

Question 1

Tania took a placement test. Her Z-score was 3.0, and her mark was 82%. The standard deviation of the group was 11.3.

a) What was the Arithmetic mean of this distribution?

$$Z = \frac{x_i - \bar{x}}{s}$$

$$\frac{3}{1} = \frac{82 - \bar{x}}{11.3} \rightarrow \bar{x} = 48.17\%$$

$$3(11.3) = 82 - \bar{x} \rightarrow \bar{x} = 82 - 33.9$$

b) Complete the following table:

Distribution	x_i	\bar{x}	s	z
A	68	76	14	A
B	1232	1228	B	0.4
C	C	21.2	5.1	-0.61

Show all your work clearly.

A $z = \frac{x_i - \bar{x}}{s} = \frac{68 - 76}{14} = \frac{-8}{14} = -0.57$

B $0.4 = \frac{1232 - 1228}{s}$

$\frac{0.4s}{0.4} = \frac{4}{0.4}$

$s = 10$

C $-0.61 = \frac{x_i - 21.2}{5.1}$

$(-0.61)(5.1) = x_i - 21.2$
 $-3.111 + 21.2 = x_i = 18.09$

Question 2

Marilyn and Carol both took the same exam in two different groups. Their Z-scores were calculated based on their respective groups. Marilyn's Z-score was higher than Carol's, even though they obtained the same mark.

- a) If the group averages were the same, what can we say about the standard deviations of the two groups? Explain.

$$\begin{array}{r} \text{Marilyn} \\ \hline x_i = x_i \\ \bar{x} = \bar{x} \\ Z > Z \end{array}$$

$$Z = \frac{x_i - \bar{x}}{s}$$

$$s_{\text{Marilyn}} < s_{\text{Carol}}$$

- b) If the standard deviations were the same, which group had the lower average? Explain.

$$\begin{array}{r} \text{Marilyn} \quad \text{Carol} \\ \hline x_i = x_i \\ s = s \\ Z > Z \end{array}$$

$$\bar{x}_{\text{Marilyn}} < \bar{x}_{\text{Carol}}$$

Question 3

Before ordering light-bulbs for street lamps, the municipal engineer decides to test two brands, A and B. He took a sample of 12 light bulbs and tested them for their durability in hours. Here are the results of his tests:

Brand A	380	380	390	390	390	390	400	400	410	420	440	450
Brand B	420	430	430	430	435	450	450	450	460	460	460	460

- a) Calculate the mean and standard deviations for both brands of light bulbs. Show all your work.

$$\bar{X}_A = 403.3 \quad \bar{X}_B = 444.6$$
$$S_A = 22.7 \quad S_B = 14.7$$

- b) Which brand should the engineer buy and why?

Brand B since they last longer

Question 4

The members of a selection committee analyzed the files of four candidates for an accounting position. The four candidates graduated from the same CEGEP but were in four different groups.

In the following table, we can see the marks for each candidate as well as their group average and standard deviation.

Candidate	Mark	Group Average	Standard Deviation
Francis	90	78	8
Lawrence	86	76	6
Gabrielle	78	62	9
Vincent	76	60	7

If the committee is basing their decisions on the candidates' Z-scores, in which order should they interview them? Show all your work.

$$Z_F = \frac{x_i - \bar{x}}{s} = \frac{90 - 78}{8} = \frac{12}{8} = 1.5$$

$$Z_L = \frac{86 - 76}{6} = \frac{10}{6} = 1.\bar{6}$$

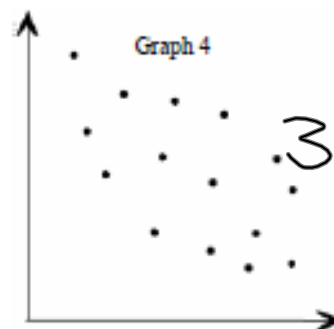
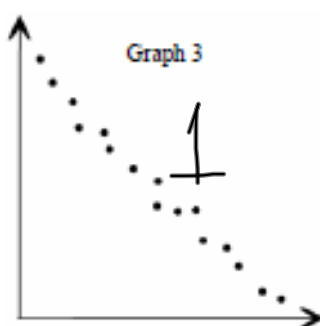
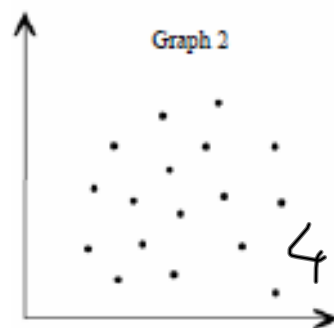
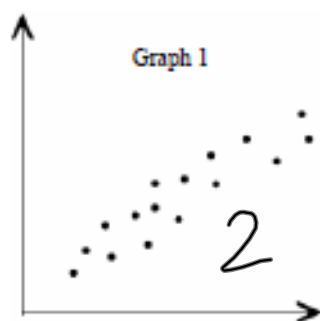
$$Z_G = \frac{78 - 62}{9} = 1.\bar{7}$$

$$Z_V = \frac{76 - 60}{7} = 2.3$$

V - G - L - F

Question 5

Each of the following scatter plots represents the correlation between the age of an adult and a variable.



What is the correct order of linear correlation, from strongest to weakest?

- a) Graph 2, Graph 4, Graph 1, Graph 3
- b) Graph 4, Graph 2, Graph 3, Graph 1
- c) Graph 3, Graph 1, Graph 4, Graph 2
- d) Graph 1, Graph 3, Graph 4, Graph 2

Question 6

Associate the given correlation coefficients with their corresponding description.

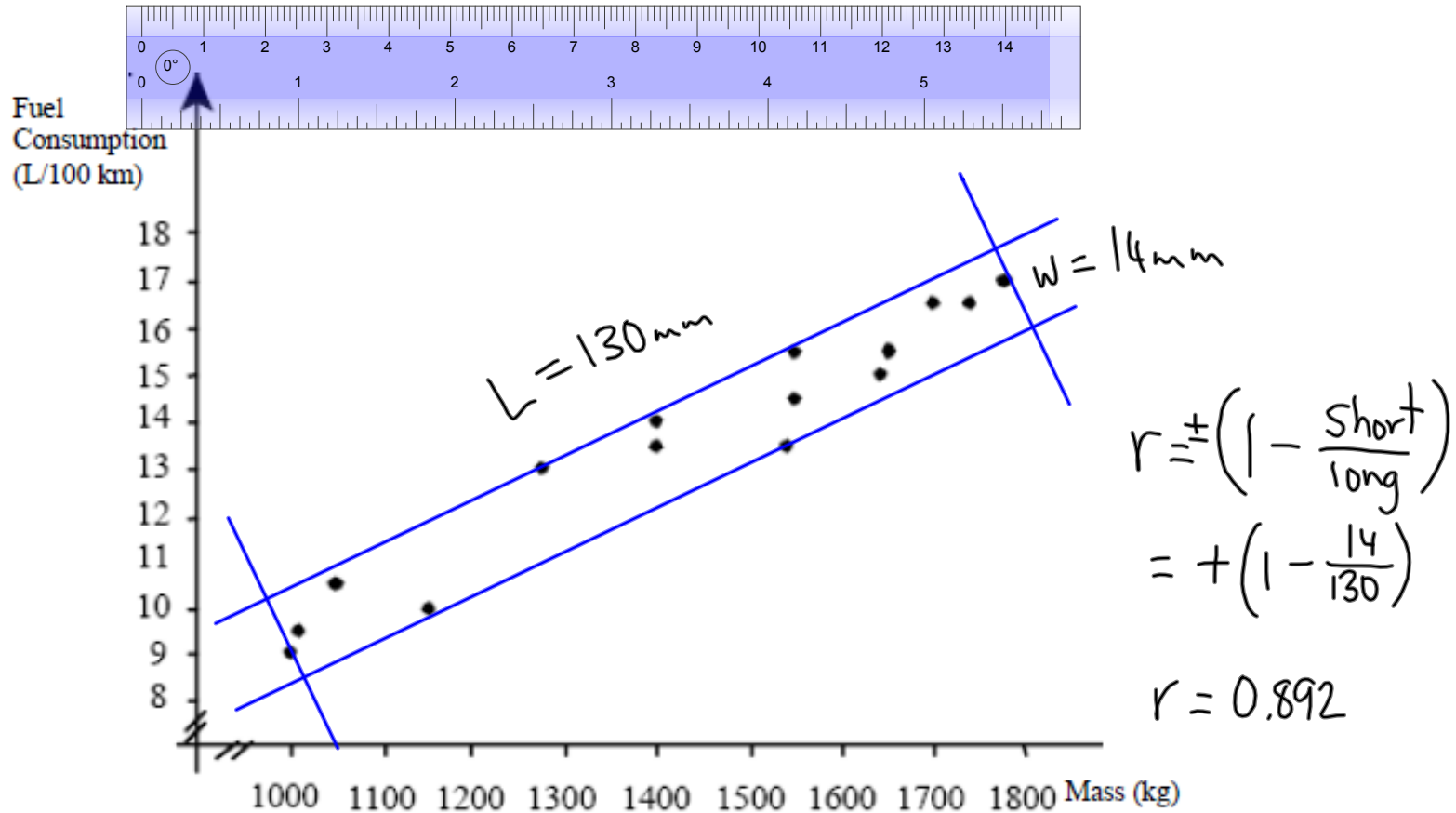
1. $r = -0.68$ 2. $r = 0.45$ 3. $r = 0$ 4. $r = 0.68$

- A. There is a weak correlation between the number of hours spent watching television and a person's age. We can see that there is a small tendency for older people to watch more television. $r = 0.45$
- B. There is no correlation between the amount of snow that falls in winter and the amount of rain that falls in summer. $r = 0$
- C. There is a moderate correlation between the ~~the~~ number of hours spent studying and the overall average on a student's report card. This correlation indicates that the student's average increases with more time spent studying.

$$r = 0.68$$

Question 7

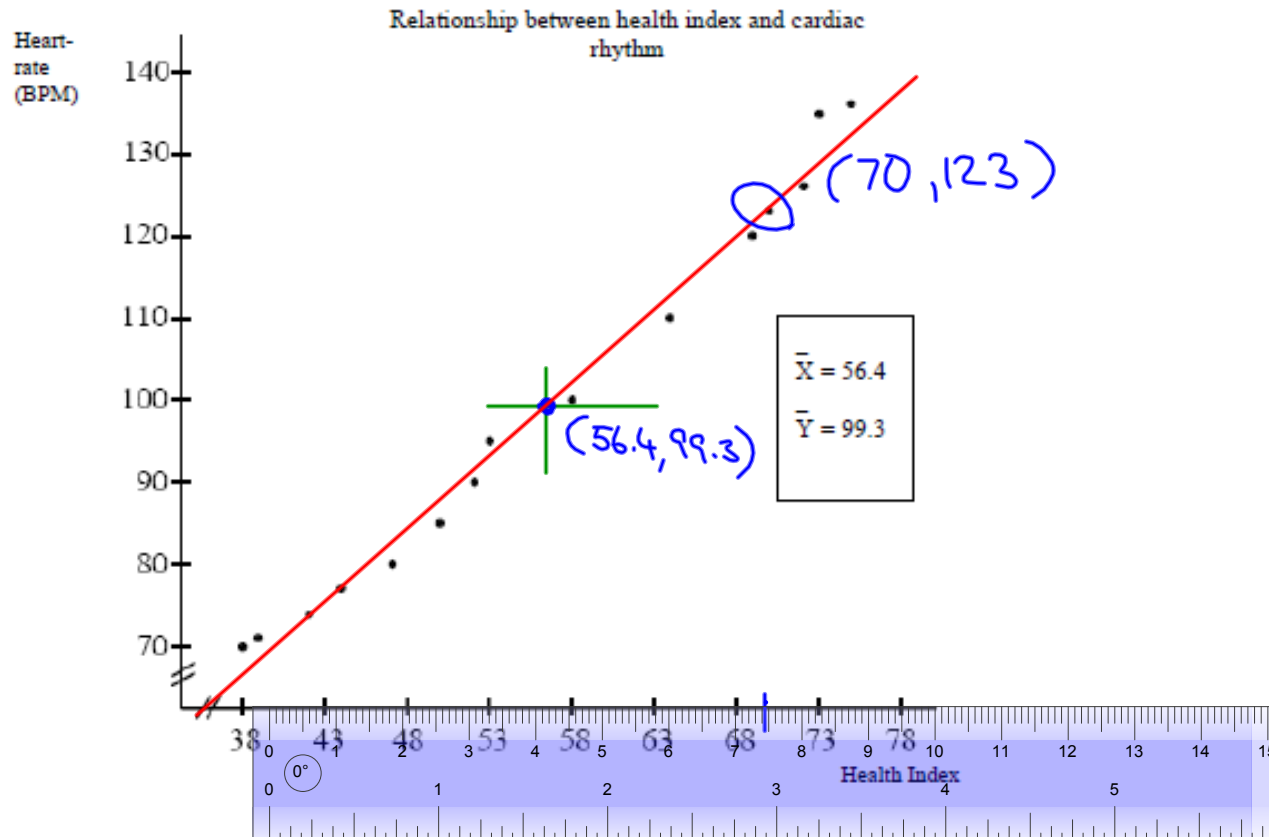
The following scatter plot illustrates the correlation between the weight of a car in kilograms and its consumption of fuel in litres per 100 km.



Using the graphical method, estimate the coefficient of linear correlation. Show all your work.

Question 8

A group of specialists analyzed the relationship between the health of an individual and their cardiac rhythm during moderately intensive physical exercise. The following scatter plot represents their findings:



mean - scatter plot

- a) Using the ~~median~~ method, estimate the equation of the regression line.

$$P_1(56.4, 99.3)$$

$$P_2(70, 123)$$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{123 - 99.3}{70 - 56.4} = 1.7426$$

$$\begin{aligned} y &= mx + b \\ y &= 1.7426x + b \\ 123 &= 1.7426(70) + b \\ b &= 123 - 1.7426(70) \\ &= 1.018 \end{aligned}$$

$$y = 1.7426x + 1.018$$

- b) According to the equation obtained in (a), what would be the health index of a person with a heart rate of 75?

$$y = 75$$

$$75 = 1.7426x + 1.018$$


$$\frac{75 - 1.018}{1.7426} = \frac{\cancel{1.7426}x}{\cancel{1.7426}}$$

$$x = 42.45$$

Question 9

Ronald is researching the quantity of fat, cholesterol and sodium contained in 14 different brands of chocolate-chip cookies.

His research is represented in the table below.



Brand	Fat (g)	Sodium (mg)	Cholesterol (mg)
A	15	16	8
B	21	14	9
C	12	18	6
D	18	10	6
E	36	6	12
F	39	2	12
G	9	12	6
H	24	10	10
I	33	2	9
J	30	8	8
K	6	16	7
L	42	2	11
M	3	22	5
N	27	4	9

$$r = -0.753$$

Based on this research, how can we describe the correlation between sodium and cholesterol content in these 14 brands of cookies? Explain your answer.

- A) Moderate and Positive B) Moderate and Negative
 C) Strong and Positive D) Strong and Negative

Explanation:

$$r = -0.753$$

Question 10

Researchers wanted to determine if there was a correlation between the Intelligence Quotient (IQ) and mathematics and music marks of students in a particular school.

These are the findings for a group of 13 students.

IQ	99	123	123	104	127	108	122	106	104	123	100	121	118
Math mark	39	45	39	50	65	34	61	42	50	75	40	64	45
Music mark	52	48	35	55	68	48	61	38	55	68	60	75	65

Here are the results of the statistical analysis of this data:

Correlation between IQ and Math:

$r = 0.589$ Weak
 $y_1 = 0.4826 x_1 + 89.6$

Correlation between IQ and Music:

$r = 0.2939$ None
 $y_2 = 0.25 x_2 + 99.4$

- a) Can we conclude that the marks obtained in mathematics and in music offer an appropriate evaluation of the intelligence quotient of an individual? Explain.

No since correlations are weak and none!

- b) If Charles obtained a 48% in math and a 62% in music, which of the above correlations could we use to ~~best~~ estimate his IQ?

Use Math since ^{better} r value is higher

$$y = 0.4826(48) + 89.6$$

$$= 112.8 \approx 113$$

Question 11

A hockey player is trying to determine if there is a link between the number of minutes he plays and the number of points he scores, as well as the number of penalty minutes he incurs.

Here is a table of his last 11 games correlating ice-time and points scored:

Ice-Time (minutes)	18	17	11	23	22	21	18	13	18	12	18	x
Points Scored	2	1	0	3	4	2	1	0	1	1	2	y

- a) Based on this data, can the player make reliable predictions for the number of points he will score during a game? Explain.

$r = 0.850$
 Strong, Positive } Yes he can since the data shows a
Strong, Positive correlation

- b) Based on the regression line obtained for his points scored, how many minutes should the player need to be on the ice in order to score one point?

$$a = 0.26 \quad y = 0.26x - 2.98 \quad \frac{3.98 = 0.26x}{0.26 \quad 0.26}$$

$$b = -2.98 \quad \downarrow = 0.26x - 2.98 \quad x = 15.3 \text{ min}$$

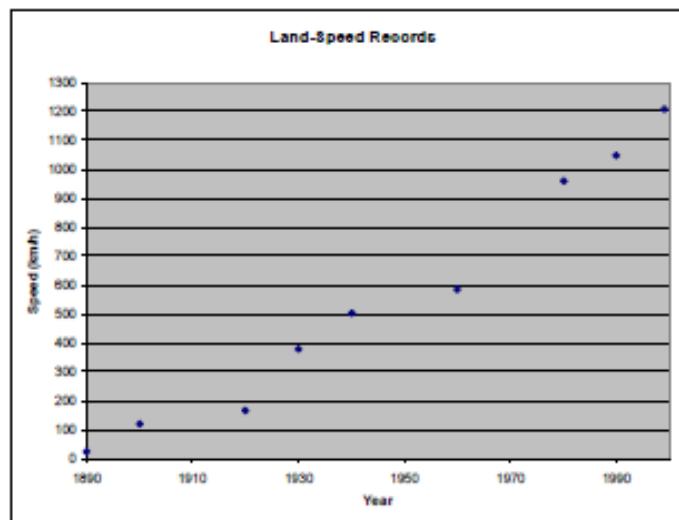
- c) Is it realistic to use this linear model to make these predictions?

No since he could play while injured.

Question 12

The data table and scatter plot below present 9 land-speed records established over the past 109 years by an automobile.

Year	Land-Speed Record (km/h)
1890	24
1900	120
1920	165
1930	380
1940	502
1960	584
1980	961
1990	1048
1999	1208



- a) Establish and qualify the statistical relationship between the record speed and the year it was established.

$$r = 0.985 \rightarrow \text{Strong, Positive}$$

- b) Using the regression line, determine the speed record that would have been set in 1974.

$$a = 10.81373253$$

$$b = -20484.05073$$

$$y = 10.8137x - 20484.05$$

$$x = 1974$$

$$y = 10.8137(1974) - 20484.05$$

$$= 862.19$$

$$\approx 862 \text{ km/h}$$