Tu-Linear-01

Sicheng Peng, Hajar El Mesaoudi

2023/7/4

In order not to be bothered with rounding the numbers, set `options(digits=3)`

Exercise 01

Install package dependencies and read dataset for tutorial 1 as linear_data

```
#install.packages("readxl")
library("readxl")
linear_data = read_excel("./Tutorial/1_LinearRegression_Tutorial_Data.xlsx")
```

We perform linear regression on profit, using crew tenure and manager tenure as DVs. The result p-value for Mtenure is lower than 0.05, thus H0 is accepted, manager tenure is a driver of store level financial performance. The result p-value for Ctenure is higher than 0.1, thus H0 is rejected, customer tenure is not a significant driver of store level financial performance.

```
performancelm = lm(linear_data$Profit~linear_data$Mtenure+linear_data$Ctenure)
summary(performancelm)
```

```
##
## Call:
## lm(formula = linear_data$Profit ~ linear_data$Mtenure + linear_data$Ctenure)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -165442  -49679   -7474   48733  194710
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)            236951      13125   18.05  < 2e-16 ***
## linear_data$Mtenure       620        167    3.72  0.00039 ***
## linear_data$Ctenure       810        543    1.49  0.14024
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 80200 on 72 degrees of freedom
## Multiple R-squared:  0.217,  Adjusted R-squared:  0.195
## F-statistic: 9.97 on 2 and 72 DF,  p-value: 0.000151
```

Exercise 02

We try different DVs. Result shows DV Visible and CrewSkill are not significant in related to financial performance, ServiceQuality is partially relative. The rest DVs are all significant drivers of financial performance.

```
totalperformancelm = lm(Profit~Mtenure+Ctenure+Pop+Comp+Visible+PedCount+Res+Hours24+CrewSkill+MgrSkill+

summary(totalperformancelm)
```

```
##
## Call:
## lm(formula = Profit ~ Mtenure + Ctenure + Pop + Comp + Visible +
##     PedCount + Res + Hours24 + CrewSkill + MgrSkill + ServiceQuality,
##     data = linear_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -98617 -36089  -7220  35118 114183
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.41e+05   9.82e+04   -1.44  0.15475
## Mtenure         6.84e+02   1.19e+02    5.72  3.1e-07 ***
## Ctenure         8.70e+02   4.00e+02    2.17  0.03350 *
## Pop             3.04e+00   1.41e+00    2.16  0.03450 *
## Comp           -2.85e+04   5.17e+03   -5.52  6.7e-07 ***
## Visible         7.04e+03   8.61e+03    0.82  0.41616
## PedCount        3.44e+04   8.31e+03    4.14  0.00011 ***
## Res             8.21e+04   3.64e+04    2.26  0.02742 *
## Hours24         6.28e+04   1.85e+04    3.39  0.00122 **
## CrewSkill      -1.41e+04   1.71e+04   -0.83  0.41137
## MgrSkill        4.52e+04   1.68e+04    2.70  0.00896 **
## ServiceQuality  9.45e+02   5.45e+02    1.73  0.08775 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 51800 on 63 degrees of freedom
## Multiple R-squared:  0.714,  Adjusted R-squared:  0.664
## F-statistic: 14.3 on 11 and 63 DF,  p-value: 2.77e-13
```