3. Write R script to compute the regression equation of y on x from the following data. Predict the value of y when x=7

X	2	4	5	6	8	11
Y	18	12	10	8	7	5

```
 x <- c(2, 4, 5, 6, 8, 11)
y <- c(18, 12, 10, 8, 7, 5)

model <- lm(y ~ x)

summary(model)

new_data <- data.frame(x = 7)
predicted_y <- predict(model, newdata = new_data)

cat("Regression Equation: y =", round(coef(model)[1], 2), "+", round(coef(model)[2], 2),"x\n")

cat("Predicted y when x = 7:", round(predicted_y, 2), "\n")
```

Output:
Regression Equation: y = 18.04 + -1.34 x
Predicted y when x = 7: 8.66

Manual:

x	x2	y	y2	xy
2	4	18	324	36
4	16	12	144	48
5	25	10	100	50
6	36	8	64	48
8	64	7	49	56
11	121	5	25	55
$\Sigma x = 36$	$\Sigma x^2 = 266$	$\Sigma y = 60$	$\Sigma y^2 = 706$	$\Sigma xy = 293$

N = 6

Mean of x, $\bar{x} = \frac{36}{6} = 6$

Mean of y, $\bar{y} = \frac{60}{6} = 10$

Regression coefficient of y on x

$$b_{yx} = \frac{N \,.\, \Sigma xy - (\Sigma x \,.\, \Sigma y)}{N \,.\, \Sigma x^2 - (\Sigma x)^2}$$

$$b_{yx} = \frac{6 \,.\, 293 - (36 \,.\, 60)}{6 \,.\, 266 - 36^2}$$

$$b_{yx} = \frac{1758 - 2160}{1596 - 1296}$$

$$b_{yx} = \frac{-402}{300} = 1.34$$

Regression equation y on x

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$y - 10 = -1.34\,(x - 6)$$
$$y - 10 = -1.34x + 8.04$$
$$y = -1.34x + 8.04 + 10$$
$$y = -1.34x + 18.04$$

The value of y when x = 7

$$y = -1.34x + 18.04$$
$$y = (-1.34 \,.\, 7) + 18.04$$
$$y = -9.38 + 18.04$$
$$y = 8.66$$

4. The times taken by a large group of students to complete a piece of homework, T minutes, are Normally distributed with a mean of 57 minutes and standard deviation of 6.5. Find the probability that the time taken by a random student from the group to complete this homework will be less than 60 minutes.
Write an R script to Find the probability of the time taken by a random student from the group to complete this homework.

A. Will be less than 60 minutes
B. Between 50 and 80 minutes

```
mean_time <- 57
std_dev_time <- 6.5

prob_less_than_60 <- pnorm(60, mean_time, std_dev_time)
cat("Probability that time taken will be less than 60 minutes:", prob_less_than_60, "\n")

prob_between_50_80 <- pnorm(80, mean_time, std_dev_time) - pnorm(50, mean_time, std_dev_time)
cat("Probability that time taken will be between 50 and 80 minutes:", prob_between_50_80)
```

Output:
Probability that time taken will be less than 60 minutes: 0.6777938
Probability that time taken will be between 50 and 80 minutes: 0.8590415

Manual:
μ = 57
σ = 6.5

The formula for calculating the z - score is

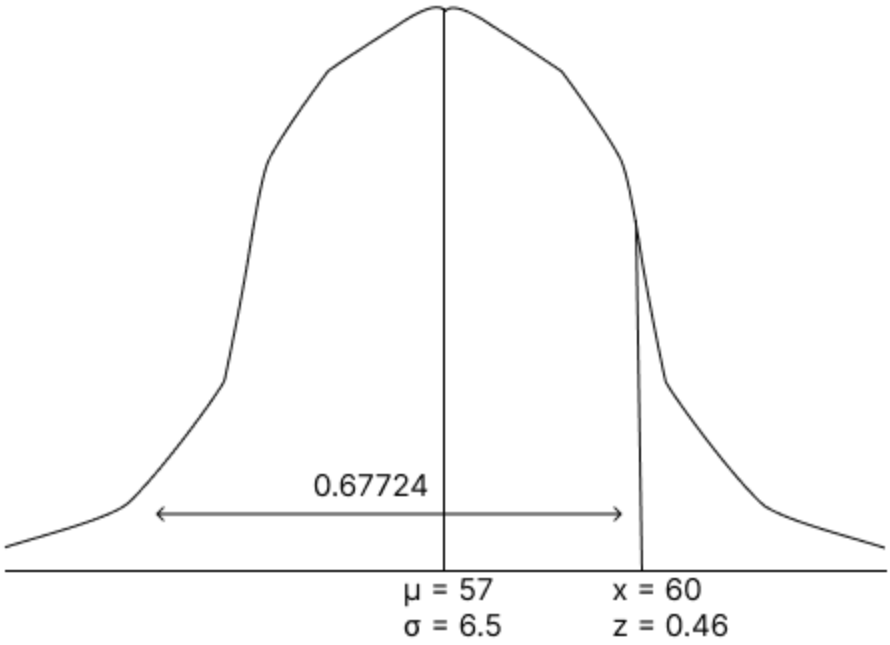
$$z = \frac{x - \mu}{\sigma}$$

a) less than 60 minutes
P(x < 60)

$$z = \frac{60 - 57}{6.5}$$

$$z = 0.4615$$

We get probability associated with this z-score using the z-table
Normal distribution table value of z = 0.4615 is 0.67724



The probability that a randomly selected student from the group will less than 60 minutes to complete the homework is 0.6772 or 67.72%

b) Between 50 and 80 minutes
P(50 < x < 80)

P(50 < x)

$$z = \frac{50 - 57}{6.5}$$

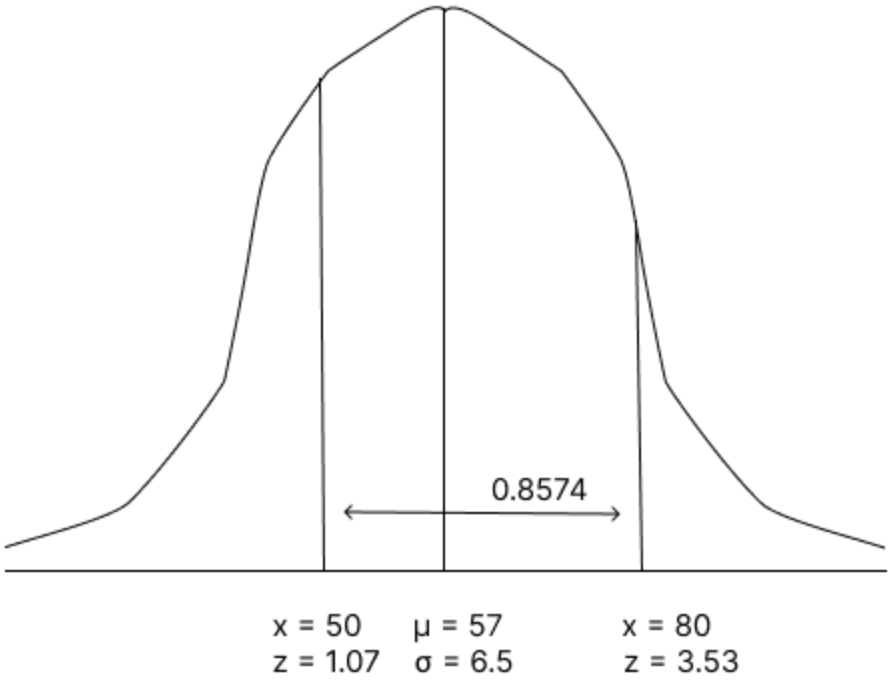
$$z = -1.076$$

P(80 < x)

$$z = \frac{80 - 57}{6.5}$$

$$z = 3.538$$

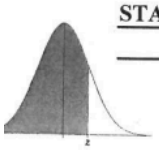
Normal distribution table value of z = -1.076 is 0.14321
Normal distribution table value of z = 3.358 is 0.99979



The probability that a randomly selected student will take between 50 and 80 minutes is
$P(50 < x < 80)$
$P(-1.076 < z < 3.358)$
$0.14321 < z < 0.99979$
$z = 0.99979 - 0.14321 = 0.8574$

The probability is 0.8574 or 85.74%
(note: b's answer is not matching)

Note: How to find Area under the curve using z-table
Identify 0.1th and 0.01th decimal from z - score (0.4615)
0.1th value 0.4 and 0.01th value 0.06
Find item with row 0.4 and column 0.06 in z-table



STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673

Using this table we get 0.67724 as probability of z - score 0.4615