1 Exercises
1.1 ISLR 2e (Gareth James, et al.), Chapter 3

Q.6)

Answer: The simple linear regression model uses the form of $y_i = ß_0 + ß_1 x_i$.

The error function is defined as sum of squares of each data row. The formula is given as

$$E(ß0, ß1) \quad = \sum_{i=1}^{n} (yi - \acute{y}i)^2 = \sum_{i=1}^{n} \big(yi - (ß0 + ß1\mathrm{x})\big)^2$$

Using error function

$$ß0 = \bar{y} - ß1\bar{x}$$

$$ß1 = \frac{\sum_{i=1}^{n}(xi - \bar{x})(yi - \bar{y})}{\sum_{i=1}^{n}(xi - \bar{x})^2}$$

To check whether $(\bar{x}, \bar{y})$ is used to satisfy the equation

$Y = ß_0 + ß_1 x_i$.

$Y = \bar{y} - ß1\bar{x} + ß1x$

When we plug x as $\bar{x}$ then

$Y = \bar{y} - ß1\bar{x} + ß1\bar{x}$

$Y = \bar{y}$

From the above result it is observed that least squares line always passes through point $(\bar{x}, \bar{y})$

1.2 ISLR 2e (Gareth James, et al.)

Q.1)

Answer: The null hypotheses obtained from the advertising budgets allocated to 'TV', 'radio' and 'newspaper' from Table has no impact on sales. Irrespective of the predicated values the coefficients are always zero. The pr values of TV and radio are positive and high and pr value of newspaper is negative. Which concludes that changing advertising budget for newspapers will affect sales negatively.

Q.4a)

Answer: Since x and y relationship is linear. As there is not much information on the data we can't say whether Training residual sum of squares of linear regression is less than or greater than training residual sum of squares of cubic regression.