Hw04ST430Yu

Haozhe (Jerry) Yu

2023-10-13

Question 1A

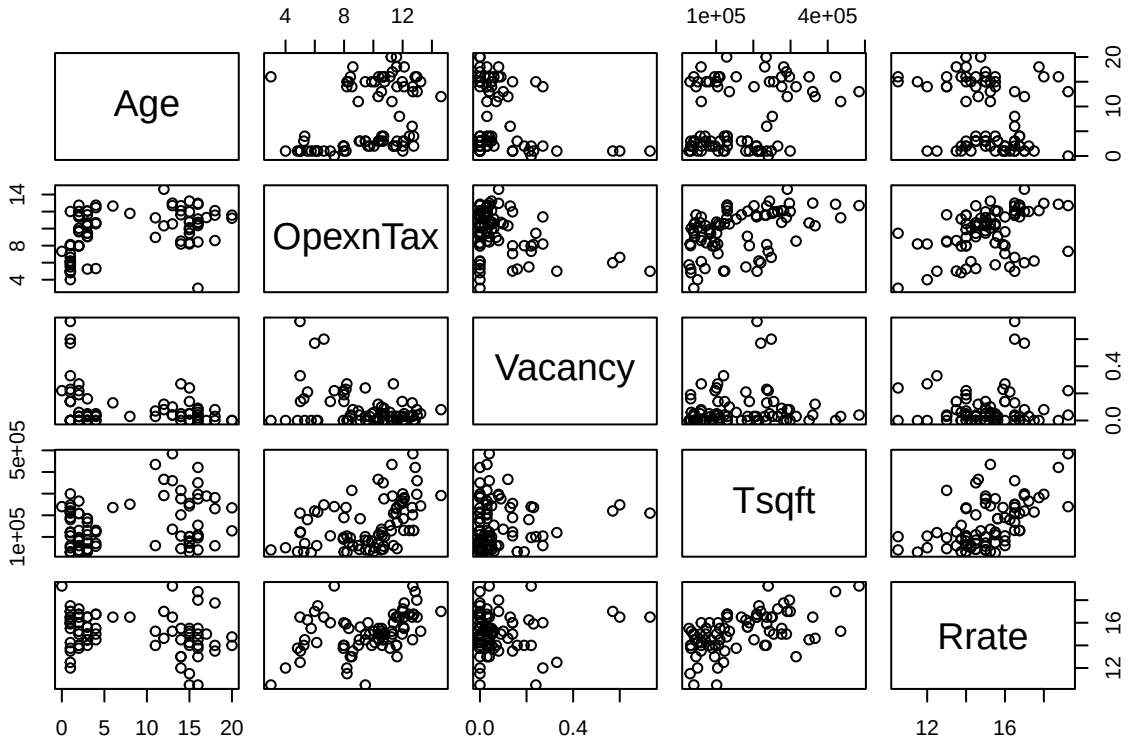
Problem1A: For this problem, use the Commercial Property data set from KNNL Problem 6.18. The response variable is rental rates. The explanatory variables are age, operating expenses and taxes, vacancy rates, and total square footage.

```
Comm <- as_tibble(read.table("https://users.stat.ufl.edu/~rrandles/sta4210/Rclassnotes/data/textdatasets
            #sep = "",
            strip.white=TRUE,
            col.name = c("Rrate","Age","OpexnTax","Vacancy","Tsqft")
            ))
Comm
```

```
## # A tibble: 81 x 5
##    Rrate   Age OpexnTax Vacancy  Tsqft
##    <dbl> <int>    <dbl>   <dbl>  <int>
##  1  13.5     1     5.02    0.14 123000
##  2  12      14     8.19    0.27 104079
##  3  10.5    16     3       0     39998
##  4  15       4    10.7     0.05  57112
##  5  14      11     8.97    0.07  60000
##  6  10.5    15     9.45    0.24 101385
##  7  14       2     8       0.19  31300
##  8  16.5     1     6.62    0.6  248172
##  9  17.5     1     6.2     0    215000
## 10  16.5     8    11.8     0.03 251015
## # i 71 more rows
```

1) Obtain a scatterplot matrix for all 5 variables (age, operating expenses and taxes, vacancy rates, total square footage, and rental rates) and give your comments about linearity of variables.

```
pairs(Comm[, c('Age', 'OpexnTax', 'Vacancy', 'Tsqft', 'Rrate')])
```



While the majority of relationships in the matrix appear nonlinear, there are some that might have a positive linear association (like Rental Rates and Total Square Feet and Operating Expenses and Operating and Rate Expenses).

2) Find the correlation of all pairs of variables and give your comments about the association between variables.

```
CommCor <- cor(Comm[, c('Age', 'OpexnTax', 'Vacancy', 'Tsqft', 'Rrate')])

CommCor
```

```
##                Age   OpexnTax     Vacancy      Tsqft       Rrate
## Age       1.0000000  0.3888264 -0.25266347 0.28858350 -0.25028456
## OpexnTax  0.3888264  1.0000000 -0.37976174 0.44069713  0.41378716
## Vacancy  -0.2526635 -0.3797617  1.00000000 0.08061073  0.06652647
## Tsqft     0.2885835  0.4406971  0.08061073 1.00000000  0.53526237
## Rrate    -0.2502846  0.4137872  0.06652647 0.53526237  1.00000000
```

There seem to be weak correlations between many variables. The strongest negative correlations seems to be between Vacancy Rate and Operating Expenses, and the strongest positive correlation seems to be between Total square Feet and Rental Rate. Like shown on the graphs, there seem to be modestly strong correlations between rental rate and Age and Vacancy rate.

3) Run the multiple regressions with age, operating expenses and taxes, vacancy rates, and total square footage as the explanatory variables and rental rates as the response variable.

```
Commm <- lm(Rrate ~ Age + OpexnTax + Vacancy + Tsqft, data=Comm)
summary(Commm)
```

```
##
## Call:
## lm(formula = Rrate ~ Age + OpexnTax + Vacancy + Tsqft, data = Comm)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1872 -0.5911 -0.0910  0.5579  2.9441
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.220e+01  5.780e-01  21.110  < 2e-16 ***
## Age         -1.420e-01  2.134e-02  -6.655 3.89e-09 ***
## OpexnTax     2.820e-01  6.317e-02   4.464 2.75e-05 ***
## Vacancy      6.193e-01  1.087e+00   0.570     0.57
## Tsqft        7.924e-06  1.385e-06   5.722 1.98e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.137 on 76 degrees of freedom
## Multiple R-squared:  0.5847, Adjusted R-squared:  0.5629
## F-statistic: 26.76 on 4 and 76 DF,  p-value: 7.272e-14
```

a.) Give the fitted (Estimated) regression equation.

Rental Rate = 1.220e+01 + -1.420e-01*Age + 2.820e-01*Operating Expenses and Taxes + 6.193e-01 *Vacancy Rate + 7.924e-06*Total Square Foot.

b) Interpret the estimated parameters in the context of the problem.

For each incremental increase in one year of age, for a given level of operating expenses, vacancy rate, and total square footage, the rental rate increases by -1.420e-01.

For each incremental increase of $1 in operating expenses and taxes, for a given age, vacancy rate, and total square footage, the rental rate increases by 2.820e-01.

For each incremental increase of 1% in vacancy rate, for a given age, operating expenses and taxes, and total square footage, the rental rate increases by 6.193e-01.

For each incremental increase of 1 square feet in total square footage, for a given age, operating expenses and taxes, and vacancy rate, the rental rate increases by 7.924e-06.

c)c) Give the value of R2 and Adj R2

Multiple R-squared: 0.5847, Adjusted R-squared: 0.5629