Solutions to Tutorial 4

1. An output of a simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, .., 10$$

is as follows

	Estimate	Std. Error	t value	P-value
Coefficients:				
(Intercept)	-0.07727	0.12005	-0.644	0.537814
x	0.97295	0.14345	6.783	0.000140

Residual standard error: 0.3761 on 8 degrees of freedom
Multiple R-squared: 0.8519, Adjusted R-squared: 0.8333
F-statistic: 46 on 1 and 8 DF, p-value: 0.0001403

(a) $b_0 = -0.07727, b_1 = 0.97295, s(b_0) = 0.12005, s(b_1) = 0.14345, t(b_0) = -0.644,$ $t(b_1) = 6.783$, $R^2 = 0.8519, \hat{\sigma} = 0.3761, r_{XY} = +\sqrt{R^2} = +\sqrt{0.8519}$, F-statistic=46

(b)

$$SSE = 0.3761^2 * 8 = 1.131610; \quad SST = SSE/(1 - R^2) = 7.640849;$$

$$SSR = SSR - SSE = 6.509239$$

Source	SS	DF	MS	F-vale
regression	6.509239	1	6.509239	46.01756
error	1.131610	8	0.1414512	
Total	7.640849	9		

(c)

$$t = \left|\frac{0.97295 - 1}{0.14345}\right| = 0.1885674 < t(1 - 0.05/2, 8) = 2.306$$

we accept H_0

2. **Sales Growth**

(R code)

i	1	2	3	4	5	6	7	8	9	10
X_i: Year	0	1	2	3	4	5	6	7	8	9
Y_i: Sales	98	135	162	178	221	232	283	300	374	395

(a) There is nonlinear pattern; see the first panel in Figure 1

(b)

$\hat{z}$	=	10.26	+	1.07 x
(SE)		(0.2129)		(0.0399)

(c) Does the regression line appear to be good to the transformed data?

Answer: YES (for the transformed data); See the second panel in Figure 1

(d) See the thrid and 4th panels in Figure 1 below.

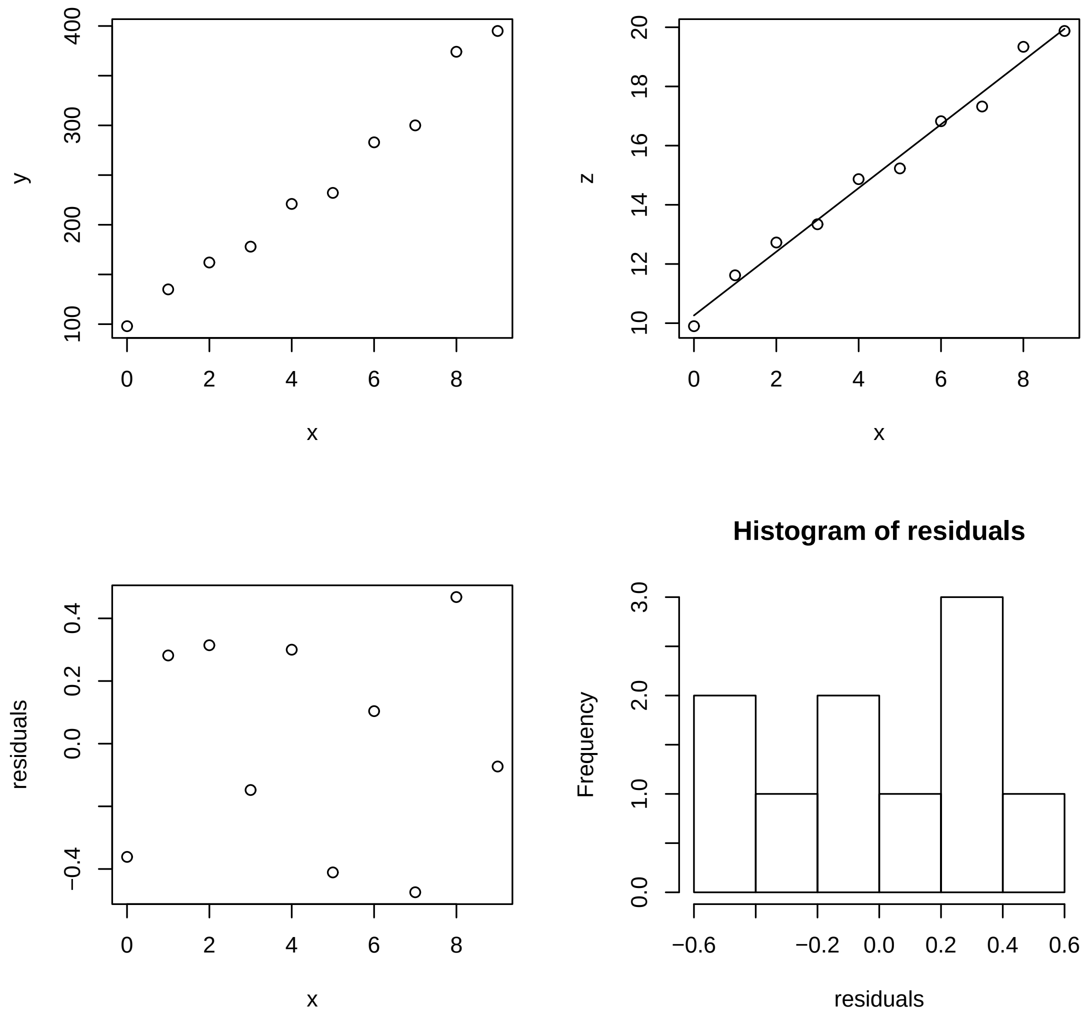

Figure 1:

3. **blood pressure (R code)**

i	1	2	3	4	5	6	7	8
X_i: age	5	8	11	7	13	12	12	6
Y_i: blood pressure	63	67	74	64	75	69	90	60

(a)

It shows that there is one outlier.

(b) comparing the two models, case 7 is very influential

(c) The interval is [60.31266, 84.73588], which does not cover case 7. This futher suggests that case 7 is an outlier.

4. write the details for

$$b = (X'X)^{-1}X'Y$$

Answer:

$$(X'X)^{-1} = \begin{bmatrix} \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} & \frac{-\bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ \frac{-\bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} & \frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{bmatrix}$$

and

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{bmatrix}$$

It follows

$$b = \begin{bmatrix} \bar{Y} - \bar{X}\frac{\sum_{i=1}^n (X_i - \bar{X})Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ \frac{\sum_{i=1}^n (X_i - \bar{X})Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{bmatrix}$$

5.

Since $r_{XY}^2 = r_{YX}^2$, both models have the same $R^2 = r_{YX}^2$.