We describe below a statistical linear regression problem. We include some applicable formulas after the statement of the problem.

Statement of the problem
Let $\beta_0$, $\beta_1$, and $\sigma > 0$ be unknown real numbers.

Assume the following:

a) Let $x_1 = 16$, $x_2 = 28$, $x_3 = 40$, $x_4 = 52$, and $x_5 = 64$.

b) Let $e_1, e_2, e_3, e_4$, and $e_5$ be independent normal random variables with mean 0, and variance $\sigma^2$.

c) Let $y_i = \beta_0 + x_i \beta_1 + e_i$ be random variables for $i \in \{1, 2, 3, 4, 5\}$.

d) A set of sample outcomes for the random variables in part c) are the following:

\[
y_1 = 130.6, \quad y_2 = 225.6, \quad y_3 = 322.1, \quad y_4 = 418.6, \quad \text{and} \quad y_5 = 514.6
\]

Then estimate $\beta_0$, $\beta_1$, and $\sigma > 0$ to the nearest hundredth.

Applicable formulas

1) Estimates for $\beta_0$ and $\beta_1$ come from

\[
\begin{pmatrix}
\hat{\beta}_0 \\
\hat{\beta}_1
\end{pmatrix} = (X^T X)^{-1} X^T Y.
\]

2) An estimate to $\sigma^2$ is the sample variance

\[
s^2 = \frac{1}{n - k - 1} \sum_{i=1}^n (y_i - \hat{y}_i)^2.
\]

In this case, $n = 5$ and $k = 1$.

3) The mean of the observed values $y_i$ is

\[
\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.
\]

The so-called coefficient of determination is given by

\[
r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}.
\]

If the value of $r^2$ is approximately 1, the approximations to $\beta_0$ and $\beta_1$ are good.

In such a case, the linear regression $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ provides a good approximation to the data set $(x_i, y_i)$, $i \in \{1, \ldots, n\}$. 
