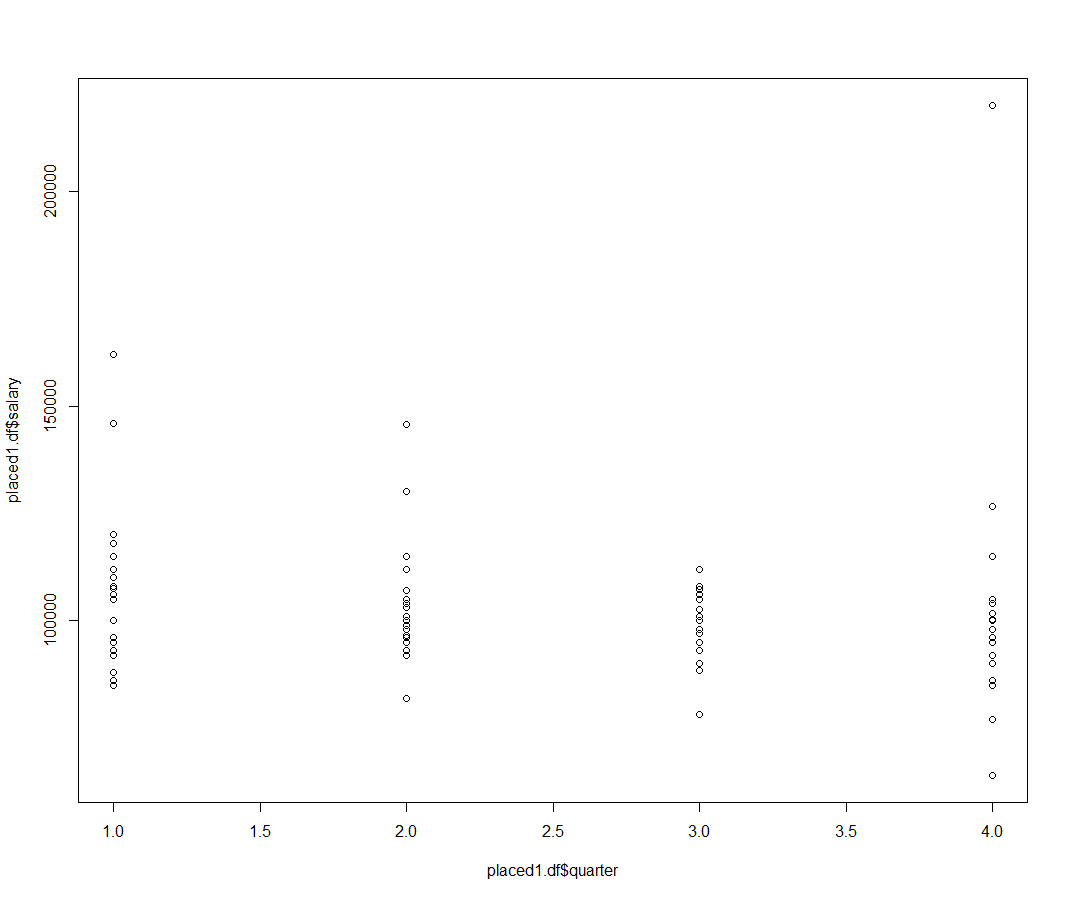
sample estimates:

cor

-0.1284853



plot(placed1.df$quarter,placed1.df$salary)



> cor.test(placed1.df$quarter,placed1.df$salary) #insignificant,low correlation

Pearson's product-moment correlation

data: placed1.df$quarter and placed1.df$salary

t = -1.3021, df = 101, p-value = 0.1959

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

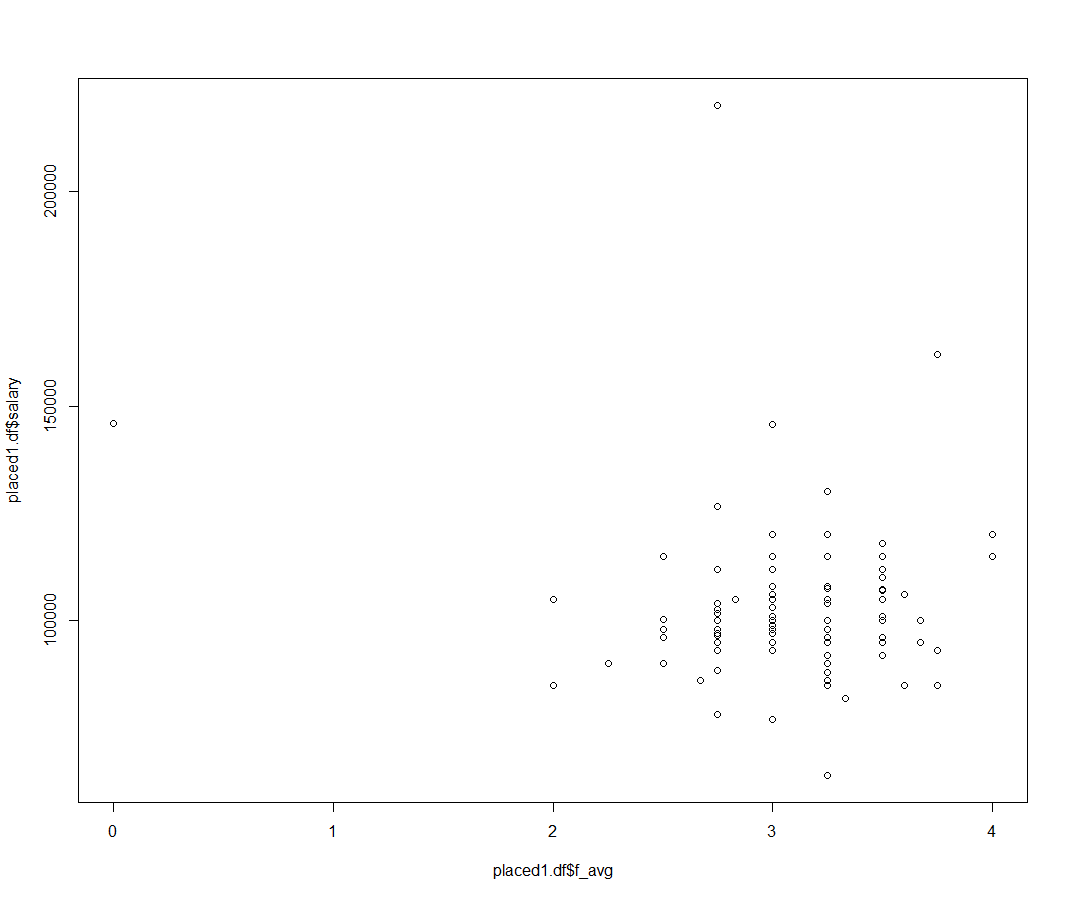
-0.31419739 0.06669785

sample estimates:

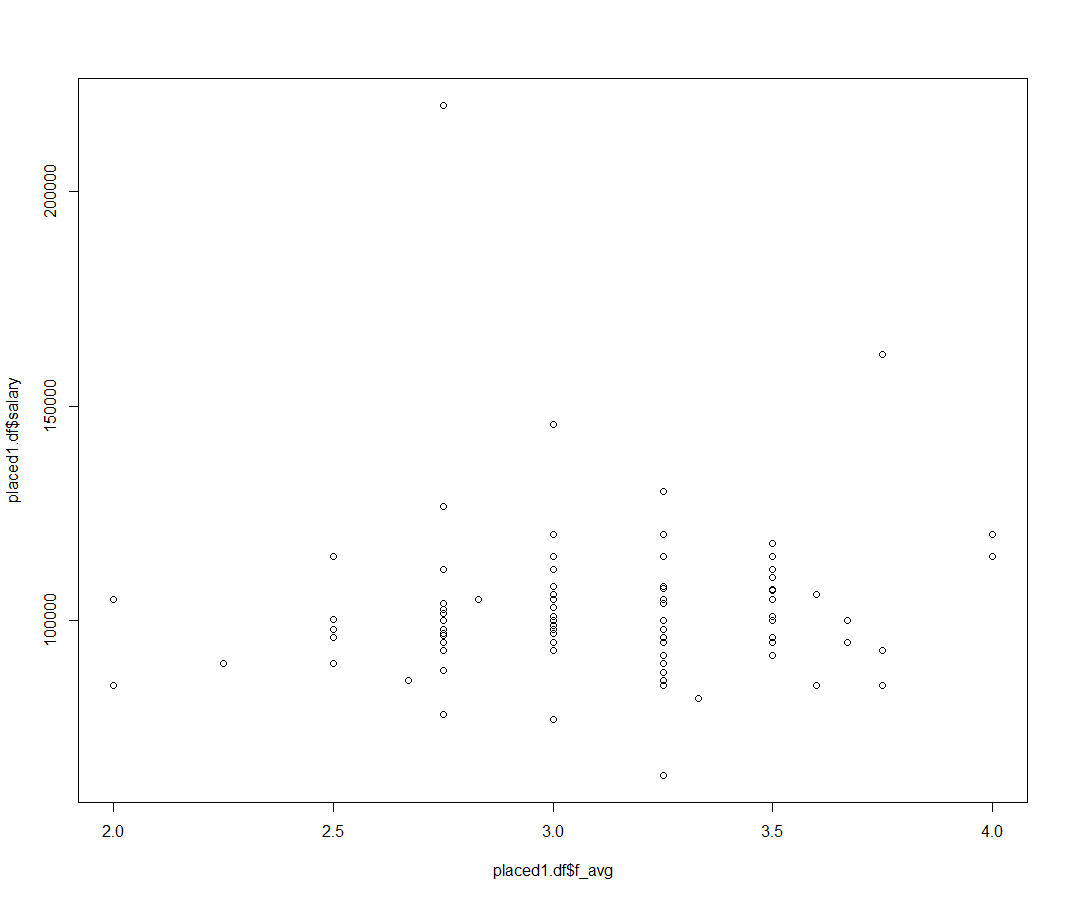
cor

-0.1284853

plot(placed1.df$f\_avg,placed1.df$salary)



plot(placed1.df$f\_avg,placed1.df$salary,xlim=c(2,4))



> cor.test(placed1.df$f\_avg,placed1.df$salary)

Pearson's product-moment correlation

data: placed1.df$f\_avg and placed1.df$salary

t = -1.0717, df = 101, p-value = 0.2864

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

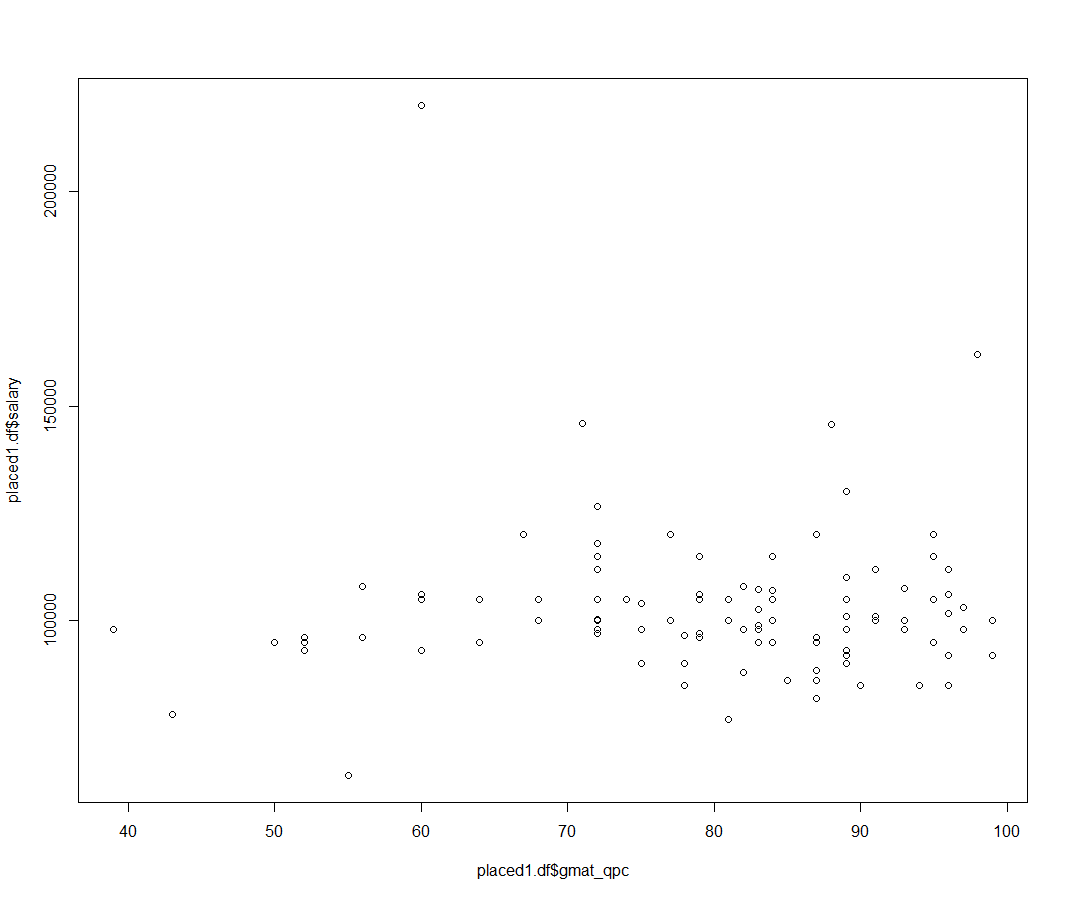
-0.29353985 0.08931862

sample estimates:

cor

-0.106039

plot(placed1.df$gmat\_tpc,placed1.df$salary,xlim=c(60,100)) # Observe the values from x=60 to x=100



cor.test(placed1.df$gmat\_qpc,placed1.df$salary) #insignificant, low correlation

Pearson's product-moment correlation

data: placed1.df$gmat\_qpc and placed1.df$salary

t = 0.14213, df = 101, p-value = 0.8873

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.1798756 0.2070992

sample estimates:

cor

0.0141413

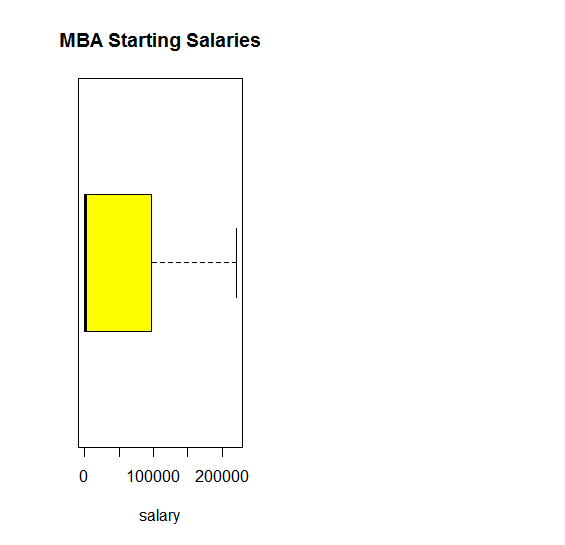
> boxplot(salary,

+ main="MBA Starting Salaries",

+ horizontal=TRUE,

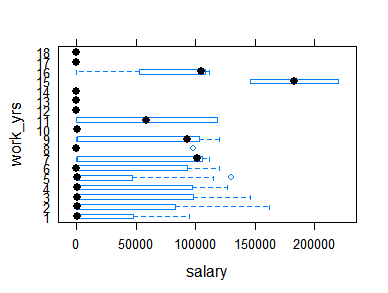
+ col=c("yellow"),

+ xlab="salary" )



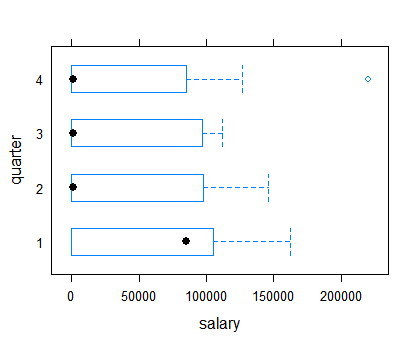
> bwplot(work\_yrs ~ salary,horizontal=TRUE,

+ xlab = "salary")

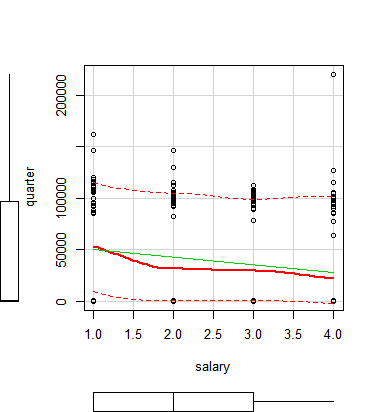


> bwplot(quarter ~ salary,horizontal=TRUE,

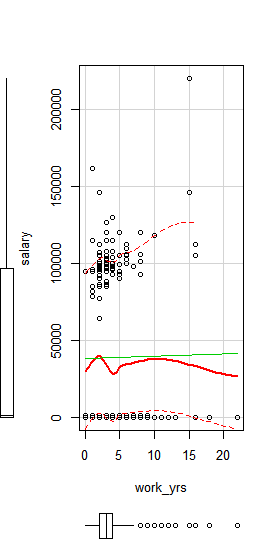
+ xlab = "salary")



scatterplot(quarter,salary, xlab="salary", ylab="quarter")

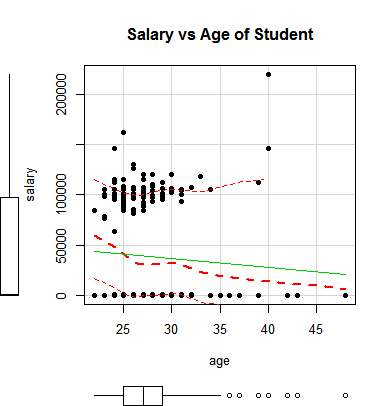


> scatterplot(work\_yrs,salary, xlab="work\_yrs", ylab="salary")



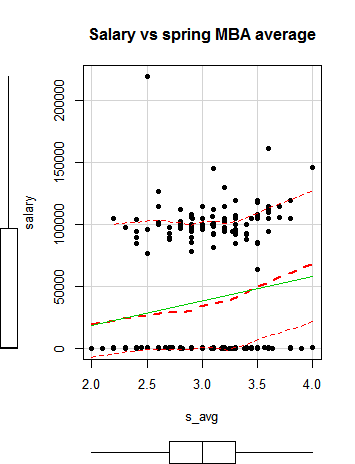
> scatterplot(salary ~ age, data=mba, spread=TRUE, smoother.args=list(lty=2), pch=19,

+ main="Salary vs Age of Student", xlab="age", ylab="salary")



> scatterplot(salary ~ s\_avg, data=mba, spread=TRUE, smoother.args=list(lty=2), pch=19,

+ main="Salary vs spring MBA average", xlab="s\_avg", ylab="salary")



> cov(mba)

age sex gmat\_tot gmat\_qpc gmat\_vpc gmat\_tpc s\_avg

age 1.4e+01 -4.5e-02 -3.1e+01 -1.2e+01 -2.76 -8.840 0.212

sex -4.5e-02 1.9e-01 -1.3e+00 -1.1e+00 0.55 -0.049 0.021

gmat\_tot -3.1e+01 -1.3e+00 3.3e+03 6.2e+02 726.00 683.991 2.480

gmat\_qpc -1.2e+01 -1.1e+00 6.2e+02 2.2e+02 38.15 135.800 -0.169

gmat\_vpc -2.8e+00 5.5e-01 7.3e+02 3.8e+01 284.25 157.493 1.314

gmat\_tpc -8.8e+00 -4.9e-02 6.8e+02 1.4e+02 157.49 196.606 0.627

s\_avg 2.1e-01 2.1e-02 2.5e+00 -1.7e-01 1.31 0.627 0.145

f\_avg -3.4e-02 2.1e-02 3.2e+00 5.8e-01 0.67 0.587 0.110

quarter -2.0e-01 -6.4e-02 -5.9e+00 6.0e-01 -3.27 -1.292 -0.322

work\_yrs 1.0e+01 -1.6e-02 -3.4e+01 -1.1e+01 -3.62 -7.858 0.159

frstlang 6.8e-02 2.1e-04 -2.5e+00 6.6e-01 -2.11 -0.466 -0.017

salary -1.2e+04 1.5e+03 -1.6e+05 -3.3e+04 -5273.85 3522.750 2831.601

satis -1.8e+02 -8.8e+00 1.8e+03 3.3e+02 392.36 484.247 -4.629

f\_avg quarter work\_yrs frstlang salary satis

age -0.0340 -2.0e-01 10.295 6.8e-02 -1.2e+04 -1.8e+02

sex 0.0208 -6.4e-02 -0.016 2.1e-04 1.5e+03 -8.8e+00

gmat\_tot 3.1547 -5.9e+00 -33.916 -2.5e+00 -1.6e+05 1.8e+03

gmat\_qpc 0.5754 6.0e-01 -11.372 6.6e-01 -3.3e+04 3.3e+02

gmat\_vpc 0.6721 -3.3e+00 -3.618 -2.1e+00 -5.3e+03 3.9e+02

gmat\_tpc 0.5870 -1.3e+00 -7.858 -4.7e-01 3.5e+03 4.8e+02

s\_avg 0.1102 -3.2e-01 0.159 -1.7e-02 2.8e+03 -4.6e+00

f\_avg 0.2757 -2.6e-01 -0.066 -6.3e-03 7.9e+02 2.1e+00

quarter -0.2608 1.2e+00 -0.309 3.6e-02 -9.3e+03 -5.2e-03

work\_yrs -0.0663 -3.1e-01 10.449 -2.9e-02 1.5e+03 -1.3e+02

frstlang -0.0063 3.6e-02 -0.029 1.0e-01 -1.4e+03 9.5e+00

salary 787.6560 -9.3e+03 1486.147 -1.4e+03 2.6e+09 -6.3e+06

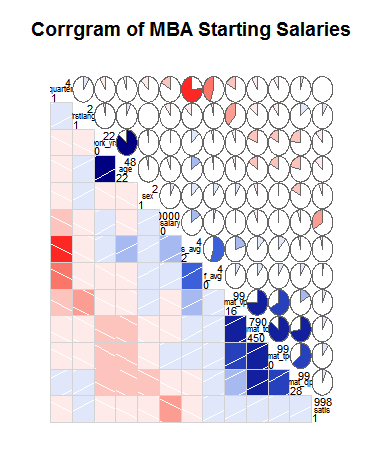
satis 2.1253 -5.2e-03 -131.241 9.5e+00 -6.3e+06 1.4e+05

> corrgram(mba, order=TRUE,

+ lower.panel=panel.shade, upper.panel=panel.pie,

+ diag.panel=panel.minmax, text.panel=panel.txt,

+ main="Corrgram of MBA Starting Salaries")



> placed <- mba[ which(mba$salary>1000) , ]

> View(placed)

> mytable <- xtabs(~ work\_yrs+salary, data=placed)

> addmargins(mytable)

salary

work\_yrs 64000 77000 78256 82000 85000 86000 88000 88500 90000 92000 93000 95000 96000 96500 97000 98000 99000 100000 100400 101000 101100 101600 102500

0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

1 0 0 1 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0

2 1 1 0 0 2 1 0 0 2 3 0 2 2 1 0 7 0 6 0 2 0 0 0

3 0 0 0 0 1 1 1 1 0 0 0 2 0 0 1 1 0 1 1 0 0 1 0

4 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 1 0 1 0 0 0 0 0

5 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 1 0 0 0 0 0 0

6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1

7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

8 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0

10 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

16 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Sum 1 1 1 1 4 2 1 1 3 3 3 7 4 1 2 10 1 9 1 2 1 1 1

salary

work\_yrs 103000 104000 105000 106000 107000 107300 107500 108000 110000 112000 115000 118000 120000 126710 130000 145800 146000 162000 220000 Sum

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

1 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 1 0 8

2 0 0 4 0 1 1 0 0 0 1 0 0 0 0 0 1 0 0 0 38

3 1 0 4 0 0 0 1 1 0 0 1 0 1 1 0 0 0 0 0 21

4 0 2 0 0 0 0 0 1 0 0 2 0 0 0 1 0 0 0 0 11

5 0 0 1 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 7

6 0 0 1 2 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 7

7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

8 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 4

10 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1

15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 2

16 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 2

Sum 1 2 11 3 1 1 1 2 1 3 5 1 4 1 1 1 1 1 1 103

> mytable # frequencies

salary

work\_yrs 64000 77000 78256 82000 85000 86000 88000 88500 90000 92000 93000 95000 96000 96500 97000 98000 99000 100000 100400 101000 101100 101600 102500

0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

1 0 0 1 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0

2 1 1 0 0 2 1 0 0 2 3 0 2 2 1 0 7 0 6 0 2 0 0 0

3 0 0 0 0 1 1 1 1 0 0 0 2 0 0 1 1 0 1 1 0 0 1 0

4 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 1 0 1 0 0 0 0 0

5 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 1 0 0 0 0 0 0

6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1

7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

8 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0

10 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

16 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

salary

work\_yrs 103000 104000 105000 106000 107000 107300 107500 108000 110000 112000 115000 118000 120000 126710 130000 145800 146000 162000 220000

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

1 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 1 0

2 0 0 4 0 1 1 0 0 0 1 0 0 0 0 0 1 0 0 0

3 1 0 4 0 0 0 1 1 0 0 1 0 1 1 0 0 0 0 0

4 0 2 0 0 0 0 0 1 0 0 2 0 0 0 1 0 0 0 0

5 0 0 1 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0

6 0 0 1 2 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0

7 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

8 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0

10 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1

16 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

> chisq.test(mytable)

Pearson's Chi-squared test

data: mytable

X-squared = 500, df = 500, p-value = 0.004

Warning message:

In chisq.test(mytable) : Chi-squared approximation may be incorrect

> mytable1 <- xtabs(~ age+salary, data=placed)

> mytable1 # frequencies

salary

age 64000 77000 78256 82000 85000 86000 88000 88500 90000 92000 93000 95000 96000 96500 97000 98000 99000 100000 100400 101000 101100 101600 102500 103000

22 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

23 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0

24 1 0 0 0 0 0 0 0 0 0 0 1 1 1 0 3 0 4 0 1 0 0 0 0

25 0 0 0 0 1 1 1 0 2 2 1 5 1 0 0 2 0 1 0 1 0 0 0 0

26 0 0 0 1 1 1 0 0 0 0 0 0 2 0 0 1 0 1 0 0 0 1 0 0

27 0 0 0 0 1 0 0 1 1 1 0 0 0 0 1 1 0 1 0 0 0 0 0 1

28 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1 1 0 0 0 0 0 0 0

29 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 0 0

30 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0

31 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0

32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

33 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

34 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

39 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

40 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

salary

age 104000 105000 106000 107000 107300 107500 108000 110000 112000 115000 118000 120000 126710 130000 145800 146000 162000 220000

22 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

23 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

24 0 1 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0

25 0 2 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0

26 0 3 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0

27 0 1 0 0 0 1 0 0 0 3 0 1 0 0 0 0 0 0

28 1 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0

29 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

30 0 1 2 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0

31 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

32 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0

33 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

34 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

39 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

40 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1

> chisq.test(mytable1)

Pearson's Chi-squared test

data: mytable1

X-squared = 700, df = 600, p-value = 4e-05

Warning message:

In chisq.test(mytable1) : Chi-squared approximation may be incorrect

> mytable2 <- xtabs(~ salary+quarter, data=placed)

> addmargins(mytable2)

quarter

salary 1 2 3 4 Sum

64000 0 0 0 1 1

77000 0 0 0 1 1

78256 0 0 1 0 1

82000 0 1 0 0 1

85000 2 0 0 2 4

86000 1 0 0 1 2

88000 1 0 0 0 1

88500 0 0 1 0 1

90000 0 0 2 1 3

92000 1 1 0 1 3

93000 1 1 1 0 3

95000 3 2 1 1 7

96000 2 1 0 1 4

96500 0 1 0 0 1

97000 0 0 2 0 2

98000 0 3 6 1 10

99000 0 1 0 0 1

100000 3 2 2 2 9

100400 0 0 0 1 1

101000 0 1 1 0 2

101100 0 0 1 0 1

101600 0 0 0 1 1

102500 0 0 1 0 1

103000 0 1 0 0 1

104000 0 1 0 1 2

105000 6 3 1 1 11

106000 2 0 1 0 3

107000 0 1 0 0 1

107300 0 0 1 0 1

107500 1 0 0 0 1

108000 1 0 1 0 2

110000 1 0 0 0 1

112000 1 1 1 0 3

115000 2 2 0 1 5

118000 1 0 0 0 1

120000 4 0 0 0 4

126710 0 0 0 1 1

130000 0 1 0 0 1

145800 0 1 0 0 1

146000 1 0 0 0 1

162000 1 0 0 0 1

220000 0 0 0 1 1

Sum 35 25 24 19 103

> mytable2 # frequencies

quarter

salary 1 2 3 4

64000 0 0 0 1

77000 0 0 0 1

78256 0 0 1 0

82000 0 1 0 0

85000 2 0 0 2

86000 1 0 0 1

88000 1 0 0 0

88500 0 0 1 0

90000 0 0 2 1

92000 1 1 0 1

93000 1 1 1 0

95000 3 2 1 1

96000 2 1 0 1

96500 0 1 0 0

97000 0 0 2 0

98000 0 3 6 1

99000 0 1 0 0

100000 3 2 2 2

100400 0 0 0 1

101000 0 1 1 0

101100 0 0 1 0

101600 0 0 0 1

102500 0 0 1 0

103000 0 1 0 0

104000 0 1 0 1

105000 6 3 1 1

106000 2 0 1 0

107000 0 1 0 0

107300 0 0 1 0

107500 1 0 0 0

108000 1 0 1 0

110000 1 0 0 0

112000 1 1 1 0

115000 2 2 0 1

118000 1 0 0 0

120000 4 0 0 0

126710 0 0 0 1

130000 0 1 0 0

145800 0 1 0 0

146000 1 0 0 0

162000 1 0 0 0

220000 0 0 0 1

> chisq.test(mytable2)

Pearson's Chi-squared test

data: mytable2

X-squared = 100, df = 100, p-value = 0.3

Warning message:

In chisq.test(mytable2) : Chi-squared approximation may be incorrect

> mytable3 <- xtabs(~ salary+s\_avg, data=placed)

> addmargins(mytable3)

s\_avg

salary 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 2.91 3 3.09 3.1 3.2 3.27 3.3 3.4 3.45 3.5 3.6 3.7 3.8 4 Sum

64000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1

77000 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

78256 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

82000 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1

85000 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 2 0 0 0 0 4

86000 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 2

88000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1

88500 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

90000 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3

92000 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 3

93000 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 3

95000 0 0 1 0 0 0 0 1 0 0 0 0 1 1 2 0 0 0 1 0 0 0 7

96000 0 0 0 1 0 0 0 0 0 0 0 0 1 0 2 0 0 0 0 0 0 0 4

96500 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1

97000 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2

98000 0 1 0 0 0 1 1 4 0 1 0 0 2 0 0 0 0 0 0 0 0 0 10

99000 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1

100000 0 0 0 0 2 0 1 1 0 1 1 0 0 0 1 2 0 0 0 0 0 0 9

100400 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

101000 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 2

101100 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

101600 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

102500 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

103000 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1

104000 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 2

105000 1 0 0 0 0 0 0 0 1 2 0 0 1 0 2 0 1 1 1 0 1 0 11

106000 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0 3

107000 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1

107300 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

107500 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1

108000 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 2

110000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1

112000 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 3

115000 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 1 0 0 5

118000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1

120000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 2 0 0 1 0 4

126710 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

130000 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1

145800 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1

146000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1

162000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1

220000 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

Sum 1 1 4 3 6 4 7 13 1 6 2 7 9 1 11 5 1 10 6 2 2 1 103

> mytable3 # frequencies

s\_avg

salary 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 2.91 3 3.09 3.1 3.2 3.27 3.3 3.4 3.45 3.5 3.6 3.7 3.8 4

64000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0

77000 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

78256 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

82000 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

85000 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 2 0 0 0 0

86000 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0

88000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0

88500 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

90000 0 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

92000 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0

93000 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0

95000 0 0 1 0 0 0 0 1 0 0 0 0 1 1 2 0 0 0 1 0 0 0

96000 0 0 0 1 0 0 0 0 0 0 0 0 1 0 2 0 0 0 0 0 0 0

96500 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

97000 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

98000 0 1 0 0 0 1 1 4 0 1 0 0 2 0 0 0 0 0 0 0 0 0

99000 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

100000 0 0 0 0 2 0 1 1 0 1 1 0 0 0 1 2 0 0 0 0 0 0

100400 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

101000 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

101100 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

101600 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

102500 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

103000 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

104000 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

105000 1 0 0 0 0 0 0 0 1 2 0 0 1 0 2 0 1 1 1 0 1 0

106000 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0

107000 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

107300 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

107500 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

108000 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0

110000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0

112000 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0

115000 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 1 0 0

118000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0

120000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 2 0 0 1 0

126710 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

130000 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

145800 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

146000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

162000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0

220000 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

> chisq.test(mytable3)

Pearson's Chi-squared test

data: mytable3

X-squared = 800, df = 900, p-value = 1

Warning message:

In chisq.test(mytable3) : Chi-squared approximation may be incorrect

> t.test(placed$salary,placed$work\_yrs, paired=TRUE)

Paired t-test

data: placed$salary and placed$work\_yrs

t = 60, df = 100, p-value <2e-16

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

99535 106519

sample estimates:

mean of the differences

1e+05

> t.test(placed$age, placed$salary, paired=TRUE)

Paired t-test

data: placed$age and placed$salary

t = -60, df = 100, p-value <2e-16

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-106496 -99512

sample estimates:

mean of the differences

-1e+05

> t.test(placed$quarter, placed$salary, paired=TRUE)

Paired t-test

data: placed$quarter and placed$salary

t = -60, df = 100, p-value <2e-16

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-106521 -99536

sample estimates:

mean of the differences

-1e+05

> t.test(placed$s\_avg, placed$salary, paired=TRUE)

Paired t-test

data: placed$s\_avg and placed$salary

t = -60, df = 100, p-value <2e-16

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-106520 -99535

sample estimates:

mean of the differences

-1e+05

> fit1 <- lm(salary ~ s\_avg, data = placed)

> summary(fit1)

Call:

lm(formula = salary ~ s\_avg, data = placed)

Residuals:

Min 1Q Median 3Q Max

-40989 -8087 -2068 3682 119814

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 88179 14558 6.06 2.4e-08 \*\*\*

s\_avg 4803 4673 1.03 0.31

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17900 on 101 degrees of freedom

Multiple R-squared: 0.0103, Adjusted R-squared: 0.000551

F-statistic: 1.06 on 1 and 101 DF, p-value: 0.307

> fit1$coefficients

(Intercept) s\_avg

88179 4803

> fitted(fit1)

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

104989 104028 104989 104508 104028 104508 104028 104028 105469 104028 104028 104508 104508 104028 104028 104989 106430 104749 104028 105469 104989 105949

57 58 59 60 61 62 63 64 65 66 67 68 69 115 116 117 118 119 120 121 122 123

104028 104989 105469 105469 105949 105469 104989 104989 106430 104989 104028 107390 105469 103068 103068 103068 103548 103884 103548 102587 103548 103548

124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 186 187 188 189 190 191

102587 103068 102587 103020 103068 103548 103548 103548 102587 102587 103020 103068 102587 103548 103548 103068 102107 101146 102107 101146 101146 102107

192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 256 257 258 259

102107 101627 102107 102107 101627 102107 101146 102107 102107 101627 101627 101627 101627 102155 102107 102107 102107 101627 104989 100186 104989 99706

260 261 262 263 264 265 266 267 268 269 270 271 272 273 274

102107 99706 104508 99706 100186 99225 100666 100666 100666 100666 99706 98745 100666 100666 100186

> confint(fit1)

2.5 % 97.5 %

(Intercept) 59299 117058

s\_avg -4468 14074

> residuals(fit1)

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

-19989 -19028 -18989 -16508 -12028 -11508 -9028 -9028 -10469 -8028 -8028 -4508 -4508 -4028 972 11 -1430 251 972 -469 1011 51

57 58 59 60 61 62 63 64 65 66 67 68 69 115 116 117 118 119 120 121 122 123

3472 3011 4531 6531 9051 9531 13011 15011 13570 15011 15972 38610 56531 -21068 -11068 -10068 -8548 -8884 -7548 -6087 -5548 -5548

124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 186 187 188 189 190 191

-4587 -4068 -2587 -3020 -2068 -548 452 1452 2413 2413 3980 8932 12413 11452 26452 42732 -23851 -12646 -12107 -11146 -8146 -7107

192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 256 257 258 259

-5107 -4627 -4107 -4107 -3627 -4107 -3146 -4107 -2107 -1627 -627 -527 873 2845 3893 5193 5893 10373 -40989 -23186 -19989 -14706

260 261 262 263 264 265 266 267 268 269 270 271 272 273 274

-16107 -9706 -12508 -4706 -4186 -1225 -666 -666 -266 934 4294 6255 14334 26044 119814

> summary(fit1)

Call:

lm(formula = salary ~ s\_avg, data = placed)

Residuals:

Min 1Q Median 3Q Max

-40989 -8087 -2068 3682 119814

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 88179 14558 6.06 2.4e-08 \*\*\*

s\_avg 4803 4673 1.03 0.31

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17900 on 101 degrees of freedom

Multiple R-squared: 0.0103, Adjusted R-squared: 0.000551

F-statistic: 1.06 on 1 and 101 DF, p-value: 0.307

> fit2 <- lm(salary ~ quarter, data = placed)

> summary(fit2)

Call:

lm(formula = salary ~ quarter, data = placed)

Residuals:

Min 1Q Median 3Q Max

-35468 -8543 -1518 3457 120532

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 107668 3970 27.1 <2e-16 \*\*\*

quarter -2050 1574 -1.3 0.2

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17800 on 101 degrees of freedom

Multiple R-squared: 0.0165, Adjusted R-squared: 0.00677

F-statistic: 1.7 on 1 and 101 DF, p-value: 0.196

> fit2$coefficients

(Intercept) quarter

107668 -2050

> fitted(fit2)

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618

57 58 59 60 61 62 63 64 65 66 67 68 69 115 116 117 118 119 120 121 122 123

105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 105618 103568 103568 103568 103568 103568 103568 103568 103568 103568

124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 186 187 188 189 190 191

103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 103568 101518 101518 101518 101518 101518 101518

192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 256 257 258 259

101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 101518 99468 99468 99468 99468

260 261 262 263 264 265 266 267 268 269 270 271 272 273 274

99468 99468 99468 99468 99468 99468 99468 99468 99468 99468 99468 99468 99468 99468 99468

> confint(fit2)

2.5 % 97.5 %

(Intercept) 99792 115545

quarter -5174 1073

> residuals(fit2)

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

-20618 -20618 -19618 -17618 -13618 -12618 -10618 -10618 -10618 -9618 -9618 -5618 -5618 -5618 -618 -618 -618 -618 -618 -618 382 382

57 58 59 60 61 62 63 64 65 66 67 68 69 115 116 117 118 119 120 121 122 123

1882 2382 4382 6382 9382 9382 12382 14382 14382 14382 14382 40382 56382 -21568 -11568 -10568 -8568 -8568 -7568 -7068 -5568 -5568

124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 186 187 188 189 190 191

-5568 -4568 -3568 -3568 -2568 -568 432 1432 1432 1432 3432 8432 11432 11432 26432 42232 -23262 -13018 -11518 -11518 -8518 -6518

192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 256 257 258 259

-4518 -4518 -3518 -3518 -3518 -3518 -3518 -3518 -1518 -1518 -518 -418 982 3482 4482 5782 6482 10482 -35468 -22468 -14468 -14468

260 261 262 263 264 265 266 267 268 269 270 271 272 273 274

-13468 -9468 -7468 -4468 -3468 -1468 532 532 932 2132 4532 5532 15532 27242 120532

> summary(fit2)

Call:

lm(formula = salary ~ quarter, data = placed)

Residuals:

Min 1Q Median 3Q Max

-35468 -8543 -1518 3457 120532

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 107668 3970 27.1 <2e-16 \*\*\*

quarter -2050 1574 -1.3 0.2

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17800 on 101 degrees of freedom

Multiple R-squared: 0.0165, Adjusted R-squared: 0.00677

F-statistic: 1.7 on 1 and 101 DF, p-value: 0.196

> fit <- lm(salary ~ age, data = placed)

> summary(fit)

Call:

lm(formula = salary ~ age, data = placed)

Residuals:

Min 1Q Median 3Q Max

-31454 -8533 -2182 4546 80886

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 29963 12698 2.36 0.02 \*

age 2729 471 5.80 7.7e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15600 on 101 degrees of freedom

Multiple R-squared: 0.25, Adjusted R-squared: 0.242

F-statistic: 33.6 on 1 and 101 DF, p-value: 7.75e-08

> fit$coefficients

(Intercept) age

29963 2729

> fitted(fit)

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

89996 103640 98182 98182 103640 106369 95454 98182 98182 98182 100911 92725 95454 103640 98182 98182 100911 100911 111826 114555 111826 111826

57 58 59 60 61 62 63 64 65 66 67 68 69 115 116 117 118 119 120 121 122 123

103640 98182 106369 136386 103640 103640 120013 103640 106369 111826 111826 139114 98182 100911 98182 114555 98182 98182 95454 95454 92725 95454

124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 186 187 188 189 190 191

100911 106369 95454 114555 98182 103640 106369 100911 103640 122742 100911 109098 95454 103640 100911 95454 92725 103640 98182 98182 98182 109098

192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 256 257 258 259

103640 106369 95454 98182 98182 103640 106369 109098 95454 98182 95454 109098 111826 95454 109098 117284 106369 95454 95454 92725 98182 100911

260 261 262 263 264 265 266 267 268 269 270 271 272 273 274

100911 103640 98182 98182 100911 95454 95454 100911 109098 100911 114555 92725 98182 100911 139114

> confint(fit)

2.5 % 97.5 %

(Intercept) 4774 55152

age 1795 3663

> residuals(fit)

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

-4996 -18640 -12182 -10182 -11640 -13369 -454 -3182 -3182 -2182 -4911 7275 4546 -3640 6818 6818 4089 4089 -6826 -9555 -5826 -5826

57 58 59 60 61 62 63 64 65 66 67 68 69 115 116 117 118 119 120 121 122 123

3860 9818 3631 -24386 11360 11360 -2013 16360 13631 8174 8174 6886 63818 -18911 -6182 -21555 -3182 -3182 546 1046 5275 2546

124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 186 187 188 189 190 191

-2911 -7369 4546 -14555 2818 -640 -2369 4089 1360 -17742 6089 2902 19546 11360 29089 50346 -14469 -15140 -8182 -8182 -5182 -14098

192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 256 257 258 259

-6640 -9369 2546 -182 -182 -5640 -8369 -11098 4546 1818 5546 -7998 -9326 9546 -3098 -9984 1631 16546 -31454 -15725 -13182 -15911

260 261 262 263 264 265 266 267 268 269 270 271 272 273 274

-14911 -13640 -6182 -3182 -4911 2546 4546 -911 -8698 689 -10555 12275 16818 25799 80886

> summary(fit)

Call:

lm(formula = salary ~ age, data = placed)

Residuals:

Min 1Q Median 3Q Max

-31454 -8533 -2182 4546 80886

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 29963 12698 2.36 0.02 \*

age 2729 471 5.80 7.7e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15600 on 101 degrees of freedom

Multiple R-squared: 0.25, Adjusted R-squared: 0.242

F-statistic: 33.6 on 1 and 101 DF, p-value: 7.75e-08

> notplaced <- mba[ which(mba$salary==0) , ]

> par(mfrow=c(1, 2))

> hist(placed$work\_yrs,

+ main="For placed students",

+ xlab="years of work experience",

+ ylab="Count",

+ breaks=2, # more columns

+ col="grey") # color the bars

> hist(notplaced$work\_yrs,

+ main="For not placed students",

+ xlab="years of work experience",

+ ylab="Count",

+ breaks=2, # more columns

+ col="grey") # color the bars

> par(mfrow=c(1, 1))



> hist(placed$s\_avg,

+ main="For placed students",

+ xlab="spring MBA average",

+ ylab="Count",

+ breaks=2, # more columns

+ col="grey") # color the bars

> hist(notplaced$s\_avg,

+ main="For not placed students",

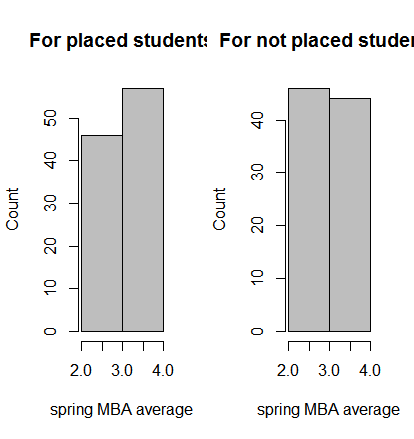
+ xlab="spring MBA average",

+ ylab="Count",

+ breaks=2, # more columns

+ col="grey") # color the bars

> par(mfrow=c(1, 1))



> mytable <- xtabs(~ work\_yrs+salary, data=notplaced)

> addmargins(mytable)

salary

work\_yrs 0 Sum

0 1 1

1 12 12

2 22 22

3 14 14

4 9 9

5 12 12

6 2 2

7 5 5

8 2 2

9 1 1

10 1 1

11 2 2

12 2 2

13 1 1

16 1 1

18 1 1

22 2 2

Sum 90 90

> mytable # frequencies

salary

work\_yrs 0

0 1

1 12

2 22

3 14

4 9

5 12

6 2

7 5

8 2

9 1

10 1

11 2

12 2

13 1

16 1

18 1

22 2

> chisq.test(mytable)

Chi-squared test for given probabilities

data: mytable

X-squared = 100, df = 20, p-value <2e-16

> mytable1 <- xtabs(~ age+salary, data=notplaced)

> mytable1 # frequencies

salary

age 0

22 1

23 3

24 13

25 9

26 10

27 14

28 6

29 11

30 2

31 2

32 5

34 3

35 3

36 2

37 1

39 1

42 1

43 2

48 1

> chisq.test(mytable1)

Chi-squared test for given probabilities

data: mytable1

X-squared = 70, df = 20, p-value = 1e-08

Warning message:

In chisq.test(mytable1) : Chi-squared approximation may be incorrect

> mytable2 <- xtabs(~ salary+quarter, data=notplaced)

> addmargins(mytable2)

quarter

salary 1 2 3 4 Sum

0 18 27 23 22 90

Sum 18 27 23 22 90

> mytable2 # frequencies

quarter

salary 1 2 3 4

0 18 27 23 22

> chisq.test(mytable2)

Chi-squared test for given probabilities

data: mytable2

X-squared = 2, df = 3, p-value = 0.6

> mytable3 <- xtabs(~ salary+s\_avg, data=notplaced)

> addmargins(mytable3)

s\_avg

salary 2 2.1 2.2 2.3 2.4 2.6 2.7 2.8 2.82 2.9 3 3.08 3.09 3.1 3.17 3.2

0 1 2 1 2 2 1 8 9 1 9 10 1 2 6 1 4

Sum 1 2 1 2 2 1 8 9 1 9 10 1 2 6 1 4

s\_avg

salary 3.25 3.27 3.3 3.38 3.4 3.45 3.5 3.6 3.64 3.8 3.9 Sum

0 1 2 9 1 7 1 2 4 1 1 1 90

Sum 1 2 9 1 7 1 2 4 1 1 1 90

> mytable3 # frequencies

s\_avg

salary 2 2.1 2.2 2.3 2.4 2.6 2.7 2.8 2.82 2.9 3 3.08 3.09 3.1 3.17 3.2

0 1 2 1 2 2 1 8 9 1 9 10 1 2 6 1 4

s\_avg

salary 3.25 3.27 3.3 3.38 3.4 3.45 3.5 3.6 3.64 3.8 3.9

0 1 2 9 1 7 1 2 4 1 1 1

> chisq.test(mytable3)

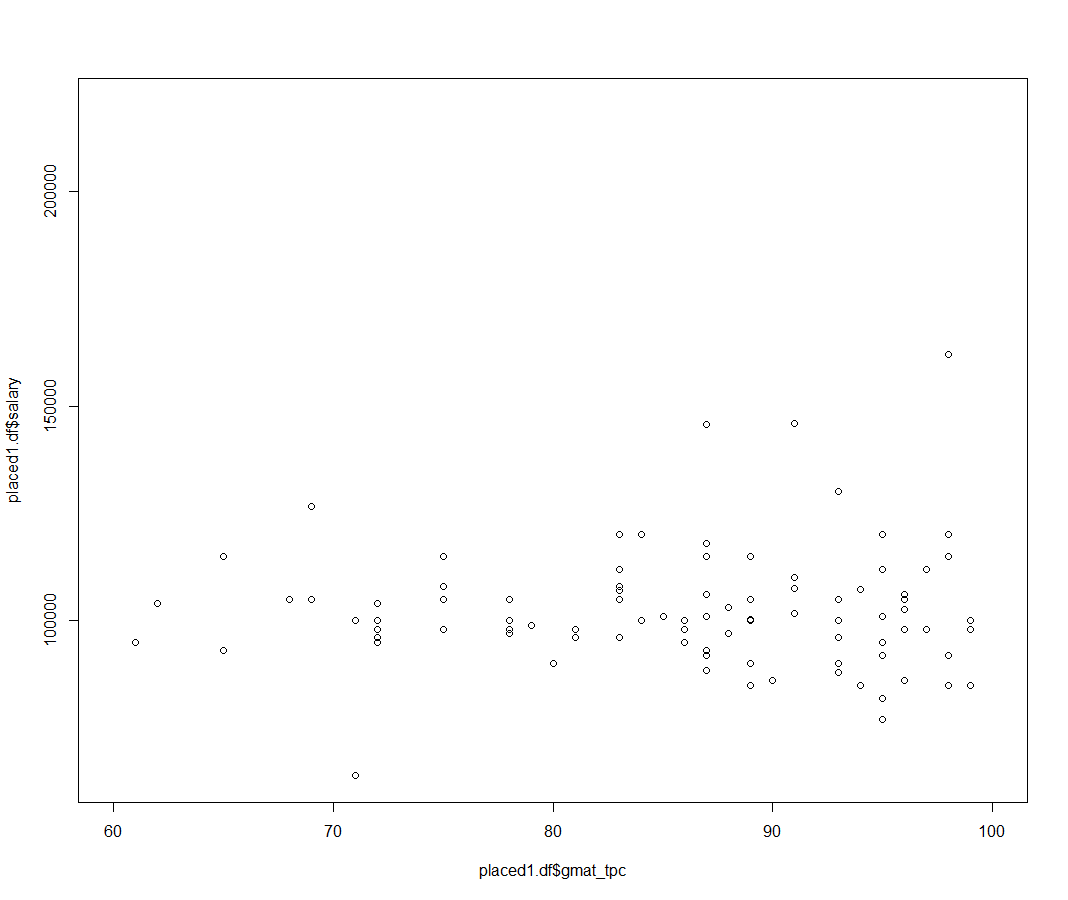
Chi-squared test for given probabilities

data: mytable3

X-squared = 80, df = 30, p-value = 4e-07

Warning message:

In chisq.test(mytable3) : Chi-squared approximation may be incorrect



> cor.test(placed1.df$gmat\_tpc,placed1.df$salary) #insignificant, low correlation

Pearson's product-moment correlation

data: placed1.df$gmat\_tpc and placed1.df$salary

t = -1.3385, df = 101, p-value = 0.1837

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.31743251 0.06311945

sample estimates:

cor

-0.1320178