Feet-propelled scooters

Introduction

Feet-propelled scooters have become very popular. These scooters require balance in order to ride on one leg, whilst pushing with the other. They seem to be as safe as bicycles. Therefore similar precautions apply. Those scooters were designed so that they could be ridden over flat and smooth surfaces. They can be dangerous if they are ridden over holes or cracks, because the wheels may get stuck in them. I wanted to carry out this experiment on a smooth, flat surface as actual roads were too dangerous because there were a lot of cars. Scooters are a very useful way of travelling in town without polluting. They are useful when it comes to travelling short distances (same uses as bikes). However one of their disadvantages is that it is very tiring to ride one, as it relies on the feet. This is why they are more convenient only for short distances.

Statement of task

The aim of this project is to find out what factor influences the speed of a person riding a foot-propelled scooter. This project will define whether the height or the weight of a person influences the speed, down a slope of 45 meters without braking, or speeding up with the aid of feet.

Hypothesis: The person who weighs the less will go faster than a heavier person. The fastest speed will be gained by the shortest person.

 H_{o} null hypothesis. Height and speed are independent. Weight and speed are independent.

H₁ alternative hypothesis. Height and speed are dependent. Weight and speed are dependent.

Plan of Investigation

- The data was collected by testing 12 people of different gender: six boys and six girls in the D2 class in order to get a range of similar heights and weights. The height and weight of each person were measured (figure 1). The data was represented on a scatter graph and the correlation of the graph was found (Figure 2)
- II. Height and weight values were classed in terms of intervals in order to organize the data by grouping it together (figure 3.0 and 4.0). One the data was summarized in a frequency table; it was represented diagrammatically with histograms (figure 3.1 and 4.1).
- III. A distance of 45 meters was measured down a slope and marked with the aid of a chalk, the path that each student was to follow (the same surface was used for the whole experiment). A letter from A to L was given to each person in order to organize the results, making them easier to place in a table and to understand.

- IV. Each student was asked to go to the top of the 45 meters with the scooter and follow the path drawn without braking or accelerating. The time they took to reach the end of the distance was calculated using a stopwatch. The results were then recorded in a table (see appendix). This process was repeated for each person during four tries.
- V. Before comparing and analyzing the result, each person's mean time was calculated by adding all their tries and dividing those by four. The speed of each person was calculated by dividing the time by the distance. (see figure 5)
- VI. Then the speed values were classed in intervals and placed in a frequency table (figure 5.1), it was then represented diagrammatically with a histogram (figure 5.2).
- VII. Two other scatter graphs were drawn to show and compare the data. One was drawn to compare the height versus speed. The second one was drawn to compare the weight versus speed. The regression line was then found for each of these and it was then made possible to see if there was any correlation between each set of data. The equation of the regression line was also found using the formula:

$$y - \overline{y} = \underbrace{Sxy}_{Sx^2} (x - \overline{x})$$

VIII. The spread of the samples was measured by calculating the standard deviation. It measures the deviation between the scores and the mean. The standard deviation of the time, the height and the weight were calculated using this formula:

$$\frac{\sqrt{\sum f \times (x-\overline{X})^2}}{n}$$

In order to do this, the midpoint of each class interval was determined,. The differences between the scores and the mean were squared and the average of these squares was then found. The standard deviation was then established by calculating the square root of this average. The coefficient of variation was then calculated in order to compare the spread of the data.

IX. In order to prove the hypothesis, the x² test of independence was applied. The formula below was used:

$$X_{calc}^2 = \sum_{e} \frac{(f_o - f_e)^2}{f_e}$$

Where f_0 is the observed frequency and f_e is the expected frequency.

Results and data collection

The variables I will be using are continuous values because they are a result of measuring. Figure 1 is the table used to record each person's height and weight. I found out the mean in order to see what the average height and weight of this group of student is. I did this by adding the heights and dividing this number by 12, as there are 12 students who are being investigated on. The same was done to find the mean of weight.

Figure 1 Table to show the height and weight of each 12 students

AT THE BERGY OF	Height (cm)	Weight (kg)
A	180	66
В	165	50
С	166	52
D	151	50
E	153 172	47
F		63
G	195	87
Н	173	65
I	187 183 170	92
J		87
K		63
L	163	67

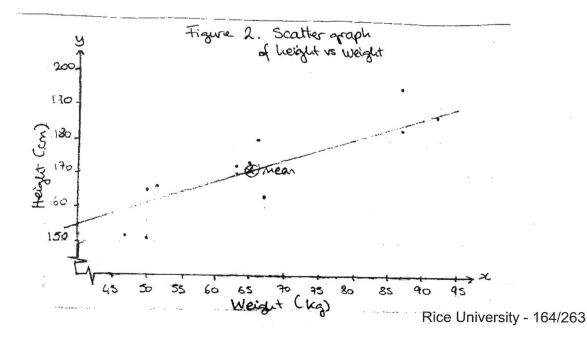
Average height: Σ height / 12 Average weight: Σ weight / 12 = 2058 / 12= 171.5 cm

The average height was found to be 171.5cm and the average weight was found to be 65.8 kg.

= 789 / 12

= 65.75 kg

The scatter graph below shows the height of each student versus the weight. This was done in order to see the data in a visual way and to see if there is any correlation between the height and the weight of a person. We can see that a correlation exists between the height and the weight. The taller a person is, the heavier he/she is.



The equation of the line of regression was calculated using the method below:

Figure 2 – Scatter graph of height versus weight through the points (80; 180), (65;170), (85.5;185)

Using

$$y - \vec{y} = \frac{Sxy}{Sx^2} (x - \overline{x})$$
 we get: $y - \frac{535}{3} = \frac{161.7}{225.2} (x-76.8)$

y=0.72x+123.2 this shows a positive correlation

(tables and calculations in appendix)

Height and weight now need to be described as a matter of what kind of distribution they are. In order to do this description, the height and weight values were classed in terms of intervals and placed in frequency tables as shown below.

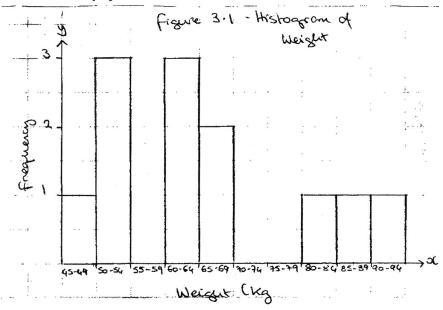
Figure 3.0

Weight interval Frequency 45 - 4950 - 543 55 - 590 3 60 - 6465 - 692 70 - 740 75 - 790 80 - 841 85 - 891 90 - 94 1

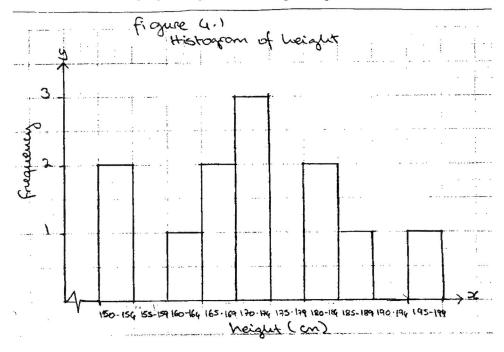
Figure 4.0

Height interval	Frequency
150 – 154	2
155 – 159	0
160 – 164	1
165 – 169	2
170 – 174	3
175 – 179	0
180 – 184	2
185 – 189	1
190 – 194	0
195 – 199	1

It is now possible to place this information on histograms. Below is figure 3.1, showing a histogram of the weight in kilograms and figure 4.1, which shows a histogram of height in centimeters.



As shown on this histogram, we can see that the distribution is positively skewed. It shows that most of the people weigh between 45kg - 69kg.



By looking at the histogram above, we can see that the height is normally distributed. The majority of the people lie around 170 cm - 174 cm. These interpretations of the histograms will later be proven using the standard deviation.

Figure 5 - Table showing the mean time and speed of each person

	Mean time	speed	speed
		calculations	m/
A	13.94	45/13.94	3.23
В	10.73	45/10.73	4.19
C	10.65	45/10.65	4.23
D	10.2	45/10.20	4.41
E	10.39	45/10.39	4.33
F	12.39	45/12.39	3.63
G	15.17	45/15.17	2.97
H	13.26	45/13.26	3.39
I	14.70	45/14.70	3.06
J	14.04	45/14.04	3.21
K	10.72	45/10.72	4.20
L	11.71	45/11.71	3.84

Average speed: $\frac{44.69}{12} = 3.72$

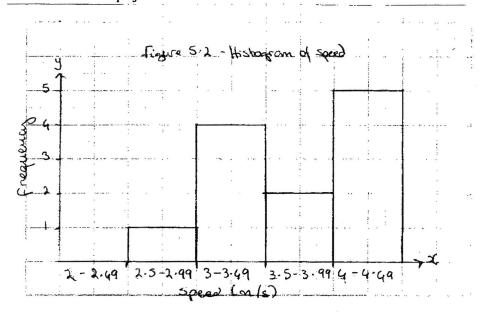
The fastest speed was 4.41m/seconds and the fastest mean time 10.2seconds were achieved by person D. The slowest speed was 2.97m/seconds and the slowest mean time 15.17 seconds and was done by person G. If we now look back at figure 1, we see that person D is the shortest and person G is the tallest and heavier.

The speed was then classified in intervals, in order to make it easier to see which one was the more frequent.

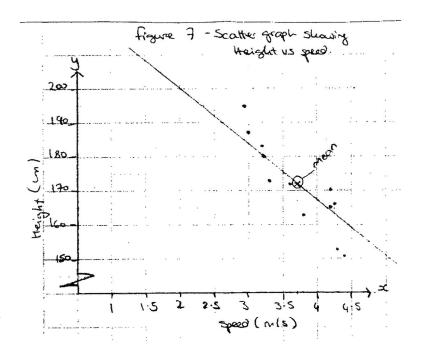
Figure 5.1- speed frequency table

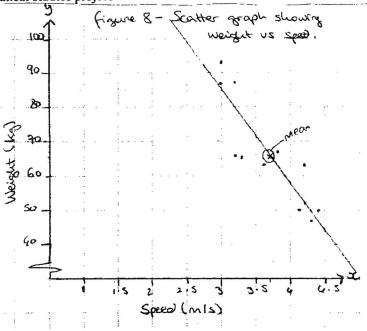
Speed interval	Frequency
2.0 - 2.49	0
2.5 - 2.9	1
3.0 - 3.49	4
3.5 – 3.99	2
4.0 - 4.49	5

Using this table, we can now draw a histogram to represent this data.



As represented on the histogram, the most frequent speed was between 4 to 4.49 m/s. It is now time to see which type of correlation, if any, there is between the height and the speed (figure 7), and between the weight and the speed (figure 8). The graphs below represent this data.





As we can see from these graphs, there is a moderate linear negative correlation for both of the graphs. This means that as one decrease, the other increases. We can therefore say that the shorter some one is, the faster they will go.

The lines of best fit were drawn through the mean of the values. The equations for the regression lines were found using the least square regression formula.

See appendix to see the table and the calculations for this equation

Figure 7 – scatter graph of height versus speed through the points (3.7; 171.5), (3.23;180) and (2,200)

Using

$$y - \overline{y} = \underline{Sxy}(x - \overline{x})$$
 we get: $y - \underline{551.5} = (-\underline{25.68})(x-2.98)$
 $y = -16.68x + 233.52$

This shows a negative correlation.

Figure 8 – scatter graph of weight versus speed through the points (3.72;65.75), (3;85) and (4.23,52)

Using

$$y - \overline{y} = \underline{Sxy}(x - \overline{x})$$
 we get: $y - \underline{202.75} = \underline{(-20.49)}(x-3.65)$
= -26.87x + 165.67

This shows that the graph has a negative correlation.

In order to calculate how spread the data is, we need to use a formula called the standard deviation.

Standard deviation =
$$\sqrt{\sum f \times (x - \overline{x})^2}$$

However, as shown on the formula, a number of things such as the mean etc. need to be calculated before we can use it. Below is a table showing the calculations for the standard deviation of height.

Figure 9 - table to show the calculations for the standard deviation of height

Class interval	Midpoint (x)	Frequency (f)	f×x	x - Z	$(x-\overline{X})^2$	$f \times (\mathbf{x} - \overline{\mathbf{x}})^2$
150-154	152	2	304	-20	400	800
155-159	157	0	0	-15	225	0
160-164	162	1	162	-10	100	100
165-169	167	2	334	-5	25	50
170-174	172	3	516	0	0	0
175-179	177	0	0	5	25	0
180-184	182	2	364	10	100	200
185-189	187	1	187	15	225	225
190-194	192	0	0	20	400	0
195-199	197	1	197	25	625	625

2 - 2000

Mean height
$$(\overline{x}) = \underbrace{\sum f \times x}_{n}$$
 standard deviation $(s) = \underbrace{\sqrt{\sum f \times (x - \overline{x})^2}}_{n}$ standard deviation $= \underbrace{\sqrt{2000}}_{12}$ standard deviation $= 12.9$

= 172 cm standard deviation of height = 12.9cm Whilst the range of heights gathered was 151 cm - 195 cm, when using standard deviation, a better measure of dispersion is $\overline{x} \pm 1$ standard deviation which is $159 \le h \le 185$.

Figure 10 - table to show the calculations for the standard deviation of weight

Below is the table which shows the calculations required to find out the standard deviation of weight.

Class interval	Midpoint (x)	Frequency (f)	f×x	x - \overline{X}	(x-₹)²	$f \times (x-\overline{X})^2$
45-49	47	1	47	-19.2	368.64	368.64
50-54	52	3	156	-14.2	201.64	604.92
55-59	57	0	0	-9.2	84.64	0
60-64	62	2	124	-4.2	17.64	35.28
65-69	67	3	201	0.8	0.64	1.92
70-74	72	0	0	5.8	33.64	0
75-79	77	0	0	10.8	116.64	0
80-84	82	0	0	15.8	249.64	0
85-89	87	2	174	20.8	432.64	865.28
90-94	92	1	92	25.8	665.64	665.64
95-99	97	0	0	30.8	948.64	0

Mean weight
$$(\overline{x}) = \frac{\sum f \times x}{n}$$
 standard deviation (s) = $\frac{\sqrt{\sum f \times (x-x)^2}}{n}$ = $\frac{794}{12}$ = $\frac{\sqrt{2541.68}}{12}$ Standard deviation of weight = 14.55 kg

The standard deviation of weight: $51.65 \le h \le 80.75$ was a better measure of dispersion than the range of weights which were gathered from the beginning of this experiment: 47kg - 92kg.

The standard deviations of the height and the weight have now being calculated; we will now look at the standard deviation of the speed.

Figure 11 - table to show calculations for the standard deviation of speed

Class interval	Midpoint (x)	Frequency (f)	f×x	x - 7	(x-\overline{x})2	$f \times (x-\overline{X})^2$
2-2.49	2.245	0	0	-1.455	2.12	0
2.5-2.99	2.745	1	2.745	-0.955	0.91	0.91
3.0-3.49	3.245	4	12.98	-0.455	0.21	0.84
3.5 - 3.99	3.745	2	7.49	0.045	0.002	0.004
4.0 - 4.49	4.245	5	21.225	0.545	0.3	1.5

Mean speed
$$(\overline{x}) = \underline{\sum f \times x}$$
 standard deviation (s) = $\underline{\sqrt{\sum f \times (x \cdot \overline{x})^2}}$ n = $\underline{44.44}$ 12 = 3.70 standard deviation of speed = 0.52m/seconds

We can see that the range of speed gathered at the beginning was not wide: 2.97m/s - 4.41m/s, however using the standard deviation allowed us to have a better measure of dispersion: $3.18\text{m/s} \le h \le 4.22\text{m/s}$.

The use of standard deviation allowed me to see that the heights and weights were widely dispersed but the speed was not. This means that speed was not necessarily dependent on height or weight.

Since the spread of each set of data has been found out, we can now compare them. This is done using the coefficient of variation ¹

$$V = \underline{s} \times 100\%$$
 v is the coefficient of variation
 X s is the standard deviation of the data
 X is the mean of the data

Height	Weight	Speed
$V = \frac{12.9}{172} \times 100$	$v = \frac{14.5}{66.2} \times 100$	$v = \frac{0.4259}{2.66} \times 100$
= 7.5%	= 21.9%	= 16.01%

These results show the ratio of the standard deviation to the mean of the data as a percentage. Since weight has a greater coefficient of variation, it has a greater relative dispersion.

X² test of independence

This test is used to find if two factors are independent. It examines the difference between observed and expected values. If any difference is found, even if it is small, it indicates that the two factors are independent.

(see appendix for the table which shows the calculations for the X^2 test of independence for height and speed.)

"Are height and speed independent?"

¹ Mathematics for international student, Mathematical Studies SL International Baccalaureate Diploma Programme, HAESE & HARRIS PUBLICATION, Chapter 5, page 153: comparing the spread.

$$X_{calc}^{2} = \sum_{e} \frac{(f_{o} - f_{e})^{2}}{f_{e}}$$
= 1.892242

Considering a 5% significant level and referring to the table of critical values, the rejection inequality is $X^2_{calc} > 19.675$. Since $X^2_{calc} < 19.675$, we accept H_0 , height and speed are independent factors.

"Are weight and speed independent?"

$$X^{2}_{calc} = \frac{\sum (f_{o} - f_{e})^{2}}{f_{e}}$$
$$= 5.327$$

Considering a 5% significant level and referring to the table of critical values, the rejection inequality is $X_{calc}^2 > 19.675$. Since $X_{calc}^2 < 19.675$, we accept H_0 , height and speed are independent factors.

Conclusion and evaluation

As seen from all the calculations, graphs and the X^2 test of independence, the hypothesis was proved to be wrong. Height, weight and speed are independent. The speed of a person riding down a slope does not depend on their weight or height.

The histograms were a good way to show the data in a visual way to make it easier to see the average number. However, in order to be more accurate, a table showing the frequencies such as relative frequency could have been drawn. Finding the mean time of the people was a very useful calculation to do, as most of the other mathematics used required this information. It was also another way to show the average time people did. Calculating the speed of each person was also an efficient calculation as it gives an accurate idea of how fast someone with similar body proportions could go. Standard deviation was the best measure to use in order to differentiate the spread of the data. The other measures to calculate the spread of the data could be by finding out the range and inter-quartile range. However, these were not appropriate for this set of data, as they do not include all the values in the calculations. The least square regression formula was very efficient to use, as it does not only show the equation of the line but also its correlation. However Pearson's correlation coefficient could have been calculated in order to see how strong the association between the two variables is. The chi-square test of independence was the method, which proved to be the most accurate in means of proving the hypothesis.

APPENDIX

Table showing the time and mean time of each person

	First try	Second try	Third try	Fourth try	Mean time
A	13.91	13.85	14.12	13.89	13.94
В	10.59	10.54	10.37	11.44	10.73
C	10.53	11.02	10.49	10.55	10.65
D	10.24	09.96	10.16	10.43	10.2
E	10.37	10.13	10.29	10.78	10.39
F	12.48	12.52	12.31	12.24	12.39
G	15.32	14.98	15.26	15.13	15.17
H	13.52	12.79	13.43	13.31	13.26
I	14.68	14.24	14.77	15.09	14.70
J	13.86	14.17	13.89	14.24	14.04
K	10.35	10.69	10.87	11.00	10.72
L	11.72	11.78	11.64	11.70	11.71

Table showing the observed values for X² test of independence of height and speed

	A	В	С	D	E	F	G	Н	I	J	K	L	Totals
Speed	3.23	4.19	4.23	4.41	4.33	3.63	2.97	3.39	3.06	3.21	4.2	3.84	44.69
leight	180	165	166	151	153	172	195	173	187	183	170	163	2058
otals	183.23	169.19	170.23	155.41	157.33	175.63	197.97	176.39	190.06	186.21	174.20	166.84	2102.69

12 × 2 contingency table

Table showing the expected values for X2 test of independence of height and speed

	Α	В	С	D	E	F	G	Н	1	J	K	L	Totals
Speed	3.89	3.60	3.62	3.30	3.34	3.73	4.21	3.75	4.04	3.96	3.70	3.55	44.69
Height	179.34	165.59	166.61	152.12	154.0	171.90	193.76	172.64	186.02	182.25	170.50	163.29	2058.02
Totals	183.23	169.19	170.2	155.42	157.34	175.6	198	176.39	190.06	186.2	174.2	166.84	2102.67

Fo	fe	fo-fe	(fo-fe) ²	(fo-fe) 2/ f
3.23	3.89	-0.66	0.4356	0.111979
4.19	3.6	0.59	0.3481	0.096694
4.23	3.62	0.61	0.3721	0.10279
4.41	3.3	1.11	1.2321	0.373364
4.33	3.34	0.99	0.9801	0.293443
3.63	3.73	-0.1	0.01	0.002681
2.97	4.21	-1.24	1.5376	0.365226
3.39	3.75	-0.36	0.1296	0.03456
3.06	4.04	-0.98	0.9604	0.237723

3.21	3.96	-0.75	0.5625	0.142045
4.2	3.7	0.5	0.25	0.067568
3.84	3.55	0.29	0.0841	0.02369
180	179.34	0.66	0.4356	0.002429
165	165.59	-0.59	0.3481	0.002102
166	166.61	-0.61	0.3721	0.002233
151	152.12	-1.12	1.2544	0.008246
153	154	-1	1	0.006494
172	171.9	0.1	0.01	5.82E-05
195	193.76	1.24	1.5376	0.007936
173	172.64	0.36	0.1296	0.000751
187	186.02	0.98	0.9604	0.005163
183	182.25	0.75	0.5625	0.003086
170	170.5	-0.5	0.25	0.001466
163	163.29	-0.29	0.0841	0.000515

1.892242

Table showing the observed values for X2 test of independence of weight and speed

	Α	В	С	D	E	F	G	Н	ı	J	К	L	Totals
Speed	3.23	4.19	4.23	4.41	4.33	3.63	2.97	3.39	3.06	3.21	4.2	3.84	44.69
Weight	66	50	52	50	47	63	87	65	92	87	63	67	789
Totals	69.23	54.19	56.23	54.41	51.33	66.63	89.97	68.39	95.06	90.21	67.2	70.84	833.69

Table showing the expected values for X2 test of independence of height and speed

	A	В	С	D	E	F	G	Н	I	J	ĸ	L	Totals
Speed	3.71	2.90	3.01	2.92	2.75	3.57	4.82	3.67	5.10	4.84	3.60	3.80	44.69
Weight	65.52	51.29	53.22	51.49	48.58	63.06	85.15	64.72	89.10	85.37	63.60	67.04	788.14
Totals	69.23	54.19	56.23	54.41	51.33	66.63	89.97	68.39	94.20	90.21	67.20	70.84	832.83

fo	fe	fo-fe	(fo-fe) ²	(fo-fe) ² / fe
3.23	3.71	-0.48	0.2304	0.062102
4.19	2.90	1.29	1.6641	0.573828
4.23	3.01	1.22	1.4884	0.494485
4.41	2.92	1.49	2.2201	0.760308
4.33	2.75	1.58	2.4964	0.907782
3.63	3.57	0.06	0.0036	0.001008
2.97	4.82	-1.85	3.4225	0.710062
3.39	3.67	-0.28	0.0784	0.021362
3.06	5.10	-2.04	4.1616	0.816
3.21	4.84	-1.63	2.6569	0.548946
4.20	3.60	0.6	0.36	0.1
3.84	3.80	0.04	0.0016	0.000421
66	65.52	0.48	0.2304	0.003516
50	51.29	-1.29	1.6641	0.032445
52	53.22	-1.22	1.4884	0.027967

50	51.49	-1.49	2.2201	0.043117
47	48.58	-1.58	2.4964	0.051387
63	63.06	-0.06	0.0036	5.71E-05
87	85.15	1.85	3.4225	0.040194
65	64.72	0.28	0.0784	0.001211
92	89.10	2.9	8.41	0.094388
87	85.37	1.63	2.6569	0.031122
63	63.60	-0.6	0.36	0.00566
67	67.04	-0.04	0.0016	2.39E-05

5.327395

Least squares regression

For figure 2- height vs weight

x	Υ	xy	x ²
80	180	14400	6400
65	170	11050	4225
85.5	185	15817.5	7310.25
230.5	535	41267.5	17935.25

$$Sxy = \Sigma xy - (\Sigma x)(\Sigma y)$$
n
$$Sxy = 41267.5 - (230.5)(535)$$

$$Sxy = 161.7$$

$$Sx^{2} = \Sigma x^{2} - (\Sigma x)^{2}$$

$$Sx^{2} = 17935.25 - (230.5)^{2}$$

$$Sx^{2} = 225.2$$

$$\bar{x} = \frac{\Sigma x}{3} = \frac{230.5}{3} = 76.8$$

For figure 7 - height vs speed

X	Y	xy	x2
3.7	171.5	634.55	13.69
3.23	180	581.4	10.43
2	200	400	4
8.93	551.5	1615.95	28.12

$$Sxy = \Sigma xy - (\Sigma x)(\Sigma y)$$
n
$$Sxy = 1615.95 - (8.93)(551.5)$$

$$Sxy = -25.68$$

$$Sx^{2} = \Sigma x^{2} - (\Sigma x)^{2}$$
n
$$Sx^{2} = 28.12 - (8.93)^{2}$$
1
$$Sx^{3} = 28.12 - (8.93)^{2}$$

$$\bar{x} = \frac{\sum x}{3} = \frac{8.93}{3} = 2.98$$

$$y - \overline{y} = \underbrace{Sxy}_{Sx^2} (x - \overline{x}) \quad \text{we get} : y - \underbrace{551.5}_{3} = (-25.68)(x-2.98)$$
$$= -16.68x + 183.3$$

For figure 8 - weight vs speed

X	У	ху	x2
3.72	65.75	244.59	13.84
3	85	255	9
4.23	52	219.96	17.89
10.95	202.75	719.55	40.73

$$\mathbf{S}\mathbf{x}\mathbf{y} = \Sigma \mathbf{x}\mathbf{y} - \underline{(\Sigma \mathbf{x})(\Sigma \mathbf{y})}$$

$$Sxy = 719.55 - \underbrace{(10.95)(202.75)}_{3}$$

$$Sxy = -20.49$$

$$Sx^2 = \Sigma x^2 - (\Sigma x)^2$$
 $Sx^2 = 40.73 - (10.95)^2$
 $Sx^2 = 0.7625$

$$\bar{x} = \underline{\Sigma x} = \underline{10.95} = 3.65$$