multi_reg_2.R

SANGHOOJEFFREY

Tue Jun 26 19:46:55 2018

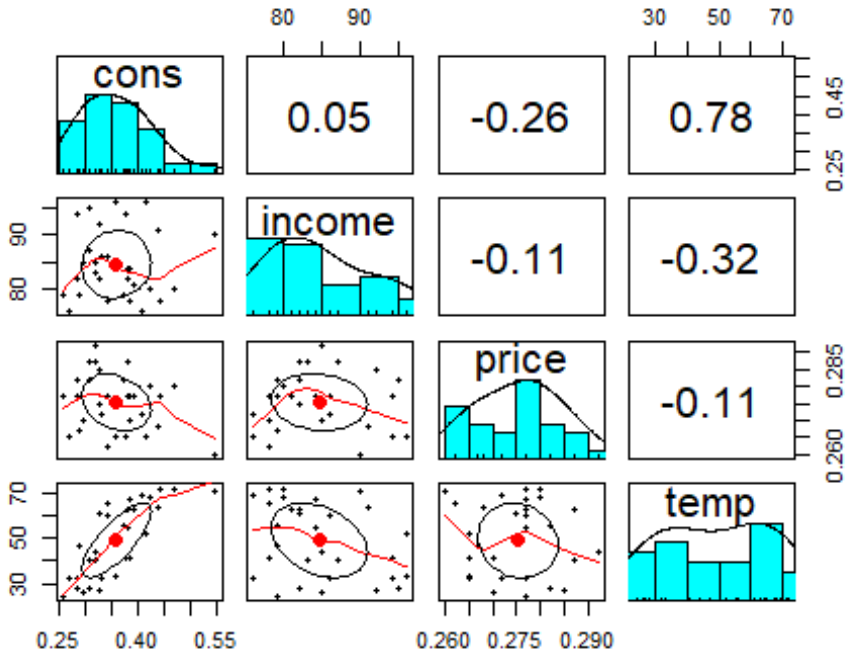
```
# 아이스크림 섭취량과 아이스크림의 가격, 가족의 월수입, 평균기온의 관계를 연구하기 위해
# 다음의 자료를 수집했다.
setwd("F:/work/2018/0. R-work/2018_R_intermediate")
load("data/icecream.rda")

if (!require(psych)) install.packages("psych"); library(psych)

## Loading required package: psych

## Warning: package 'psych' was built under R version 3.4.4

pairs.panels(ice[,1:4])
```




```
r.full = lm(cons~price+income+temp, data=ice)
summary(r.full)
```

```
##
## Call:
## lm(formula = cons ~ price + income + temp, data = ice)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.065302 -0.011873  0.002737  0.015953  0.078986
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.1973151  0.2702162   0.730  0.47179
## price       -1.0444140  0.8343573  -1.252  0.22180
## income       0.0033078  0.0011714   2.824  0.00899 **
## temp         0.0034584  0.0004455   7.762  3.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03683 on 26 degrees of freedom
## Multiple R-squared:  0.719,  Adjusted R-squared:  0.6866
## F-statistic: 22.17 on 3 and 26 DF,  p-value: 2.451e-07
```

```
# model selection-stepwise
step(r.full, direction="both")
```

```
## Start:  AIC=-194.38
## cons ~ price + income + temp
##
##          Df Sum of Sq      RSS     AIC
## - price   1  0.002126 0.037399 -194.62
## <none>                0.035273 -194.38
## - income  1  0.010817 0.046090 -188.35
## - temp    1  0.081741 0.117013 -160.40
##
## Step:  AIC=-194.62
## cons ~ income + temp
##
##          Df Sum of Sq      RSS     AIC
## <none>                0.037399 -194.62
## + price   1  0.002126 0.035273 -194.38
## - income  1  0.012611 0.050009 -187.90
## - temp    1  0.087836 0.125235 -160.36

##
## Call:
## lm(formula = cons ~ income + temp, data = ice)
##
## Coefficients:
## (Intercept)       income         temp
##   -0.113195     0.003530     0.003543
```

```r
r.reduce = lm(cons~income+temp, data=ice)

anova(r.full, r.reduce) # 두 모형에 차이가 없다. price 설명변수를 제거해도 된다.
```

```
## Analysis of Variance Table
##
## Model 1: cons ~ price + income + temp
## Model 2: cons ~ income + temp
##   Res.Df      RSS Df  Sum of Sq      F Pr(>F)
## 1     26 0.035273
## 2     27 0.037399 -1 -0.0021257 1.5669 0.2218
```

```r
# Q1. 다중공선성 확인

# Q2. 정규성 확인

# Q3. 독립성 확인

# Q4. 등분산성 확인

# Q5. 영향점, 이상점, 지렛대점 확인

# Q6. 이상점을 제거한 회귀모형은?
```