Amountspend = a + b\* # children + c\*income + d\* catalogs + e

Nullhypo: no.of child has no effect on amt spend

P = 5.29\*10^-16

Actual b H0:= 0

P(b^ = -1.98\*10|Ho) = p = 0

So, Reject null hypo

Alpha =5%

Coefficients:

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

(Intercept) -4.428e+01 5.372e+00 -8.242 5.29e-16 \*\*\*

mydata$Children -1.987e+01 1.709e+00 -11.628 < 2e-16 \*\*\*

mydata$Income 2.041e-03 5.929e-05 34.417 < 2e-16 \*\*\*

mydata$Catalogs 4.770e+00 2.755e-01 17.310 < 2e-16 \*\*\*

amt spnt

= 150 150-6.956 age old-94.037 age young ppl

middle : 150-0-0

old:150-6.956

y:150-94..

best is middle

3 imp rules:

E(epsilon) = 0

Significant predictor can be hidden in epsilon

residual

(pet) ~

o

fitted(# children)

> mydata=DirectMarketing

> fit<-lm(mydata$AmountSpent~mydata$Children+mydata$Income+mydata$Catalogs)

> summary(fit)

Call:

lm(formula = mydata$AmountSpent ~ mydata$Children + mydata$Income +

mydata$Catalogs)

Residuals:

Min 1Q Median 3Q Max

-177.59 -34.87 -3.87 25.55 321.13

Coefficients:

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

(Intercept) -4.428e+01 5.372e+00 -8.242 5.29e-16 \*\*\*

mydata$Children -1.987e+01 1.709e+00 -11.628 < 2e-16 \*\*\*

mydata$Income 2.041e-03 5.929e-05 34.417 < 2e-16 \*\*\*

mydata$Catalogs 4.770e+00 2.755e-01 17.310 < 2e-16 \*\*\*

---

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

Residual standard error: 56.25 on 996 degrees of freedom

Multiple R-squared: 0.6584, Adjusted R-squared: 0.6574

F-statistic: 640 on 3 and 996 DF, p-value: < 2.2e-16

> plot(fit)

> layout(matrix(c(1,2,3,4),2,2))

> summary(mydata)

Age Gender OwnHome Married Location

Length:1000 Length:1000 Length:1000 Length:1000 Length:1000

Class :character Class :character Class :character Class :character Class :character

Mode :character Mode :character Mode :character Mode :character Mode :character

History Children Income Catalogs AmountSpent

Length:1000 Min. :0.000 Min. : 10100 Min. : 6.00 Min. : 3.80

Class :character 1st Qu.:0.000 1st Qu.: 29975 1st Qu.: 6.00 1st Qu.: 48.83

Mode :character Median :1.000 Median : 53700 Median :12.00 Median : 96.20

Mean :0.934 Mean : 56104 Mean :14.68 Mean :121.68

3rd Qu.:2.000 3rd Qu.: 77025 3rd Qu.:18.00 3rd Qu.:168.85

Max. :3.000 Max. :168800 Max. :24.00 Max. :621.70

> mydata$ageold=ifelse(mydata$Age=="Old",1,0)

> mydata$ageyoung=ifelse(mydata$Age=="Young",1,0)

> fit<-lm(mydata$AmountSpent~mydata$ageold+mydata$ageyoung,data=mydata)

> summary(fit)

Call:

lm(formula = mydata$AmountSpent ~ mydata$ageold + mydata$ageyoung,

data = mydata)

Residuals:

Min 1Q Median 3Q Max

-136.71 -54.61 -15.21 39.06 478.49

Coefficients:

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

(Intercept) 150.169 3.842 39.085 <2e-16 \*\*\*

mydata$ageold -6.956 7.165 -0.971 0.332

mydata$ageyoung -94.307 6.395 -14.748 <2e-16 \*\*\*

---

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

Residual standard error: 86.6 on 997 degrees of freedom

Multiple R-squared: 0.1897, Adjusted R-squared: 0.1881

F-statistic: 116.7 on 2 and 997 DF, p-value: < 2.2e-16

> fit<-lm(mydata$AmountSpent~mydata$ageold+mydata$ageyoung,data=mydata,

+ subset=(mydata$Gender=="Male"))

> summary(fit)

Call:

lm(formula = mydata$AmountSpent ~ mydata$ageold + mydata$ageyoung,

data = mydata, subset = (mydata$Gender == "Male"))

Residuals:

Min 1Q Median 3Q Max

-148.14 -56.69 -14.89 44.96 452.55

Coefficients:

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

(Intercept) 163.835 5.039 32.512 <2e-16 \*\*\*

mydata$ageold 5.316 11.238 0.473 0.636

mydata$ageyoung -99.516 9.566 -10.403 <2e-16 \*\*\*

---

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

Residual standard error: 87.57 on 491 degrees of freedom

Multiple R-squared: 0.1929, Adjusted R-squared: 0.1896

F-statistic: 58.66 on 2 and 491 DF, p-value: < 2.2e-16

>

>

> summary(fit)

Call:

lm(formula = mydata$AmountSpent ~ mydata$ageold + mydata$ageyoung,

data = mydata, subset = (mydata$Gender == "Male"))

Residuals:

Min 1Q Median 3Q Max

-148.14 -56.69 -14.89 44.96 452.55

Coefficients:

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

(Intercept) 163.835 5.039 32.512 <2e-16 \*\*\*

mydata$ageold 5.316 11.238 0.473 0.636

mydata$ageyoung -99.516 9.566 -10.403 <2e-16 \*\*\*

---

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

Residual standard error: 87.57 on 491 degrees of freedom

Multiple R-squared: 0.1929, Adjusted R-squared: 0.1896

F-statistic: 58.66 on 2 and 491 DF, p-value: < 2.2e-16

> mydata$Far=ifelse(mydata$Location=="Far",1,0)

> fit<-lm(mydata$AmountSpent~mydata$Income+mydata$Far)

> summary(fit)

Call:

lm(formula = mydata$AmountSpent ~ mydata$Income + mydata$Far)

Residuals:

Min 1Q Median 3Q Max

-204.50 -37.45 -0.65 26.17 331.52

Coefficients:

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

(Intercept) -2.048e+01 4.413e+00 -4.641 3.93e-06 \*\*\*

mydata$Income 2.229e-03 6.545e-05 34.050 < 2e-16 \*\*\*

mydata$Far 5.906e+01 4.414e+00 13.380 < 2e-16 \*\*\*

---

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

Residual standard error: 63.29 on 997 degrees of freedom

Multiple R-squared: 0.5672, Adjusted R-squared: 0.5663

F-statistic: 653.2 on 2 and 997 DF, p-value: < 2.2e-16

Am sp = -2.048e + 2.229e-03 \* income + 5.906e+01 \* far

Close = -2.048e + 2.229e-03 \* income as far =0

Far = -2.048e + 2.229e-03 \* income + 5.906e+01 \* 1 as far =1

am

Am= a +b. inc + c. far + d.inc\*far

= a +(b+d.far).inc + c.far

---------- -> slope