1- Given a simple linear regression model with slope β_1 and intercept β_0, show that $b_0 \sim N\left(\beta_0, \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right]\right)$?

2- Following on Q1, what is the distribution of b_0 when σ^2 is estimated?

3- Following on Q1 and Q2, construct a 99% CI of b_0

4. The table below lists the USA social security costs for 7 specific years between 1965 and 1992.

Year	1965	1970	1975	1980	1985	1990	1992
x = Year − 1960	5	10	15	20	25	30	32
y = social security cost ($ billion)	17.1	29.6	63.6	117.1	186.4	246.5	285.1

(a) Plot the data using y against x (a hand-drawn graph is acceptable).

(b) Compute $\sum_i x_i$, $\sum_i y_i$, $\sum_i x_i^2$, $\sum_i y_i^2$ and $\sum_i x_i y_i$. Use these figures to fit the data with the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$.

(c) Test the hypothesis $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 > 0$ at the 5% significance level. What can be concluded about social security costs from this test?

(d) Plot the residuals against x (a hand-drawn graph is acceptable). Are you happy with the fitted model? If not, discuss what you might try to achieve a better fit.

5. The following is a portion of Minitab output for a linear regression analysis relating maintenance expenses (y, in pounds per month) to usage (x, in hours per week) of a particular brand of computer terminal.

```
The regression equation is
y = 4.6849 + 0.9475x

Predictor      Coef       SE Coef
Constant       4.6849     0.8470
X              0.9475     0.1025

Analysis of Variance

SOURCE           DF    SS
Regression        1    1683
Residual Error   23     453
Total            24    2136
```

(a) Test at the 5% significance level if maintenance expenses are related to usage, and also test if the intercept is equal to 5, again at the 5% significance level.

(b) Find an estimate of the variance σ^2.

(c) Predict the monthly maintenance expense when the usage is 7 hours.

(d) Calculate R^2 (the coefficient of determination) and interpret the value.

Q6 The Tri-City Office Equipment Corporation sells an imported copier on a franchise basis and performs preventive maintenance and repair service on this copier. The data below have been collected from 45 recent calls on users to perform routine preventive maintenance service; for each call, X is the number of copiers serviced and Y is the total number of minutes spent by the service person. Assume that first-order regression model is appropriate. Use the data in the Copier maintenance.xlsx file

a. Write the regression model, the regression function and its fitted version.

b. Plot the estimated regression function and the data. How well does the estimated regression function fit the data?

c. Interpret $\boldsymbol{b0}$ in your estimated regression function. Does $\boldsymbol{b0}$ provide any relevant information here? Explain.

d. Obtain a point estimate of the mean service time when X = 5 copiers are serviced.

e. Obtain point estimates of σ^2 and $\boldsymbol{\sigma}$. In what units is $\boldsymbol{\sigma}$ expressed?

f. Given the following

$$n = 45, \sum_{i=1}^{n=45} X_i = 230, \sum_{i=1}^{45} Y_i = 3432, \sum_{i=1}^{45} X_i^2 = 1516, \quad \sum_{i=1}^{45} X_i Y_i = 22660$$

$$SSE = 3416.377$$

Estimate the change in the mean service time when the number of copiers serviced increases by one. First use a 90% CI then a 99%CI. Interpret both intervals and which one provides more certainty about the parameter value.

g. Determine whether there is a linear association between X and Y at 0.01 level of significance.

h. Are your results in parts (f) and (g) consistent? Explain.

i. The manufacturer has suggested that the mean required time should not increase by more than 14 minutes for each additional copier that is serviced on a service call. Using p-value, conduct a test to decide whether this standard is being satisfied by Tri-City.