Exercise 1
A new profit-sharing plan was introduced at an automobile parts manufacturing plant last year. Both management and union representatives were interested in determining how a worker’s years of experience influence his or her productivity gains. After the plan had been in effect for a while, the data shown below were collected:

<table>
<thead>
<tr>
<th>Years of experience (x)</th>
<th>Number of units daily (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.1</td>
<td>110</td>
</tr>
<tr>
<td>7.0</td>
<td>105</td>
</tr>
<tr>
<td>18.6</td>
<td>115</td>
</tr>
<tr>
<td>23.7</td>
<td>127</td>
</tr>
<tr>
<td>11.5</td>
<td>98</td>
</tr>
<tr>
<td>16.4</td>
<td>103</td>
</tr>
<tr>
<td>6.3</td>
<td>87</td>
</tr>
<tr>
<td>15.4</td>
<td>108</td>
</tr>
<tr>
<td>19.9</td>
<td>112</td>
</tr>
</tbody>
</table>

For your convenience:
\[ \sum_{i=1}^{9} y_i = 965, \sum_{i=1}^{9} x_i = 133.9, \sum_{i=1}^{9} y_i^2 = 104469, \sum_{i=1}^{9} x_i^2 = 2258.73, \sum_{i=1}^{9} x_i y_i = 14801.2. \]

a. Find the least-squares regression line (perform all calculations by hand).

b. Use R: Verify your answer to part (b). Submit the R results.

c. Use R: Construct the scatterplot of the number of units manufactured daily against the number of years of experience and add the fitted line.

d. Predict the number of units manufactured daily by an employee who has 10 years of experience on the assembly line.

Exercise 2
For the regression model \( y_i = \beta_0 + \beta_1 x_i + \epsilon_i \) show that

a. \( \sum_{i=1}^{n} e_i = 0. \)

b. \( \sum_{i=1}^{n} e_i x_i = 0. \)

c. \( \sum_{i=1}^{n} e_i \hat{y}_i = 0. \)

Exercise 3
Suppose in the model \( y_i = \beta_0 + \beta_1 x_i + \epsilon_i \), where \( i = 1, \ldots, n \), \( E(\epsilon_i) = 0 \), \( \text{var}(\epsilon_i) = \sigma^2 \) the measurements \( x_i \) were in inches and we would like to write the model in centimeters, say, \( z_i \). If one inch is equal to \( c \) centimeters (\( c \) is known), we can write the above model as follows \( y_i = \beta_0^* + \beta_1^* z_i + \epsilon_i \).

a. Suppose \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) are the least squares estimates of \( \beta_0 \) and \( \beta_1 \) of the first model. Find the estimates of \( \beta_0^* \) and \( \beta_1^* \) in terms of \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \).

b. Show that the value of \( R^2 \) remains the same for both models.

c. Find the variance of \( \hat{\beta}_1^* \).

Exercise 4
Consider the regression model
\[ y_i = (\beta_0 + \beta_1 \bar{x}) + \beta_1 (x_i - \bar{x}) + \epsilon_i \]

This model is called the centered version of the regression model \( y_i = \beta_0 + \beta_1 x_i + \epsilon_i \) that was discussed in class. If we let \( \gamma_0 = \beta_0 + \beta_1 \bar{x} \) we can rewrite the centered version as \( y_i = \gamma_0 + \beta_1 (x_i - \bar{x}) + \epsilon_i \). Find the least squares estimates of \( \gamma_0 \) and \( \beta_1 \).
Exercise 5
Consider the regression model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$. Show that $\text{Cov}(\bar{Y}, \hat{\beta}_1) = 0$ where $\bar{Y}$ is the sample mean of the $y$ values, and $\hat{\beta}_1$ is the estimate of $\beta_1$.

Exercise 6
Consider the regression model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$. Find $\text{cov}(e_i, e_j)$.

Exercise 7
Suppose $Y_i = \beta x_i + \epsilon_i$. In this equation $X$ is non-random, $\beta$ is a parameter (unknown), and $\epsilon \sim N(0, \sigma)$.
   
   a. Find the mean of $Y$.
   b. Find the variance of $Y$.
   c. What distribution does $Y$ follow?
   d. Write down the likelihood function based on $n$ observations of $Y$ and $x$.
   e. Find the maximum likelihood estimate of $\beta$. Denote it with $\hat{\beta}$.
   f. Show that the estimate of part (e) is unbiased estimator of $\beta$.
   g. Find the variance of this estimate.

Exercise 8
Consider the model of exercise 7. Find the covariance between $e_i$ and $e_j$. Also find the variance of $e_i$. 