

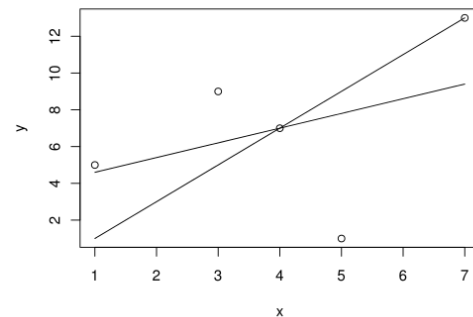
Regression Exercises with R

1. Use R to calculate the example from pages 132–133.

```
> x <- c(1, 3, 4, 5, 7)
> y <- c(5, 9, 7, 1, 13)
> source("http://www.adjoint-functors.net/SDline.R")
> SDline(x, y)
$meanX=4, $meanY=7, $slope=2, $correlation=0.4
> linearModel <- lm(y~x)
> summary(linearModel)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.800      4.733   0.803   0.481
x              0.800      1.058   0.756   0.505
Residual standard error: 4.733 on 3 degrees of freedom
Multiple R-squared: 0.16, Adjusted R-squared: -0.12
F-statistic: 0.5714 on 1 and 3 DF, p-value: 0.5046
```

2. Use R to plot the example from pages 132–133.

```
> x <- c(1, 3, 4, 5, 7)
> y <- c(5, 9, 7, 1, 13)
> plot(x, y)
> lines(x, 2*x - 1, type="l")
> lines(x, 0.8*x + 3.8, type="l")
```



3. Use the following to get the equation for the SD line of the `cars` data that is available in R.

```
> SDline(cars$speed, cars$dist)
> plot(cars$speed, cars$dist)
Add the SD line to your plot.
Use the following to get the regression line
> linearModel <- lm(cars$dist~cars$speed)
Add the regression line to your plot.
```

Regression Exercises

1. Consider the following data.

x: average ≈ 100 , SD ≈ 15
y: average ≈ 110 , SD ≈ 30 , $r \approx 0.8$

a) Predict the value of y if $x = 85$, $x = 130$ and $x = \text{unknown}$.

b) Predict the value of x if $y = 95$, $y = 140$ and $y = \text{unknown}$.

2. Consider the following data.

x: average ≈ 10 , SD ≈ 2
y: average ≈ 12 , SD ≈ 6 , $r \approx -0.2$

a) Predict the value of y if $x = 6$, $x = 16$ and $x = \text{unknown}$.

b) Predict the value of x if $y = 18$, $y = 6$ and $y = \text{unknown}$.

3. Consider the data from exercise 1.

a) Calculate the RMS error for regression.

b) Suppose $x = 85$. In what range does 68% of the y -data fall?

c) Suppose $y = 140$. In what range does 95% of the x -data fall?

4. Consider the data from exercise 2.

a) Calculate the RMS error for regression.

b) Suppose $x = 16$. In what range does 68% of the y -data fall?

c) Suppose $y = 18$. In what range does 95% of the x -data fall?