Newfood

ECM

10/10/2022

```
library(car)
```

```
## Loading required package: carData
```

```
setwd("/Users/ecm/teach/data")
newfood = read.csv("newfood.csv")
names(newfood)
```

```
## [1] "sales"  "price"  "ad"     "loc"    "income" "volume" "city"
```

```
head(newfood)
```

```
##   sales price ad loc income volume city
## 1   225    24  0   0    7.3     34    3
## 2   323    24  0   0    8.3     41    4
## 3   424    24  1   0    6.9     32    1
## 4   268    24  1   0    6.5     28    2
## 5   224    24  0   1    7.3     34    3
## 6   331    24  0   1    8.3     41    4
```

```
summary(newfood)
```

```
##      sales           price          ad           loc          income
##  Min.   :128.0   Min.   :24   Min.   :0.0   Min.   :0.0   Min.   :6.10
##  1st Qu.:162.5   1st Qu.:24   1st Qu.:0.0   1st Qu.:0.0   1st Qu.:6.50
##  Median :224.5   Median :29   Median :0.5   Median :0.5   Median :6.75
##  Mean   :236.0   Mean   :29   Mean   :0.5   Mean   :0.5   Mean   :7.05
##  3rd Qu.:273.0   3rd Qu.:34   3rd Qu.:1.0   3rd Qu.:1.0   3rd Qu.:7.50
##  Max.   :492.0   Max.   :34   Max.   :1.0   Max.   :1.0   Max.   :8.40
##      volume          city
##  Min.   :19.00   Min.   :1.00
##  1st Qu.:28.75   1st Qu.:1.75
##  Median :32.00   Median :2.50
##  Mean   :32.33   Mean   :2.50
##  3rd Qu.:37.00   3rd Qu.:3.25
##  Max.   :43.00   Max.   :4.00
```

1. Scatterplot matrix

```
plot(newfood, cex=.5, pch=16)
```
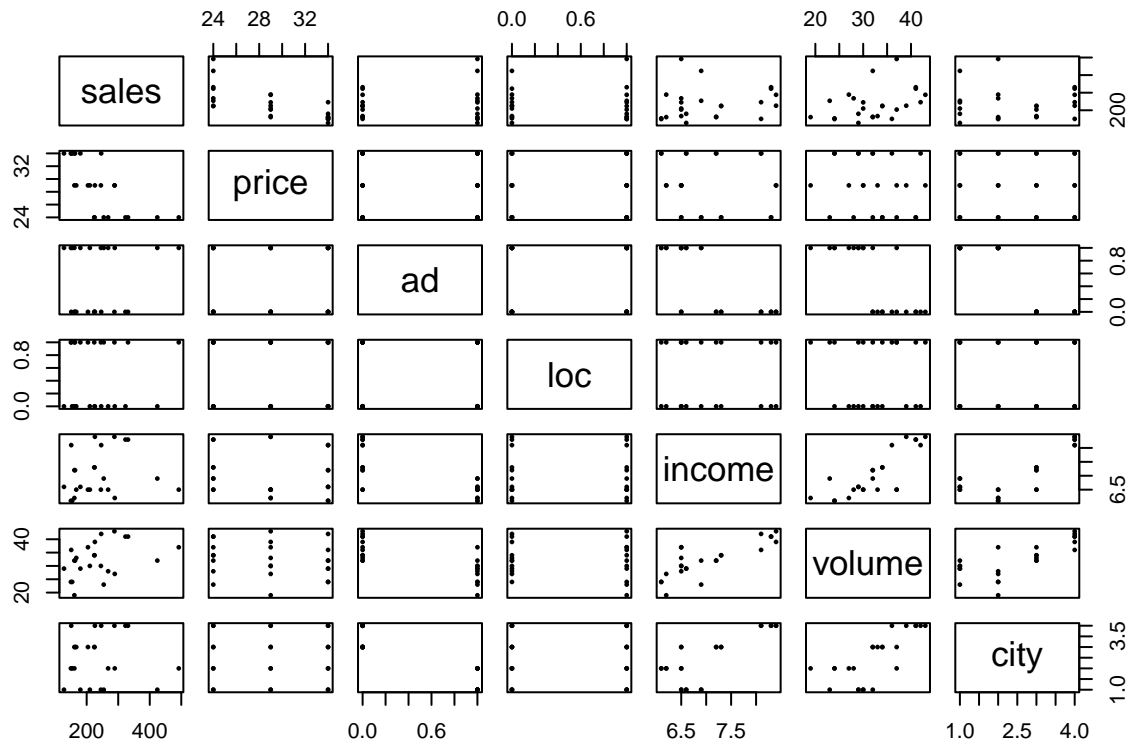
2. Crosstab showing design

```
ans = with(newfood, table(city, ad, loc, price))
ftable(ans, row.vars=1:2, cols.vars = 3:4)
```

```
##          loc    0          1
##          price 24 29 34 24 29 34
## city ad
## 1    0          0  0  0  0  0  0
##      1          1  1  1  1  1  1
## 2    0          0  0  0  0  0  0
##      1          1  1  1  1  1  1
## 3    0          1  1  1  1  1  1
##      1          0  0  0  0  0  0
## 4    0          1  1  1  1  1  1
##      1          0  0  0  0  0  0
```

3. Correlation matrix

```
round(cor(newfood), 3)
```

```
##         sales  price     ad   loc income volume   city
## sales   1.000 -0.704  0.117  0.01  0.186  0.393  0.000
## price  -0.704  1.000  0.000  0.00 -0.131 -0.179  0.000
## ad      0.117  0.000  1.000  0.00 -0.746 -0.742 -0.894
## loc     0.010  0.000  0.000  1.00  0.000 -0.040  0.000
## income  0.186 -0.131 -0.746  0.00  1.000  0.809  0.792
## volume  0.393 -0.179 -0.742 -0.04  0.809  1.000  0.741
```

```
## city     0.000   0.000 -0.894   0.00   0.792   0.741   1.000
```

Note that cor(ad, volume) = -0.742. Why?

4. Price model

```
summary(lm(sales ~ price, newfood))
```

```
##
## Call:
## lm(formula = sales ~ price, data = newfood)
##
## Residuals:
##     Min     1Q Median     3Q    Max
## -87.50 -35.25  -8.00  19.50 180.50
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  673.900     95.134   7.084 4.18e-07 ***
## price        -15.100      3.248  -4.648 0.000124 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 64.97 on 22 degrees of freedom
## Multiple R-squared:  0.4955, Adjusted R-squared:  0.4726
## F-statistic: 21.61 on 1 and 22 DF,  p-value: 0.0001238
```

$$\hat{y} = 674 - 15.1^{***}P$$

5. Price + ad model

```
summary(lm(sales ~ price+ad, newfood))
```

```
##
## Call:
## lm(formula = sales ~ price + ad, data = newfood)
##
## Residuals:
##     Min     1Q Median     3Q    Max
## -85.25 -45.50  -0.50  23.75 170.25
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  663.650     96.971   6.844 9.14e-07 ***
## price        -15.100      3.279  -4.604 0.000153 ***
## ad            20.500     26.777   0.766 0.452438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 65.59 on 21 degrees of freedom
## Multiple R-squared:  0.5092, Adjusted R-squared:  0.4625
## F-statistic: 10.89 on 2 and 21 DF,  p-value: 0.0005683
```

$$\hat{y} = 674 - 15.1^{***}P$$

$$\hat{y} = 664 - 15.1^{***}P + 20.5\text{ad}$$

6. Price + ad + location model

```
summary(lm(sales ~ price+ad+loc, newfood))
```

```
##
## Call:
## lm(formula = sales ~ price + ad + loc, data = newfood)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -86.167 -44.583  -1.167  24.208 169.333
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  662.733    100.297   6.608 1.95e-06 ***
## price        -15.100      3.360  -4.494 0.000222 ***
## ad            20.500     27.435   0.747 0.463623
## loc            1.833     27.435   0.067 0.947385
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 67.2 on 20 degrees of freedom
## Multiple R-squared:  0.5093, Adjusted R-squared:  0.4357
## F-statistic: 6.919 on 3 and 20 DF,  p-value: 0.002227
```

$$\hat{y} = 674 - 15.1^{***}P$$

$$\hat{y} = 664 - 15.1^{***}P + 20.5\text{ad}$$

$$\hat{y} = 663 - 15.1^{***}P + 20.5\text{ad} + 1.83\text{loc}$$

7. Price + ad + location + volume model

```
fit = lm(sales ~ price+ad+loc+volume, newfood)
summary(fit)
```

```
##
## Call:
## lm(formula = sales ~ price + ad + loc + volume, data = newfood)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -71.013 -25.921  -6.456  16.743  71.904
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  125.931    116.274   1.083    0.292
## price        -11.836      2.217  -5.340 3.74e-05 ***
```

```
## ad               131.283      26.587    4.938 9.13e-05 ***
## loc                 7.768      17.475    0.445    0.662
## volume             11.870       2.150    5.521 2.52e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 42.72 on 19 degrees of freedom
## Multiple R-squared:  0.8116, Adjusted R-squared:  0.7719
## F-statistic: 20.46 on 4 and 19 DF,  p-value: 1.132e-06
```

```
vif(fit)
```

```
##    price        ad       loc    volume
## 1.076598 2.323486 1.003798 2.403883
```

$$\hat{y} = 674 - 15.1^{***}P$$

$$\hat{y} = 664 - 15.1^{***}P + 20.5\text{ad}$$

$$\hat{y} = 663 - 15.1^{***}P + 20.5\text{ad} + 1.83\text{loc}$$

$$\hat{y} = 126 - 11.8^{***}P + 131^{***}\text{ad} + 7.77\text{loc} + 11.9^{***}V$$