Solutions to Exercises - chapter 2

2.1) Explain what is meant by an indeterminate and a determinate error, giving examples of each.

Indeterminate (or random) errors are those which cause a random distribution of the data around a mean point. Errors of this type are normally associated with the net effect of a number of small unpredictable fluctuations that may not readily be identified or eliminated. Errors of this type lead to poor precision.

Indeterminate or random errors arise from a number of small unpredictable variations, which in some cases add or sum together; in other cases these may partially nullify each other. The source of error may be due to many factors such as human error, fluctuations in temperature or small differences in the quantities of reagents used. Since there are several different sources of error which may sometimes randomly lower or raise the reading, the data are scattered around the true value.

Determinate (or systematic) errors, however, cause all of the data to be shifted in one direction. The results are therefore typically shifted to values that are either all too low or all too high. Errors of this type lead to poor accuracy.

Determinate or systematic errors cause all of the data to be shifted in one direction. This behaviour is caused by the same type of error which keeps occurring every time a measurement is made. An example of a determinate error is if an analytical top-pan balance is not zeroed or *tarred* prior to the first measurement, and so gives a reading of, for example, of 0.5g when nothing is placed on the pan. Every mass of a material that is subsequently weighed on the balance will in fact be 0.5g *less* than the value recorded on the balance.

2.2) How many significant figures do each of the following data values possess?

(a) 7.9×10^{5}, (b) 300.45, (c) 5.043×10^{-4}.

All of the figures are data directly quoted not involving estimates. The data can for each example can therefore be quoted as:

(a) 79000, (b) 300.45, (c) 0.00054043

2.3) Replicate samples for the iron content of an alloy were determined to contain 94.67, 94.54, 94.62 and 94.93% Fe. Calculate the standard deviation and the relative standard deviation of these analyses.

Standard deviation for the figures above:

$$s = \sqrt{\frac{\sum_{1=1}^{N} x_1^{\,2} - \frac{(\sum_{i=1}^{N} x_i)^2}{N}}{N-1}}$$

Step 1: Calculate $\sum_{i=1}^{N} x_1^{\ 2}$

$\sum_{i=1}^{N} x_i^{\ 2}$ = (94.67)^2 + (94.54)^2 + (94.62)^2 + (94.93)^2

$\therefore \sum_{i=1}^{N} x_i^{\ 2}$ = 8962.41 + 8937.81 + 8952.94 + 9011.7

$\therefore \sum_{i=1}^{N} x_i^{\ 2}$ = 35864.86

Step 2: Calculate $(\sum_{i=1}^{N} x_i)^2$

$\therefore (\sum_{i=1}^{N} x_i)^2$ = (94.67 + 94.54 + 94.62 + 94.93)^2

= (378.76)^2

= 143459.13

N = 4 since there are 4 data points, so:

$$s = \sqrt{\frac{35{,}864.86 - \frac{143{,}459.13}{4}}{4-1}}$$

s = 0.16 % Fe

2.4) Using the data for Problem 2.3, calculate the 90% and 95% confidence limits for the mean of this data.

Method: Calculate the confidence limits according to Eqn 2.15 in book:

Mean of data = (94.67 + 94.54 + 94.62 + 94.93) / 4 = 94.69 % Fe

CL 90% = 94.69 ± $\frac{2.235\text{x}0.16}{\sqrt{4}}$ = ±0.18 % Fe

CL 95% = 94.69 ± $\frac{3.182\text{x}0.16}{\sqrt{4}}$ = ±0.25 % Fe

2.5) Five Potassium chromate samples were weighed with the following results; (a) 123.3 (b) 124.2, (c) 121.5, (d) 123.6 and (e) 124.1g. Calculate the median, mean and the range of data.

Median: There are an odd number of data points when arranged arithmetically. 123.6g is the datum value that falls in the middle of the data set.

The median is therefore 123.6g potassium chromate

Mean: $\bar{x}$ = (123.3 + 124.2 + 121.5 + 123.6 + 124.1) / 5 = 123.3g potassium chromate

Range: Highest mass = 124.2g potassium chromate

Lowest mass = 121.5g potassium chromate

Range = Highest mass – lowest mass = 124.2 – 121.5g = 2.7g

2.6) Two samples of potassium chloride were weighed using an analytical balance and were recorded as being 34.5645g, and 35.5664g respectively. Express these figures to four decimal places.

34.56g and,

34.57g

2.7) A burette has calibrated divisions of 0.1cm^3; when performing a titrimetric analysis titre results are recorded between 10.5 and 10.7

cm^3. To how many significant figures should the titre values be recorded? Explain your reasoning.

A direct reading to within 0.1cm^3 can me made. An estimate can be made to one further significant figure.

The readings should therefore be recorded to four significant figures – ie:

10. _ _ cm^3.

2.8) Five samples of soil were weighed prior to analysis: The weight of these samples were recorded as:

(a)	**23.67g**
(b)	**34.53g**
(c)	**31.56g**
(d)	**26.34g**
(e)	**42.19g**

Calculate the mean and median for the weight of these five samples.

The mean of these readings is:

(23.67 + 34.53 + 31.56 + 26.34 + 43.19) / 5 g = 31.86g

The median is the middle number when arranged in ascending order:

ie: 23.67 + 26.34 + <u>31.56</u> + 34.53 + 43.19

The median is therefore 31.56g.

2.9) A water sample taken from a lake is analysed for its cadmium content. Six replicate measurements were recorded as below:

Cadmium Content (ppm)

(a)	**20.2**
(b)	**18.5**
(c)	**21.4**
(d)	**19.2**
(e)	**21.8**

(f) 18.8

Calculate the spread of the data.

The lowest value in this data set is: 18.5ppm

The greatest value in this data set is: 21.8g

The spread of the data is therefore: 21.8-18.5ppm = 3.3ppm

2.10) Calculate the relative standard deviation of the data of problem 2.9.

Standard deviation for:

Cadmium Content (ppm)	
(a)	20.2
(b)	18.5
(c)	21.4
(d)	19.2
(e)	21.8
(f)	18.8

$$s = \sqrt{\frac{\sum_{1=1}^{N} x_1^{\ 2} - \frac{(\sum_{i=1}^{N} x_i)^2}{N}}{N-1}}$$

Step 1: Calculate $\sum_{i=1}^{N} x_1^{\ 2}$

$$\sum_{i=1}^{N} x_i^{\ 2} = (20.2)^2 + (18.5)^2 + (21.4)^2 + (19.2)^2 + (21.8)^2 + (18.8)^2$$

$$\therefore \sum_{i=1}^{N} x_i^{\ 2} = 408.04 + 342.25 + 457.96 + 368.64 + 475.24 + 353.44$$

$$\therefore \sum_{i=1}^{N} x_i^{\ 2} = 2405.57$$

Step 2: Calculate $(\sum_{i=1}^{N} x_i)^2$

$$\therefore (\sum_{i=1}^{N} x_i)^2 = (20.2 + 18.5 + 21.4 + 19.2 + 21.8 + 18.8)^2$$

$$= (119.9)^2$$

= 14376.01

N = 6 since there are 6 data points, so:

$$s = \sqrt{\frac{2405.57 - \frac{14376.01}{6}}{6-1}}$$

s = 1.38ppm

mean ($\bar{x}$) = (20.2 + 18.5 + 21.4 + 19.2 + 21.8 + 18.8) / 6 = 19.98

RSD = (s/$\bar{x}$)

∴ = 1.38 / 19.98

∴ RSD = 0.7ppm

2.11) Calculate the relative error *in percentage terms* for the analysis of a river water sample that gives a value of 15.7ppm Cu when the true value is, in fact, 18.0ppm Cu.

The true value, x_t = 18.0ppm

The measured value x_i = 17.7ppm

The relative error E_r in percentage terms = $\frac{18.0 - 17.7}{17.7}$ *x* 100 = 1.69%

2.12) An acid base titration gives the concentration for a HCl solution to be 0.104M. The true concentration was in fact 0.110M. Express the relative error for this analysis in parts per thousand terms.

The true value, x_t = 0.110M

The measured value x_i = 0.104

The relative error E_r in parts per thousand terms = $\frac{0.110 - 0.104}{0.11}$ *x* 1000 =

54.54 ppt

2.13) Calculate the coefficient of variance for the following replicate measurements for the iron content of a water sample:

Fe^{2+} content for water sample (ppm)

(a)	34.6
(b)	29.5
(c)	32.2
(d)	33.7
(e)	34.6
(f)	32.4
(g)	35.1

Relative standard deviation expressed in percentage terms is the same as the coefficient of variance or CV:

Standard deviation:

$$s = \sqrt{\frac{\sum_{1=1}^{N} x_1^{\ 2} - \frac{(\sum_{i=1}^{N} x_i)^2}{N}}{N-1}}$$

Step 1: Calculate $\sum_{i=1}^{N} x_1^{\ 2}$

$\sum_{i=1}^{N} x_i^2$ = (34.6)^2 + (29.5)^2 + (32.2)^2 + (33.7)^2 + (34.6)^2 + (32.4)^2 + (35.1)^2

$\therefore \sum_{i=1}^{N} x_i^2$ = 1197.16 + 870.25 + 1036.84 + 1135.69 + 1197.16 + 1049.76 + 1232.01

$\therefore \sum_{i=1}^{N} x_i^2$ = 7718.87

Step 2: Calculate $(\sum_{i=1}^{N} x_i)^2$

$\therefore (\sum_{i=1}^{N} x_i)^2$ = (34.6 + 29.5 + 32.2 + 33.7 + 34.6 + 32.4 + 35.1)^2

= (232.1)^2

= 53870.41

N = 7 since there are 7 data points, so:

$$s = \sqrt{\frac{7718.87 - \frac{53870.41}{7}}{7-1}}$$

s = 1.96ppm

mean ($\bar{x}$) = (34.6 + 29.5 + 32.2 + 33.7 + 34.6 + 32.4 + 35.1) / 7 = 33.157

RSD = (s/$\bar{x}$)

$\therefore$ = 1.96 / 33.157

$\therefore$ RSD = 0.059ppm

$\therefore$CV = 0.059 x100 = 5.9%

2.14) Calculate the variance for the data of problem 2.13.

The variance is equal to the square of standard deviation.

$\therefore$ the variance = $(1.96)^2$ = 3.6ppmm^2

2.15) A series of replicate measurements for the water content in a sample of ethanol by the Karl-Fischer approach gave the following data:

(a)	**0.77%**
(b)	**0.67%**
(c)	**0.71%**
(d)	**0.90%**
(e)	**0.78%**

With what confidence may point (d) be rejected if the *Q*-test is used?

Q-test Approach:

$$Q_{exp} = \frac{d}{w} = \frac{x_q - x_n}{x_h - x_l}$$

Step 1:

x_q = 0.90% - as the suspect data value.
x_n = 0.78% - as the nearest neighbouring value.
x_h = 0.90% as the highest data value.
x_i = 0.67% as the lowest data value.

Step 2: Compare Q_{exp} with the Q-test table for the appropriate values corresponding to 5 data points.

Q_{exp} = 0.52

There are five data points. From the table 2.1 we can see that we cannot safely reject data point (d) even with (the lowest criteria) a 90% confidence.

2.16) Six replicate measurements for the concentration of an H_2SO_4 acid concentration are recorded as below:

(a) 0.152M
(b) 0.153M
(c) 0.149M
(d) 0.148M
(e) 0.151M

Calculate the 90% and 95% confidence limits for the mean of the data values.

Method: Calculate the confidence limits according to Eqn 2.15 in book:

Mean of data $\bar{x}$ = (0.152 + 0.153 + 0.149 + 0.148 + 0.151) / 5 = 0.1506 M

Sample standard deviation, *s*, = 0.00207

CL 90% = 0.1506 ± $\frac{2.132 \text{x} 0.00207}{\sqrt{5}}$ = ±0.002 M H_2SO_4

CL 95% = 0.1506 ± $\frac{2.776 \text{x} 0.00207}{\sqrt{5}}$ = ±0.003 M % H_2SO_4

2.17) Calculate the relative standard deviation in percentage terms for the data of problem 2.15.

Standard deviation for:

(a) 0.77%
(b) 0.67%
(c) 0.71%
(d) 0.90%
(e) 0.78%

$$s = \sqrt{\frac{\sum_{1=1}^{N} x_1{}^2 - \frac{(\sum_{i=1}^{N} x_i)^2}{N}}{N-1}}$$

Step 1: Calculate $\sum_{i=1}^{N} x_1{}^2$

$\sum_{i=1}^{N} x_i^{\,2}$ = $(0.77)^2 + (0.67)^2 + (0.71)^2 + (0.90)^2 + (0.78)^2$

$\therefore \sum_{i=1}^{N} x_i^{\,2}$ = 0.5929 + 0.4489 + 0.5041 + 0.81 + 0.6084

$\therefore \sum_{i=1}^{N} x_i^{\,2}$ = 2.9643

Step 2: Calculate $(\sum_{i=1}^{N} x_i)^2$

$\therefore (\sum_{i=1}^{N} x_i)^2 = (0.77 + 0.67 + 0.71 + 0.90 + 0.78)^2$

$= (3.83)^2$

= 14.669

N = 5 since there are 5 data points, so:

$$s = \sqrt{\frac{2.943 - \frac{14.669}{5}}{5-1}}$$

s = 0.048% water content

mean ($\bar{x}$) = (0.77 + 0.67 + 0.71 + 0.90 + 0.78) / 5 = 0.766 % water

RSD = (s/$\bar{x}$)

$\therefore$ = 0.048 / 0.766

$\therefore$ RSD = 0.0626

In percentage terms RSD = 0.0626 x 100 = 6.26%

2.18) Replicate measurements for the chloride content of a water sample gave the following results: (i) 0.81mM, (ii) 0.83mM, (iii)

0.82mM and (iv) 0.91mM. With what confidence can the data point (iv) be rejected as an outlier by (a) the Q-test or (b) the T-test?

(a) Q-test Approach:

$$Q_{exp} = \frac{d}{w} = \frac{x_q - x_n}{x_h - x_l}$$

Step 1:

x_q = 0.91% - as the suspect data value.

x_n = 0.83% - as the nearest neighbouring value.

x_h = 0.91% as the highest data value.

x_i = 0.81% as the lowest data value.

Step 2: Compare Q_{exp} with the Q-test table for the appropriate values corresponding to 5 data points.

<u>Q_{exp} = 0.8</u>

Rejecting with 90% demands a Q_{exp} value of 0.765 for 4 data points (see Table 2.1). Rejecting with 95% certainty demands a Q_{exp} value of 0.829.

It follows that the experimental point (iv) may be rejected with a certainty of >90% but <95%.

(b) T-test Approach:

Step 1: Calculate standard deviation, *s*,

Then calculate T_n

x_q is the suspect data point in question: 0.91% Fe

x_n is the nearest neighbouring data value: ie: 0.83% Fe

s = 0.046% Fe

$T_n = (x_q - x_n) / s$, so:

$T_n = (0.91 – 0.83) / 0.046$

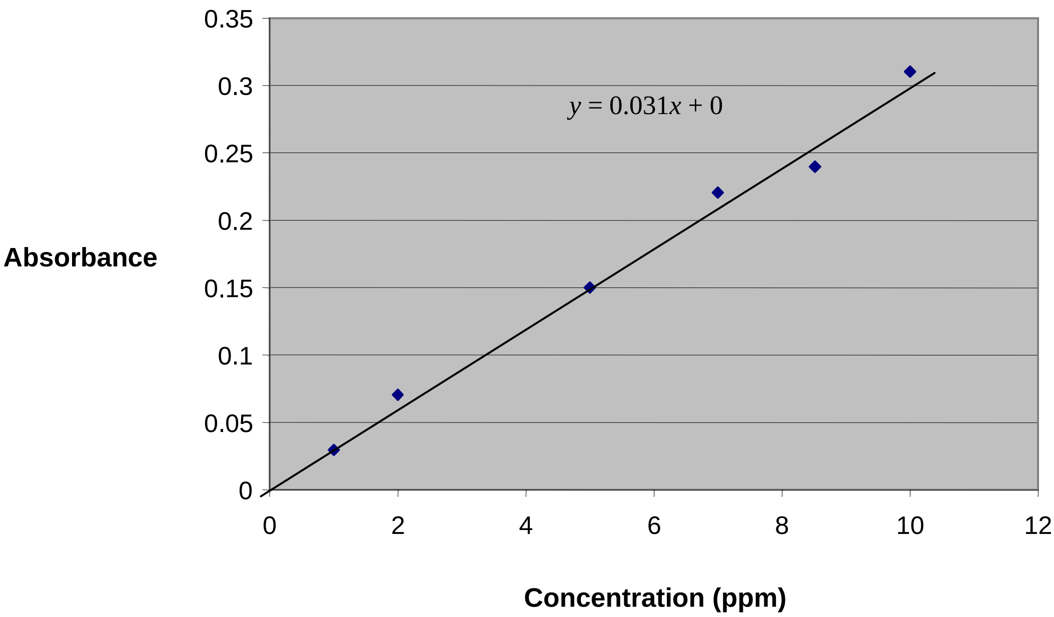
$\therefore T_n = 1.739$

Since the Tn of 1.739 > 1.49 the data point can be rejected with > 99% confidence.

2.19) A UV / visible determination for potassium permanganate gave the following results:

Concentration (ppm)	Absorbance
1	0.03
2	0.07
5	0.15
7	0.22
8.5	0.24
10	0.31

(a) Plot the data in the form of a calibration curve.



(b) Using the least squares fit approach, determine the best fit line for this calibration plot.

$\bar{y}$ = (0.03 + 0.07 + 0.15 + 0.22 + 0.24 + 0.31) / 6 = 0.17

$\bar{x}$ = (1 + 2 + 5 + 7 + 8.5 10) / 6 = 5.58

c must be 0 since a concentration of 0 cannot exhibit any absorption.

So: $\bar{y}$ = m $\bar{x}$

$\therefore m = \bar{y} / \bar{x}$

so m = 0.17 / 5.58 = 0.031

It follows that the equation for the best fit line is:

y = 0.031 x + c

(c) Calculate the Pearson correlation coefficient for this data set.

$$r = \frac{\sum x_i y_i - (\sum x_i)(\sum y_i)/n}{\sqrt{[\sum x_i^2 - (\sum x_i)^2/n][\sum y_i^2 - (\sum y_i)^2/n]}}$$

Method:
(a): Assign *x* and *y* values and calculate each term for Eqn 2.17
(b): Substitute values and calculate *r* and r^2.

Step (a): n = 6 for 6 data points.

Calculate n $\Sigma x_i y_i$

X	Y	X_iY_i
1	0.03	0.03
2	0.07	0.14
5	0.15	0.75
7	0.22	1.54
8.5	0.24	2.04
10	0.31	3.1
		$\Sigma x_i y_i = 7.6$

$\therefore n \Sigma x_i y_i = 6 \times 7.6$

$\therefore n \Sigma x_i y_i = 45.6$

Calculate Σx_i Σy_i:

$\Sigma x_i = 1.0 + 2.0 + 5.0 + 7.0 + 8.5 + 10.0$

$\Sigma x_i = 33.5$

$\Sigma y_i = 0.03 + 0.07 + 0.15 + 0.22 + 0.24 + 0.31$

$\Sigma y_i = 1.02$

$\Sigma x_i \Sigma y_i = 33.5 \times 1.02$

$\Sigma x_i \Sigma y_i = 34.17$

Calculate $[n(\Sigma x_i^2) - (\Sigma x_i)^2]$

$\Sigma x_i^2 = (1.0)^2\ (2.0)^2 + (5.0)^2 + (7.0)^2 + (8.5)^{2\,+}\ (10.0)^2$

$\Sigma x_i^2 = 1 + 4 + 25 + 49 + 72.25 + 100$

$\underline{\Sigma x_i^2 = 251.25}$

$(\Sigma x_i)^2 = 33.5^2$

$\underline{(\Sigma x_i)^2 = 1{,}122.25}$

$[n(\Sigma x_i^2) - (\Sigma x_i)^2] = 6(251.25) – 1{,}122.25$

$[n(\Sigma x_i^2) - (\Sigma x_i)^2] = 1{,}507.5 – 1{,}122.25$

$\underline{[n(\Sigma x_i^2) - (\Sigma x_i)^2] = 385.25}$

Calculate $[n(\Sigma y_i^2) - (\Sigma y_i)^2]$:

$\Sigma y_i^2 = (0.03)^2 + (0.07)^2 + (0.15)^2 + (0.22)^2 + (0.24)^2 + (0.31)^2$

$\Sigma y_i^2 = 0.0009 + 0.0049 + 0.0225 + 0.0484 + 0.0576 + 0.0961$

$\Sigma y_i^2 = 0.2304$

$n(\Sigma y_i^2) = 6 \times 0.2304$

$\underline{n(\Sigma y_i^2) = 1.3824}$

$(\Sigma y_i)^2 = (0.03 + 0.07 + 0.15 + 0.22 + 0.24 + 0.31)^2$

$(\Sigma y_i)^2 = (125.5)^2$

$\underline{(\Sigma y_i)^2 = 1.04}$

$[n(\Sigma y_i^2) - (\Sigma yi)^2] = 1.3824 – 1.04$

$\underline{[n(\Sigma y_i^2) - (\Sigma yi)^2] = 0.3424}$

Step (b): $r = \dfrac{7.6 - (33.5x1.02)/6}{\sqrt{[251.25 - (1{,}122.25/6)][0.2304 - (1.04/6)]}}$

$\therefore\ r = \dfrac{7.6 - 5.695}{\sqrt{(251.25 - 187.04)(0.2304 - 0.1733)}}$

$$\therefore\ r = \frac{1.905}{\sqrt{64.21x0.0571}}$$

$$\therefore r = \frac{1.905}{\sqrt{3.666}}$$

$$\therefore r = \frac{1.905}{1.915}$$

$\therefore r = 0.9947$

$\underline{\therefore r^2 = 0.989}$

Note: r^2 is a unitless quantity. In this case r^2 is still very close to 1, indicating an excellent correlation coefficient and hence a close fit to a perfect straight line.

2.20) A flame phototometer is used to determine the Mg^{2+} concentration of a water sample. The instrument was calibrated via a standard additions method, the additions are listed below. Assuming that no interferences are present, determine the Mg^{2+} content.

Standard addition concentration (mg dm^{-3})	Instrument reading (arbitrary units)
0 (Blank)	15.6
2.5	22.1
5	35.1
10	48.1
15	63.7
20	79.3

Determine $\bar{y} = m\bar{x} + c$ equation:

$\bar{x}$ = (0 + 2.5 + 5 + 10 + 15 + 20) / 6 = 8.75

$\bar{y}$ = (15.6 + 22.1 + 35.1 + 48.1 + 63.7 + 79.3) / 6 = 43.98

c = 15.6

$$\frac{\bar{y}-c}{x} = \frac{43.98}{8.75} = 5.026 \text{ arbitrary units per mg dm}^{-3}$$

It follows that the sample (blank) contains (15.6 / 5.026) mg dm^{-3} Mg^{2+}

∴ Mg^{2+} content = 3.1mg dm^{-3}

2.21) Calculate the Pearson correlation coefficient for the data of Problem 2.20.

Standard addition concentration (mg dm^{-3})	Instrument reading (arbitrary units)
0 (Blank)	15.6
2.5	22.1
5	35.1
10	48.1
15	63.7
20	79.3

$$r = \frac{\sum x_i y_i - (\sum x_i)(\sum y_i)/n}{\sqrt{[\sum x_i^2 - (\sum x_i)^2/n][\sum y_i^2 - (\sum y_i)^2/n]}}$$

Method:

(a): Assign *x* and *y* values and calculate each term for Eqn 2.17

(b): Substitute values and calculate *r* and r^2.

Step (a): n = 6 for 6 data points.

Calculate n $\Sigma x_i y_i$

X	Y	X_iY_i
0	15.6	0
2.5	22.1	55.25
5	35.1	175.5
10	48.1	481
15	63.7	955.5
20	79.3	1586
		$\Sigma x_i y_i$ = 3253.25

∴ n $\Sigma x_i y_i$ = 6 x 3253.25

∴ n $\Sigma x_i y_i$ = 19519.5

Calculate Σx_i Σy_i:

Σx_i = 0 + 2.5 + 5.0 + 10 + 15 + 20.0

Σx_i = 52.5

Σy_i = 15.6 + 22.1 + 35.1 + 48.1 + 63.7 + 79.3

Σy_i = 263.9

$\Sigma x_i \Sigma y_i$ = 52.5 x 263.9

$\Sigma x_i \Sigma y_i$ = 13854.75

Calculate $[n(\Sigma x_i^2) - (\Sigma x_i)^2]$

$\Sigma x_i^2 = (0)^2\ (2.5)^2 + (5.0)^2 + (10.0)^2 + (15)^{2\,+}\ (20.0)^2$

Σx_i^2 = 0 + 6.25 + 25 + 100 + 225 + 400

Σx_i^2 = 756.25

$(\Sigma x_i)^2 = (52.5)^2$

$(\Sigma x_i)^2$ = 2756.25

$[n(\Sigma x_i^2) - (\Sigma x_i)^2]$ = 6(756.25) – 2756.25

$[n(\Sigma x_i^2) - (\Sigma x_i)^2]$ = 4537.5 – 2756.25

$[n(\Sigma x_i^2) - (\Sigma x_i)^2]$ = 1781.25

Calculate $[n(\Sigma y_i^2) - (\Sigma y_i)^2]$:

$\Sigma y_i^2 = (15.6)^2 + (22.1)^2 + (35.1)^2 + (48.1)^2 + (63.7)^2 + (79.3)^2$

Σy_i^2 = 243.36 + 488.41 + 1232.01 + 2313.61 + 4057.69 + 6288.49

Σy_i^2 = 14623.57

$n(\Sigma y_i^2)$ = 6 x 14623.57

$n(\Sigma y_i^2)$ = 87741.42

$(\Sigma y_i)^2 = (15.6 + 22.1 + 35.1 + 48.1 + 63.7 + 79.3)^2$

$(\Sigma y_i)^2 = (263.9)^2$

<u>$(\Sigma y_i)^2 = 69643.21$</u>

$[n(\Sigma y_i^2) - (\Sigma yi)^2] = 87741.42 - 69643.21$

<u>$[n(\Sigma y_i^2) - (\Sigma yi)^2] = 18098.21$</u>

Step (b): $r = \dfrac{3253.25 - (52.5x263.9)/6}{\sqrt{[756.25 - (2756.25/6)][14623.57 - (69643.21/6)]}}$

$$\therefore r = \frac{3253.25 - 2309.125}{\sqrt{(756.25 - 459.375)(14623.57 - 11607.202)}}$$

$$\therefore r = \frac{944.125}{\sqrt{296.875x3016.368}}$$

$$\therefore r = \frac{944.125}{\sqrt{895484.25}}$$

$$\therefore r = \frac{944.125}{946.3}$$

$\therefore r = 0.998$

<u>$\therefore r^2 = 0.995$</u>

Note: r^2 is a unitless quantity. In this case r^2 is still very close to 1, indicating an excellent correlation coefficient and hence a close fit to a perfect straight line.