

Simple Linear Regression

Dr. Kiah Wah Ong

Using R and RStudio

- ▶ R is a statistical programming language that is open source and free.
- ▶ RStudio is a convenient interface for R, it is an integrated development environment for R.
- ▶ An analogy is that R is a car's engine while R Studio is like a car's dashboard.

Installation of R and RStudio

Go to

<https://www.rstudio.com/products/rstudio/download/>



PRODUCTS ▾ SOLUTIONS ▾ LEARN & SUPPORT ▾ EXPLORE MORE ▾ PRICING



DOWNLOAD

RStudio Desktop

Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python.

1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 208.07 MB | [SHA-256: 8DCB7103](#) | Version: 2023.03.1+446
| Released: 2023-05-12

RStudio Interface

The screenshot displays the RStudio interface with three main components highlighted by red boxes:

- Source Editor:** The top-left pane shows an R script titled "18-Generating data.R". The script contains the following code:

```
1 # Attempting to generate data for transformation
2 set.seed(26)
3 x=1:100
4 delta=rnorm(100,40,1)
5 y=2*(x*(1/2))^delta
6 plot(x,y)
7 model1=lm(y~x)
8 abline(model1)
9 summary(model1)
10 plot(model1)
11 lx=log(x)
12 ly=log(y)
13 plot(lx,ly)
14 model12=lm(ly~lx)
15 abline(model12)
16 plot(model12)
17 #summary(model12)
18 #summary(model13)
19 model13=lm(ly~x)
20 plot(x,ly)
21 abline(model13)
22 plot(model13)
```
- Environment/History:** The top-right pane shows the Global Environment with the following variables:

Variable	Value
n	32
p	3
sigma	1.0
SSR1	448.763636363636
SSR2	82.3
Sxx	82.3
Syy	54.9
- Console:** The bottom-left pane shows the output of the R script, including the residuals, coefficients, and summary statistics for the linear models.

```
lm(formula = y ~ x)

Residuals:
    Min       1Q   Median       3Q      Max
-150.142  -18.360    7.591   27.572   57.834

Coefficients:
(Intercept)  219.631    7.913   27.76  <2e-16 ***
x             6.252    0.136   45.95  <2e-16 ***

Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 39.27 on 98 degrees of freedom
Multiple R-squared:  0.9557, Adjusted R-squared:  0.9552
F-statistic: 2112 on 1 and 98 DF, p-value: < 2.2e-16
```

The bottom-right pane shows a scatter plot of y versus x, with a linear regression line. The plot is titled "Files/Plots/Packages/Help/Viewer".

R Output

