MMM_MarketMixModeling_MultipleLinearRegression

RohitDhankar

24/05/2020

This is sample code for the task being done along with Rohan Mathur = https://github.com/RohanMathur17
https://github.com/digital-cognition-co-in/DigitalCognition/issues/24

```
library(readr)
LungCapData <- read_csv("LungCapData.csv")
```

```
## Parsed with column specification:
## cols(
##   LungCap_cc = col_double(),
##   Age_years = col_double(),
##   Height_inches = col_double(),
##   Smoke = col_character(),
##   Gender = col_character(),
##   Caesarean = col_character()
## )
```

```
attach(LungCapData)
View(LungCapData)
names(LungCapData);head(LungCapData)
```

```
## [1] "LungCap_cc"    "Age_years"     "Height_inches" "Smoke"
## [5] "Gender"        "Caesarean"
```

```
## # A tibble: 6 x 6
##   LungCap_cc Age_years Height_inches Smoke Gender Caesarean
##        <dbl>     <dbl>         <dbl> <chr> <chr>  <chr>
## 1       6.48         6          62.1 no    male   no
## 2      10.1         18          74.7 yes   female no
## 3       9.55        16          69.7 no    female yes
## 4      11.1         14          71   no    male   no
## 5       4.8          5          56.9 no    male   no
## 6       6.22        11          58.7 no    female no
```

```
#
```

For the Multiple Linear Regression that we are performing the - Dependent Variable == LungCapacity in CC
The Multiple Independent vartiables are == Age() , Height() , Smoke() , Gender() , Caesarean()

We fit initial Linear Regression Model with Two Independent variables == Age() and Height()

```
init_multiple_linear_m <- lm(LungCap_cc ~ Age_years + Height_inches)
#typeof(init_multiple_linear_m) # list
#class(init_multiple_linear_m) # lm - Linear Model
summary(init_multiple_linear_m)
```

```
##
```

```
## Call:
## lm(formula = LungCap_cc ~ Age_years + Height_inches)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4080 -0.7097 -0.0078  0.7167  3.1679
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -11.747065   0.476899 -24.632  < 2e-16 ***
## Age_years       0.126368   0.017851   7.079 3.45e-12 ***
## Height_inches   0.278432   0.009926  28.051  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.056 on 722 degrees of freedom
## Multiple R-squared:  0.843,  Adjusted R-squared:  0.8425
## F-statistic:  1938 on 2 and 722 DF,  p-value: < 2.2e-16
```

```
# the - Multiple R-squared:  0.843 -- 84.3% Variability in LUNG CAPACITY can be
# explained by the linear relationship between - Age_years + Height_inches and LUNG CAPACITY
```

CORRELATION seen through SCATTER PLOTS

As seen below the - Height_inches is Positivly Correlated to Lung Capacity

```
plot(LungCap_cc~Height_inches)
```