R Notebook

GUOTAI SUN

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.0.5
```

```
## -- Attaching packages --------------------------------------- tidyve
rse 1.3.1 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## Warning: package 'tibble' was built under R version 4.0.5
```

```
## Warning: package 'tidyr' was built under R version 4.0.5
```

```
## Warning: package 'dplyr' was built under R version 4.0.5
```

```
## -- Conflicts ------------------------------------------ tidyverse_co
nflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
school <- read_csv("ICE1_Data.csv")
```

```
##
## -- Column specification --------------------------------------------
------------
## cols(
##   DBN = col_character(),
##   Quality_Review_Score = col_character(),
##   `Progress_Rpt_10-11` = col_character(),
##   `Student_Progress_10-11` = col_character(),
##   `graduation 2010-11` = col_double(),
##   `college enroll 2010-11` = col_double()
## )
```

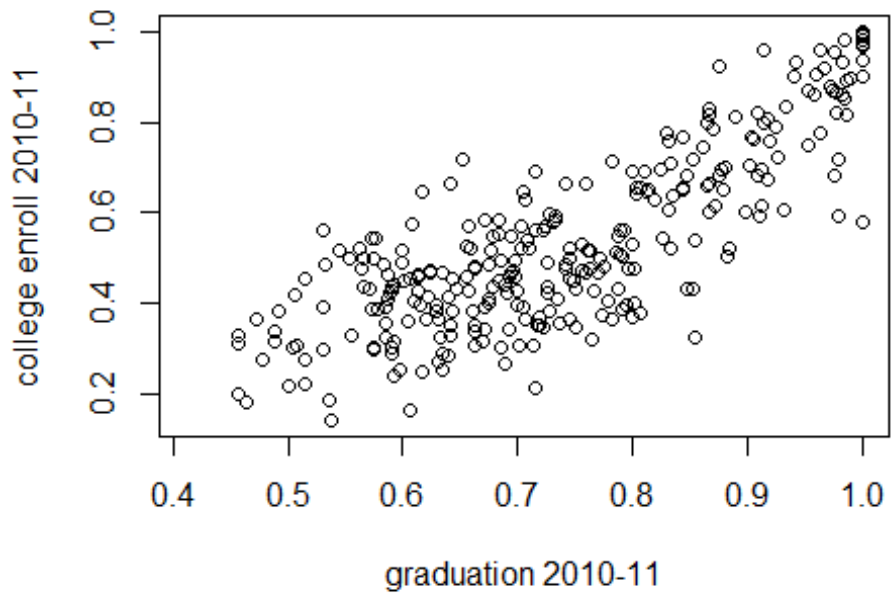
```
school
```

```
## # A tibble: 422 x 6
##    DBN    Quality_Review_S~ `Progress_Rpt_10~ `Student_Progres~ `gra
duation 201~
##    <chr>  <chr>             <chr>             <chr>
       <dbl>
##  1 01M292 Developing        C                 C
       0.563
##  2 01M448 Developing        C                 B
       0.707
##  3 01M450 Well Developed    A                 B
       0.716
##  4 01M509 Proficient        C                 C
       0.564
##  5 01M539 Proficient        A                 A
       0.953
##  6 01M696 Well Developed    B                 C
       0.976
##  7 02M047 Proficient        C                 D
       0.696
##  8 02M288 Proficient        A                 B
       0.82
##  9 02M294 Well Developed    B                 B
       0.675
## 10 02M296 Proficient        A                 A
       0.793
## # ... with 412 more rows, and 1 more variable: college enroll 2010-1
1 <dbl>
```

```
graduationCollege <- school %>%
  select(`graduation 2010-11`, `college enroll 2010-11`)

plot(graduationCollege)

model <- lm(`college enroll 2010-11`~`graduation 2010-11`, data = gradu
ationCollege)
summary(model)
```

```
##
## Call:
## lm(formula = `college enroll 2010-11` ~ `graduation 2010-11`,
##     data = graduationCollege)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33512 -0.08226  0.00445  0.07970  0.28101
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -0.27965    0.03603  -7.761 1.46e-13 ***
## `graduation 2010-11`  1.09915    0.04798  22.910  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1153 on 289 degrees of freedom
##   (131 observations deleted due to missingness)
## Multiple R-squared:  0.6449, Adjusted R-squared:  0.6437
## F-statistic: 524.9 on 1 and 289 DF,  p-value: < 2.2e-16

plot(graduationCollege)
```




```
#abline(a = coef(model)[1], b = coef(model)[2], col = "red")

videoData = read_csv("ICE3_data.csv")

##
## -- Column specification --------------------------------------------
------------
## cols(
##   stid = col_double(),
##   year = col_double(),
##   video = col_character(),
##   participation = col_double(),
##   watch.time = col_double(),
##   confusion.points = col_double(),
##   key.points = col_double()
## )
```

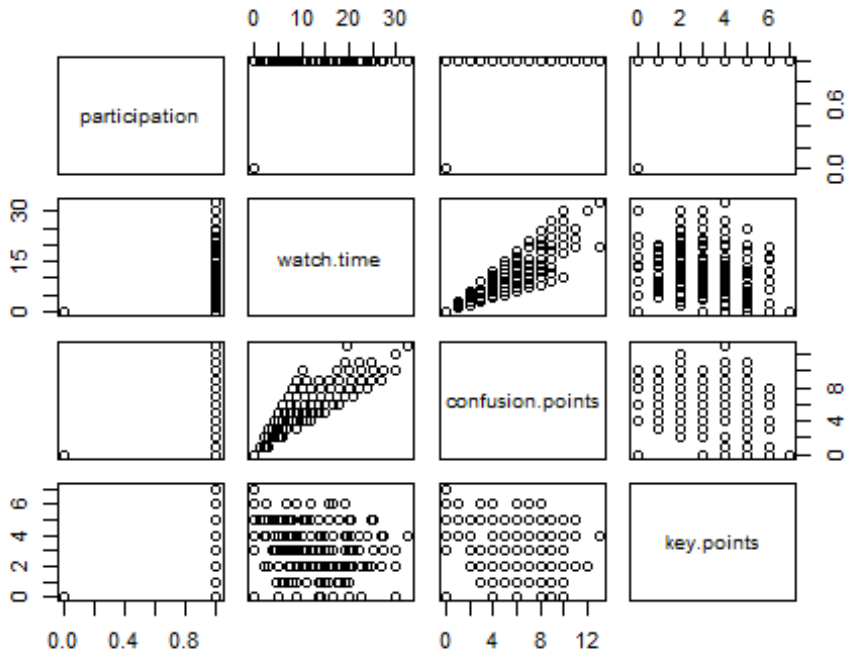
```
videoData
```

```
## # A tibble: 300 x 7
##     stid  year video participation watch.time confusion.points key.p
oints
##    <dbl> <dbl> <chr>         <dbl>      <dbl>             <dbl>
<dbl>
##  1     1  2018 A                 1       16.5                 6
    6
##  2     2  2018 A                 0        0                   0
    0
##  3     3  2018 A                 1        9                   4
    6
##  4     4  2018 A                 1       20                   8
    5
##  5     5  2018 A                 1       12                   8
    5
##  6     6  2018 A                 1       15                   5
    4
##  7     7  2018 A                 1       24.8                11
    5
##  8     8  2018 A                 1       12                   8
    6
##  9     9  2018 A                 1       15                   5
    2
## 10    10  2018 A                 1        0                   0
    5
## # ... with 290 more rows
```

```
summary(videoData)
```

```
##       stid            year         video           participation
##  Min.   : 1.00   Min.   :2018   Length:300         Min.   :0.0000
##  1st Qu.:15.75   1st Qu.:2018   Class :character   1st Qu.:0.0000
##  Median :30.50   Median :2018   Mode  :character   Median :1.0000
##  Mean   :30.50   Mean   :2018                      Mean   :0.7433
##  3rd Qu.:45.25   3rd Qu.:2019                      3rd Qu.:1.0000
##  Max.   :60.00   Max.   :2019                      Max.   :1.0000
##    watch.time     confusion.points   key.points
##  Min.   : 0.000   Min.   : 0.000   Min.   :0.000
##  1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.:0.000
##  Median : 8.375   Median : 5.000   Median :2.000
##  Mean   : 9.303   Mean   : 4.427   Mean   :2.327
##  3rd Qu.:15.750   3rd Qu.: 8.000   3rd Qu.:4.000
##  Max.   :32.500   Max.   :13.000   Max.   :7.000
```

```
videoDataRegression <- videoData %>% select(participation, watch.time,
confusion.points, key.points)
plot(videoDataRegression)
```



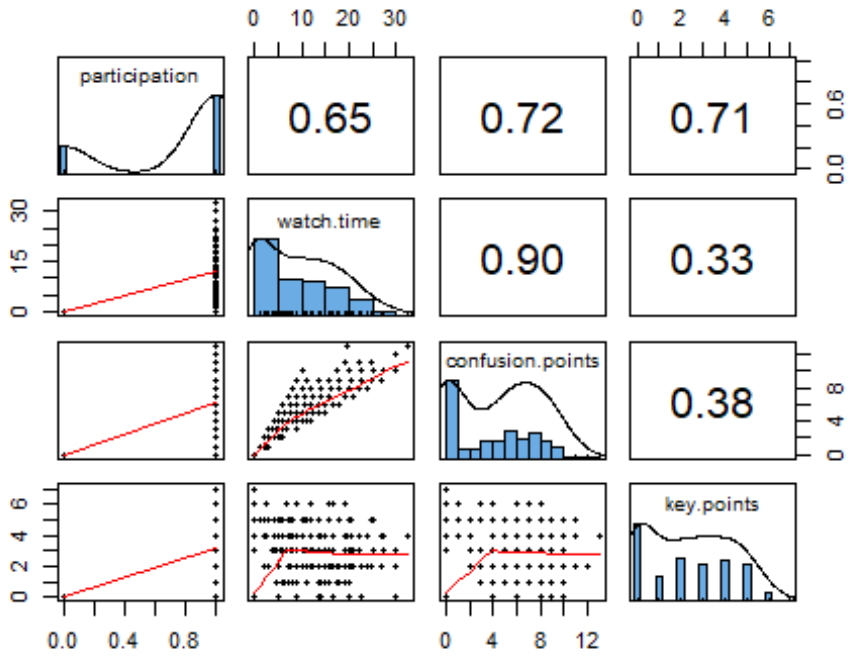

```
library(psych)
```

```
## Warning: package 'psych' was built under R version 4.0.5
```

```
##
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
pairs.panels(videoDataRegression,
             hist.col = "#6cace4", # You can use any color you like.
             ellipses= FALSE # FALSE turns off the correlation ellipses
, but feel free to set it as TRUE or just leave it as default, which is
 TRUE
             )
```




```
videoModel <- lm(watch.time ~ participation + confusion.points + key.po
ints, data = videoDataRegression)
summary(videoModel)

##
## Call:
## lm(formula = watch.time ~ participation + confusion.points +
##     key.points, data = videoDataRegression)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.675  -1.334   0.000   1.721   9.023
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.984e-14  4.100e-01   0.000    1.000
## participation     5.576e-01  9.398e-01   0.593    0.553
## confusion.points  2.087e+00  8.662e-02  24.097   <2e-16 ***
## key.points       -1.512e-01  1.603e-01  -0.943    0.346
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.597 on 296 degrees of freedom
## Multiple R-squared:  0.8183, Adjusted R-squared:  0.8164
## F-statistic: 444.3 on 3 and 296 DF,  p-value: < 2.2e-16

summary(model)
```

```
##
## Call:
## lm(formula = `college enroll 2010-11` ~ `graduation 2010-11`,
##     data = graduationCollege)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33512 -0.08226  0.00445  0.07970  0.28101
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -0.27965    0.03603  -7.761 1.46e-13 ***
## `graduation 2010-11`   1.09915    0.04798  22.910  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1153 on 289 degrees of freedom
##   (131 observations deleted due to missingness)
## Multiple R-squared:  0.6449, Adjusted R-squared:  0.6437
## F-statistic: 524.9 on 1 and 289 DF,  p-value: < 2.2e-16
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.