HW4

Yiming

2/13/2019

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the**Knit**button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

library("wooldridge")  
  
  
data <-bwght  
l1 <-lm(data$bwght~bwght$cigs+bwght$faminc)  
summary(l1)

##   
## Call:  
## lm(formula = data$bwght ~ bwght$cigs + bwght$faminc)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -96.061 -11.543 0.638 13.126 150.083   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 116.97413 1.04898 111.512 < 2e-16 \*\*\*  
## bwght$cigs -0.46341 0.09158 -5.060 4.75e-07 \*\*\*  
## bwght$faminc 0.09276 0.02919 3.178 0.00151 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.06 on 1385 degrees of freedom  
## Multiple R-squared: 0.0298, Adjusted R-squared: 0.0284   
## F-statistic: 21.27 on 2 and 1385 DF, p-value: 7.942e-10

l2 <-lm(data$bwght~bwght$cigs)  
summary(l2)

##   
## Call:  
## lm(formula = data$bwght ~ bwght$cigs)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -96.772 -11.772 0.297 13.228 151.228   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 119.77190 0.57234 209.267 < 2e-16 \*\*\*  
## bwght$cigs -0.51377 0.09049 -5.678 1.66e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.13 on 1386 degrees of freedom  
## Multiple R-squared: 0.02273, Adjusted R-squared: 0.02202   
## F-statistic: 32.24 on 1 and 1386 DF, p-value: 1.662e-08

data2 <-hprice1  
l3 <-lm(data2$price ~data2$sqrft +data2$bdrms)  
summary(l3)

##   
## Call:  
## lm(formula = data2$price ~ data2$sqrft + data2$bdrms)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -127.627 -42.876 -7.051 32.589 229.003   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -19.31500 31.04662 -0.622 0.536   
## data2$sqrft 0.12844 0.01382 9.291 1.39e-14 \*\*\*  
## data2$bdrms 15.19819 9.48352 1.603 0.113   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 63.04 on 85 degrees of freedom  
## Multiple R-squared: 0.6319, Adjusted R-squared: 0.6233   
## F-statistic: 72.96 on 2 and 85 DF, p-value: < 2.2e-16

data3 <-data.frame(discrim)  
mean(data3$prpblck, na.rm =TRUE)

## [1] 0.1134864

mean(data3$income, na.rm =TRUE)

## [1] 47053.78

sd(data3$prpblck, na.rm =TRUE)

## [1] 0.1824165

sd(data3$income, na.rm =TRUE)

## [1] 13179.29

l4 <-lm(data3$psoda ~data3$prpblck +data3$income)  
summary(l4)

##   
## Call:  
## lm(formula = data3$psoda ~ data3$prpblck + data3$income)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.29401 -0.05242 0.00333 0.04231 0.44322   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 9.563e-01 1.899e-02 50.354 < 2e-16 \*\*\*  
## data3$prpblck 1.150e-01 2.600e-02 4.423 1.26e-05 \*\*\*  
## data3$income 1.603e-06 3.618e-07 4.430 1.22e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.08611 on 398 degrees of freedom  
## (9 observations deleted due to missingness)  
## Multiple R-squared: 0.06422, Adjusted R-squared: 0.05952   
## F-statistic: 13.66 on 2 and 398 DF, p-value: 1.835e-06

l5 <-lm(data3$psoda ~data3$prpblck)  
summary(l5)

##   
## Call:  
## lm(formula = data3$psoda ~ data3$prpblck)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.30884 -0.05963 0.01135 0.03206 0.44840   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.03740 0.00519 199.87 < 2e-16 \*\*\*  
## data3$prpblck 0.06493 0.02396 2.71 0.00702 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0881 on 399 degrees of freedom  
## (9 observations deleted due to missingness)  
## Multiple R-squared: 0.01808, Adjusted R-squared: 0.01561   
## F-statistic: 7.345 on 1 and 399 DF, p-value: 0.007015

l6 <-lm(data3$psoda ~data3$prpblck +log(data3$income))  
summary(l6)

##   
## Call:  
## lm(formula = data3$psoda ~ data3$prpblck + log(data3$income))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.29484 -0.05085 0.00346 0.04283 0.44069   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.18553 0.18800 0.987 0.324   
## data3$prpblck 0.12583 0.02697 4.665 4.23e-06 \*\*\*  
## log(data3$income) 0.07882 0.01739 4.533 7.71e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.08602 on 398 degrees of freedom  
## (9 observations deleted due to missingness)  
## Multiple R-squared: 0.06628, Adjusted R-squared: 0.06159   
## F-statistic: 14.13 on 2 and 398 DF, p-value: 1.184e-06

l7 <-lm(data3$psoda ~data3$prpblck +log(data3$income) +data3$prppov)  
summary(l7)

##   
## Call:  
## lm(formula = data3$psoda ~ data3$prpblck + log(data3$income) +   
## data3$prppov)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.28083 -0.05006 0.00305 0.04247 0.44286   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.51208 0.30777 -1.664 0.09693 .   
## data3$prpblck 0.07501 0.03214 2.334 0.02011 \*   
## log(data3$income) 0.14180 0.02804 5.058 6.5e-07 \*\*\*  
## data3$prppov 0.39629 0.13915 2.848 0.00463 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.08526 on 397 degrees of freedom  
## (9 observations deleted due to missingness)  
## Multiple R-squared: 0.08497, Adjusted R-squared: 0.07806   
## F-statistic: 12.29 on 3 and 397 DF, p-value: 1.053e-07

cor\_data<-na.omit(cbind.data.frame(log(data3$income), data3$prppov))  
cor(cor\_data, method =c("pearson", "kendall", "spearman"))

## log(data3$income) data3$prppov  
## log(data3$income) 1.000000 -0.838467  
## data3$prppov -0.838467 1.000000

