Tu-Linear-02

Sicheng Peng, Hajar El Mesaoudi

2023/7/4

```r
library("readxl")
linear_data = read_excel("./Tutorial/1_LinearRegression_Tutorial_Data.xlsx")
```

We squared Ctenure and Mtenure and used both as DVs. The result shows both them are significant drivers for financial profit.

```r
Ctenure2 = linear_data$Ctenure^2
Mtenure2 = linear_data$Mtenure^2
performancelmsq = lm(Profit~Ctenure+Ctenure2+Mtenure+Mtenure2+Pop+Comp+Visible+PedCount+Res+Hours24+Crew
summary(performancelmsq)
```

```
##
## Call:
## lm(formula = Profit ~ Ctenure + Ctenure2 + Mtenure + Mtenure2 +
##     Pop + Comp + Visible + PedCount + Res + Hours24 + CrewSkill +
##     MgrSkill + ServiceQuality, data = linear_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -99094 -27677  -4817  28339 102418
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -79305.417  92141.801  -0.861 0.392780
## Ctenure          2909.344    948.874   3.066 0.003230 **
## Ctenure2          -21.687      9.840  -2.204 0.031319 *
## Mtenure          1509.284    304.652   4.954 6.08e-06 ***
## Mtenure2           -3.955      1.384  -2.858 0.005825 **
## Pop                 3.560      1.361   2.617 0.011177 *
## Comp           -27112.103   4817.217  -5.628 4.88e-07 ***
## Visible         11123.893   8323.300   1.336 0.186357
## PedCount        39131.188   7837.037   4.993 5.27e-06 ***
## Res             50892.746  35556.162   1.431 0.157438
## Hours24         61433.631  17267.683   3.558 0.000731 ***
## CrewSkill      -29072.466  16388.600  -1.774 0.081063 .
## MgrSkill        33534.580  16107.128   2.082 0.041548 *
## ServiceQuality    853.431    503.420   1.695 0.095124 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 47690 on 61 degrees of freedom
```

```
## Multiple R-squared:  0.7655, Adjusted R-squared:  0.7155
## F-statistic: 15.31 on 13 and 61 DF,  p-value: 1.321e-14
```

Based on the previous results, we find Visible may not be a relative DV for profit. We removed Visible from the linear regression function and the results had no major changes. Hence we conclude Visible is not a significant driver for profit.

```
performance = lm(Profit~Ctenure+Ctenure2+Mtenure+Mtenure2+Pop+Comp+PedCount+Res+Hours24+CrewSkill+MgrSki
summary(performance)
```

```
##
## Call:
## lm(formula = Profit ~ Ctenure + Ctenure2 + Mtenure + Mtenure2 +
##     Pop + Comp + PedCount + Res + Hours24 + CrewSkill + MgrSkill +
##     ServiceQuality, data = linear_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -113036  -28239   -3942   30624  102919
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -40635.689  88033.914  -0.462 0.645990
## Ctenure          3085.505    945.613   3.263 0.001795 **
## Ctenure2          -23.565      9.801  -2.404 0.019201 *
## Mtenure          1426.830    300.225   4.753 1.23e-05 ***
## Mtenure2           -3.439      1.337  -2.572 0.012534 *
## Pop                 3.820      1.355   2.818 0.006469 **
## Comp           -26957.924   4846.272  -5.563 6.02e-07 ***
## PedCount        37375.646   7775.000   4.807 1.01e-05 ***
## Res             48919.394  35750.016   1.368 0.176133
## Hours24         63943.198  17273.769   3.702 0.000457 ***
## CrewSkill      -33486.271  16153.844  -2.073 0.042339 *
## MgrSkill        34879.274  16177.272   2.156 0.034966 *
## ServiceQuality    936.987    502.679   1.864 0.067060 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 47990 on 62 degrees of freedom
## Multiple R-squared:  0.7586, Adjusted R-squared:  0.7119
## F-statistic: 16.23 on 12 and 62 DF,  p-value: 7.31e-15
```