PART 3: QUESTION 4

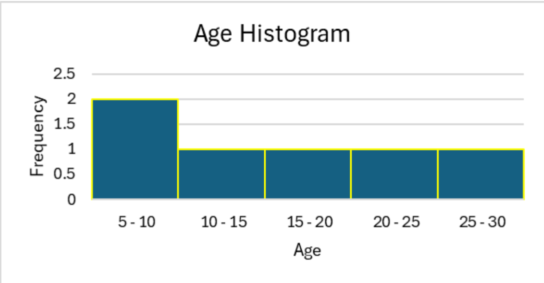
(a-1) Draw histograms (by hand) for both Age and Strength

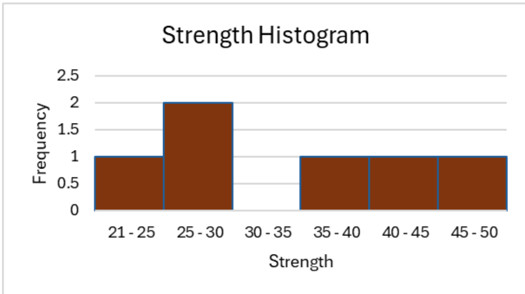
Computation Tables (Grouping)

Age groups	frequency
5 - 10	2
10 - 15	1
15 - 20	1
20 - 25	1
25 - 30	1

Strength groups	frequency
21 - 25	1
25 - 30	2
30 - 35	0
35 - 40	1
40 - 45	1
45 - 50	1

Histograms





(a-2) Comment on the shape of each distribution.

Both distributions can be described as right skewed (with more data to the left)

(b-1) Apply a natural log transformation

Computation Tables (Applying Natural Logarithms, ln)

Original Alien Beam Data		**Log-transformed Data**	
Age, x (years)	**Strength, y (MPa)**	ln(Age), x'	ln(Strength), y'
5	48	1.6094	3.8712
10	42	2.3026	3.7377
15	37	2.7081	3.6109
20	30	2.9957	3.4012
25	27	3.2189	3.2958
30	21	3.4012	3.0445

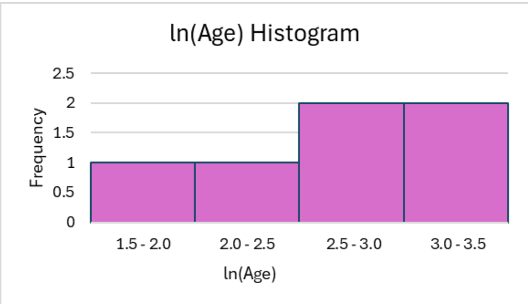
(b-1) Re-draw histograms for the transformed variables

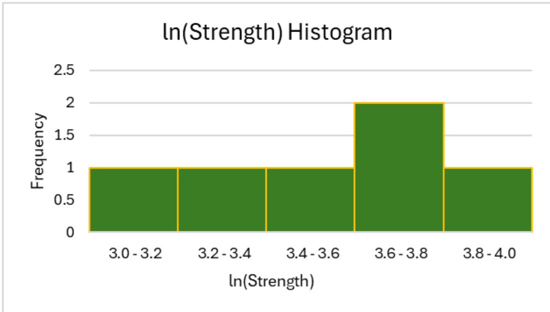
Computation Tables (Grouping)

ln(Age) groups	frequency
1.5 - 2.0	1
2.0 - 2.5	1
2.5 - 3.0	2
3.0 - 3.5	2

ln(Strength) groups	frequency
3.0 - 3.2	1
3.2 - 3.4	1
3.4 - 3.6	1
3.6 - 3.8	2
3.8 - 4.0	1

Histograms





(c) Compute the regression coefficients β0 and β1 for the log-log model using the least squares method

According to the least squares paradigm, the best fitting regression line is obtained with optimal coefficients, i.e.:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad \text{and} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

For β_1:

	Computations on Log-transformed Data					
	ln(Age), x'	**ln(Strength), y'**	**x'-x***	**(x'-x*)^2**	**y'-y***	**(x'-x*) . (y'-y*)**
	1.6094	3.8712	-1.0965	1.20240	0.3776	-0.41410
	2.3026	3.7377	-0.4034	0.16273	0.2441	-0.09847
	2.7081	3.6109	0.0021	0.00000	0.1174	0.00024
	2.9957	3.4012	0.2898	0.08396	-0.0924	-0.02676
	3.2189	3.2958	0.5129	0.26306	-0.1977	-0.10141
	3.4012	3.0445	0.6952	0.48333	-0.4490	-0.31218
Summation	**16.23588**	**20.96135**		2.19548		-0.95268
Mean (x* & y*)	x* = 16.23588/6	y* = 20.96135/6				
Mean (x* & y*)	**2.7060**	**3.4936**		**β1**	-0.43393	

For β_0:

From $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

β0	4.66776

(d) Determine the R^2 value and interpret its meaning

The formula below is used:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \in [0,1]$$

	Computing Predicted values [ln(Strength), y^]			
ln(Strength), y^ is computed using the regression coefficients, β0 and β1				
ln(Strength), y'	**ln(Strength), y^**	**Residuals (y' - y^)**	**RSS =(y' - y^)^2**	**(y'-y*)^2**
3.8712	3.96938	-0.09818	0.00964	0.14261
3.7377	3.66860	0.06907	0.00477	0.05959
3.6109	3.49266	0.11826	0.01399	0.01377
3.4012	3.36783	0.03337	0.00111	0.00853
3.2958	3.27100	0.02484	0.00062	0.03909
3.0445	3.19188	-0.14736	0.02172	0.20163
Summation			0.05184	0.46524

R^2	1 - [sum((y' - y^)^2)/sum((y'-y*)^2)]
R^2	0.88857

(e) Compute the p-value for β_1 and explain whether Age significantly affects Strength.

Testing the Null Hypothesis

$$T_{H_0:\beta_1=0} = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$$

Computing the P-value for β1		
n (data points)		6
df (degrees of freedom)	= (n-2)	4
Residual Sum of Squares (RSS) (errors^2) = sum(y' - y^)^2	= Sum(y' - y^)^2	0.05184
Residual Standard Error (RSE)	= RSS/df	**0.11384**
Computing the t-statistic (using the β1 = 0 null hypothesis)		
β1 (Slope Coefficient)		-0.43393
[(x'-x*)^2]^0.5		1.48172
SE(β1) : Standard error		0.07683
t-statistic for β1		-5.64781
Two-tailed p-value; p = 2*P(T>|t|)	(From Excel)	0.004841
Two-tailed p-value; p = 2*P(T>|t|)	(From Tables) = 1-0.995	0.005

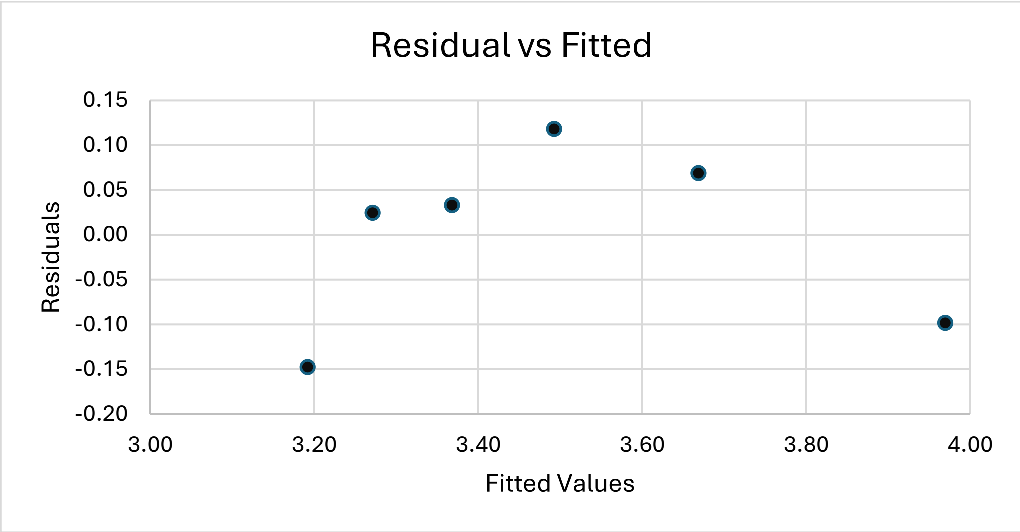
The p-value for β_1 = **0.005**. We reject the null hypothesis at the 5% significance level. This means that age has a statistically significant effect on the bending strength of alien beams, and we can be fairly confident (with 95% confidence) that the relationship isn't due to chance.

(f-1) Compute the residuals for the transformed model. Manually plot the residuals against the fitted values.

Computation of the residuals

Residuals vs Fitted	
Residuals (y' - y^)	**Fitted [ln(Strength), y^]**
-0.09818	3.96938
0.06907	3.66860
0.11826	3.49266
0.03337	3.36783
0.02484	3.27100
-0.14736	3.19188

(f-2) Manually plot the residuals against the fitted values.



VALIDATE YOUR HAND CALCULATIONS IN R.

------ (Code and Output shown in the following section) ----