

MODULE 4

Data Management: Introduction to Statistics

Learning Activity 1

Direction. Answer the given item.

Mr. Jose Mari Khan, a candidate for mayor in San Sebastian, wants to find out if there is a need to intensify his campaign efforts against his opponents. He requested the services of a group of students to interview 1,000 of the 3,000 registered voters in San Isidro. The survey results showed that 75% of the 1,000 voters in the sample will vote for him as vice-mayor.

a. Identify the population

- number of registered voters in San Isidro in population = 3,000 voters
- number of registered voters in San Isidro in sample = 1,000 voters

b. Identify the variable of interest

- before the survey, after the survey

c. Identify the parameter or statistic

- $p = \frac{\text{number of voters they interview}}{\text{number of voters} \in \text{the population}} = \frac{1,000}{3,000} = 0.33$
- $p = \frac{\text{number of voters that vote him for vice mayor}}{\text{number of votes} \in \text{the sample}} = \frac{750}{1,000} = 75\%$

Learning Activity 2

Direction. For each of the following statements below, decide whether it belongs to the field of descriptive statistics or inferential statistics.

Learning Activity 3

1. A badminton player wants to know his average score for the past 10 games. **Q** If the variable is qualitative, write **Q** if it is discrete and **C** if continuous.

2. A car manufacturer wishes to estimate the average lifetime of batteries by testing a sample of 50 batteries.

DESCRIPTIVE STATISTICS

- | | |
|---|---|
| 1. Speed of cars C | 8. Weight of children C |
| 3. A politician wants to determine the total number of votes his rival obtained in the past election based on his | 9. Height of the building C |
| 2. Brand of watches Q | 10. ID number Q |
| 3. Height of children Q | 11. Place of residence Q |
| 4. Student number Q | 12. Rank of teachers Q |
| 5. IQ score C | 13. Time required to take the attainment examination Q |
| 6. Educational Q | 14. Political school Affiliation Q |
| 7. Number of years in D | |

Learning Activity 4

Direction. State the level of measurement used to measure the following variables:

- Postal zip code
- Performance rating of an employee as excellent, very good, good, fair, and bad
- Student number
- Ranking of a student in class
- Annual salary of employee

Learning Activity 5

Direction. Answer the following items.

1. Classify each of the following as a survey, an experiment, or an observation.

- a. A local TV network asked voters to indicate whom they voted as they exited the polling booth. **Survey** _____
- b. A private hospital divides terminally ill patients into two groups, with one group receiving medication *A* and the other group receiving medication *B*. After a month, they measured each subject's improvement. **Observation** _____
- c. A researcher investigates the level of pollution in key points in Metro Manila by setting up pollution measuring devices at selected intersections. **Observation** _____
- d. The school administration asked students whether they are willing to have an increase in laboratory fees if there is an upgrade of computers. **Survey** _____

2. What method of data collection is most appropriate for the following cases?

- a. Studying two groups of patients and determining if exercise lowers the blood pressure. **Experiment** _____
- b. The Department of Health monitors and evaluates the benefits of the family planning methods given to a certain community. **Survey** _____
- c. A group of medical intern students studies the effects of laughter to patients in a hospital **Experiment** _____
- d. A nongovernment organization compares the **Survey or Use Document**

household expenditures of two districts in Isabela.

B. Identification

1. The average weekly allowance of students last year at a private high school was Php 600.00 per week, based on an enrollment of 1,080 students. The third year students who did not have this information interviewed 50 students and found their average weekly allowance last year to be Php 550.00. Identify the following:
 - a. Population – 1,080 students
 - b. Sample – 50 students
 - c. Variable of interest – before interview and after interview
 - d. Parameter – $p = \frac{\text{last year weekly allowance}}{\text{number of enrolled students} \in \text{population}} = \frac{600}{1080} = 0.53$
 - e. Sample Statistic – $p = \frac{\text{third year students weekly allowance last year}}{\text{number of enrolled students} \in \text{sample}} = \frac{550}{55} = 11$
2. Observe the use of the number seven in the following statements. Classify each statement according to the level of measurement used to get the value 7.
 - a. Mark is in the 7th grade - **Ordinal or Rank Data**
 - b. Mark measured the temperature of the object as 7°C.- **Interval Data**
 - c. Mark has a score of 7 in the Math quiz. - **Ratio Data**
 - d. Mark's basketball shirt number is 7. - **Nominal Data**
 - e. Mark has 7 cousins. – **Nominal Data**
3. What method of data collection is most appropriate for the following cases?
 - a. A group of Anthropology students studies the culture and norms of two ethnic groups. - **Observation**
 - b. A social welfare organization gathers information on hospital patients with mental disorder.- **experiment**
 - c. A construction contractor deciding on future house construction wishes to determine the type of house demand among families.- **Experiment**
 - d. A milk manufacturer wishes to determine the preference in drink flavor among children in planning for new products.- **Survey**

- e. A car manufacturer studies the preference of cars for the next production.- **observation**
4. Indicate the type of chart you would choose to present the information given in each of the following cases.
- a. Percentage distribution of monthly expenditures of a Filipino family for clothing, footwear, house maintenance, and food – **Pie chart**
 - b. The number of hardware stores in Isabela for the year 2015 to 2020- **column chart**
 - c. Log production for 1990 to 2005- **line chart**
 - d. Location Number of Savemore supermarkets in Region 2- **horizontal bar graph**
 - e. Distribution of employees by civil status- **Risto graph**

MODULE 5

Data Management: Measures of Central Tendency, Dispersion and Position

Learning Activity 1

Direction. Tell whether the following statements describe the Mean, Median, or mode

- | | |
|---|---|
| <ol style="list-style-type: none">1. The most preferred descriptive measure in a skewed distribution2. Will have the largest value in a negatively skewed distribution3. Will have the largest value in a positively skewed distribution.4. The point above and below where half of the distribution of the data falls5. Will have the same value in a bimodal distribution6. The “center of gravity” of a distribution7. Is equivalent to the 50th percentile of a distribution?8. The most popular score in a distribution9. Influenced by the specific value of every observation10. Is most appropriate to use when extreme scores are given. | <ol style="list-style-type: none">1. mean2. mode3. mean4. median5. mean and median6. mean7. median8. mode9. mean10. median |
|---|---|

5.6 Assessment Task

1. For each of the following data sets determine the mean:

- a. 72, 14, 8, 11, 57, 54, 31, 11, 67, 11, 19, 3, 66

$$\bar{x} = \frac{\sum x}{n} = \frac{424}{13} = 32.6153... \approx 32.62$$

- b. 63, 9, 87, 16, 2, 96, 13, 67, 34

$$\bar{x} = \frac{\sum x}{n} = \frac{387}{9} = 43$$

- c. 1, 6, 8, 2, 7, 2, 9, 4, 8, 9, 8, 6

$$\bar{x} = \frac{\sum x}{n} = \frac{70}{12} = 5.8333... \approx 5.83$$

- d. 12, 0, 4, 5, 8, 3, 6, 35, 47

$$\bar{x} = \frac{\sum x}{n} = \frac{120}{9} = 13.3333... \approx 13.33$$

- e. 2, 7, 17, 33, 67, 73, 88, 33, 92, 57, 33

$$\bar{x} = \frac{\sum x}{n} = \frac{502}{11} = 45.6363... \approx 45.64$$

2. For each of the following data sets determine the median:

- a. 8, 10, 20, 40, 45, 48, **57**, 63, 66, 66, 67, 78, 96

$$\frac{n+1}{2} = \frac{13+1}{2} = \frac{14}{2} = 7 \rightarrow 57$$

- b. 3, 8, 10, 10, 11, 23, 28, 45, 60, 97

$$\frac{n}{2} = \frac{10}{2} = 5 \rightarrow 11$$

$$\bar{x} = \frac{11+23}{2} = \frac{34}{2} = 17$$

- c. 0, 1, 1, 1, 3, **4**, 6, 6, 7, 8, 8, 9

$$\frac{n}{2} = \frac{12}{2} = 6 \rightarrow 4$$

$$\dot{x} = \frac{4+6}{2} = \frac{10}{2} = 5$$

d. 2, 9, 16, 24, **34**, 40, 63, 67, 86, 97

$$\frac{n}{2} = \frac{10}{2} = 5 \rightarrow 34$$

$$\dot{x} = \frac{34+40}{2} = \frac{74}{2} = 37$$

e. 4, 8, 11, 15, 42, **42**, 54, 64, 77, 88, 99

$$\frac{n+1}{2} = \frac{11+1}{2} = \frac{12}{2} = 6 \rightarrow 42$$

3. For each of the following data sets determine the mode:

a. 98, 37,5, **33**, 96, 67, 43, **33**, 91, **33**,32,8,11
 $\rightarrow 33$

b. **104**, 2, 51, 31, 8, 101, **104**, 18, 47
 $\rightarrow 104$

c. 3, 4, **5**, 3, 9, 8, **5**, 7, 2, **5**, 1
 $\rightarrow 5$

d. 19, 1, 9, 6, 4, **2**, 13, 15, 24, **2**
 $\rightarrow 2$

e. 8, 9, 39, **44**, 55, 90, 19, **44**, 28, 69, **44**
 $\rightarrow 44$

4. A quiz on the classification of research by general methodology was administered to a group of 34 students at the College of Arts and Sciences. The scores are reported below:

Male	Female
8 10 20 14 13 9	~~10~~ ~~10~~ 13 ~~10~~
17 17 12 14 14	~~16~~ 8 5 18 ~~17~~
12 10 9 18	6 16 10

- a. Consider all the members of the group and compute the mean, median and mode.

- Mean $\mu = \frac{\sum x}{N} = \frac{441}{34} = 12.9705... \approx 12.97$
- Median
6, 8, 8, 8, 9, 9, 10, 10, 10, 10, 10, 10, 10, 12, 12, 13, 13, **13**, 14, 14, 14, 14, 14, 15, 15, 16, 16, 17, 17, 17, 17, 18, 18, 20
 $\frac{n}{2} = \frac{34}{2} = 17 \rightarrow 13$
- Mode
6, 8, 8, 8, 9, 9, **10, 10, 10, 10, 10, 10**, 12, 12, 13, 13, 13, **14, 14, 14, 14, 14**, 15, 15, 16, 16, 17, 17, 17, 17, 18, 18, 20
 $\rightarrow 10 \wedge 14$

- b. Calculate the mean, median and mode for male and female students.

1. Male

- Mean
 $\bar{x} = \frac{\sum x}{n} = \frac{197}{15} = 13.1333... \approx 13.13$
- Median
8, 9, 9, 10, 10, 12, 12, **13**, 14, 14, 14, 17, 17, 18, 20
 $\frac{n+1}{2} = \frac{15+1}{2} = \frac{16}{2} = 8 \rightarrow 13$

- Mode

8, 9, 9, 10, 10, 12, 12, 13, **14, 14, 14**, 17, 17, 18, 20
 → 14

2. Female

- Mean

$$\bar{x} = \frac{\sum x}{n} = \frac{244}{19} = 12.8421 \dots \approx 12.84$$

- Median

6, 8, 8, 10, 10, 10, 10, 13, 13, **14**, 14, 14, 15, 15, 16, 16, 17, 17, 18

$$\frac{n+1}{2} = \frac{19+1}{2} = \frac{20}{2} = 10 \rightarrow 14$$

- Mode

6, 8, 8, **10, 10, 10, 10**, 13, 13, 14, 14, 14, 15, 15, 16, 16, 17, 17, 18
 → 10

- c. Compare the mean and median within each group. Which has the higher value? Why?

For items 5-6, find the mean, median, mode, range, variance, standard deviation, and the coefficient of variation.

5. Cost of notebooks for these sample prizes:

7.95	9.98	5.58	4.99	10.75	6.25	7.63	8.50	8.88
4.99	5.58	6.35	7.63	7.95	8.50	8.88	9.98	10.75

- Mean

$$\bar{x} = \frac{\sum x}{n} = \frac{70.51}{9} = 7.8344... \approx 7.83$$

- Median

4.99, 5.58, 6.35, 7.63, **7.95**, 8.50, 8.88, 9.98, 10.75

$$\frac{n+1}{2} = \frac{9+1}{2} = \frac{10}{2} = 5 \rightarrow 7.95$$

- Mode (none)

- Range

$$10.75 - 4.99 = 5.76$$

- Variance

x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
4.99	7.83	-2.84	8.07
5.58	7.83	-2.25	5.06
6.25	7.83	-1.58	2.50
7.63	7.83	-0.2	0.04
7.95	7.83	0.12	0.01
8.50	7.83	0.67	0.49
8.88	7.83	1.05	1.10
9.98	7.83	2.15	4.62
10.75	7.83	2.92	8.53
$\Sigma x = 70.51$			$\Sigma (x - \bar{x})^2 = 30.42$

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{n-1} = \frac{30.42}{9-1} = \frac{30.42}{8} = 3.8025 \approx 3.80$$

- Standard Deviation

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} = \sqrt{\frac{30.42}{9-1}} = \sqrt{\frac{30.42}{8}} = \sqrt{3.80} = 1.9493... \approx$$

1.95

- Coefficient of Variation

$$cv = \frac{s}{\bar{x}} \times 100\% = \frac{1.95}{7.38} \times 100\% = 0.2642... \approx 0.26$$

6. Age of college students in ISU:

Age(x)	No. of students (f)	f(x)
16	2	32
17	10	170
18	8	144
19	5	95
	N=25	$\sum fx = 441$

- Mean

$$\mu = \frac{\sum fx}{N} = \frac{441}{25} = 17.64$$

- Median

Age(x)	No. of students (f)	cf	
16	2	2	1-2
17	10	12	3-12
18	8	20	13-20
19	5	25	21-25

$$\frac{N+1}{2} = \frac{25+1}{2} = \frac{26}{2} = 13 \rightarrow 18$$

- Mode

Age(x)	No. of students (f)
16	2
17	10
18	8
19	5

→ 17, because it has the most frequencies

- Range
 $19 - 16 = 3$

- Variance

x	μ	$x - \mu$	$(x - \mu)^2$
16	17.5	-1.5	2.25
17	17.5	-0.5	0.25
18	17.5	0.5	0.25
19	17.5	1.5	2.25
$\Sigma = 70$			$\Sigma(x - \mu)^2 = 5$

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{N} = \frac{5}{4} = 1.25$$

- Standard Deviation

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} = \sqrt{\frac{5}{4}} = \sqrt{1.25} = 1.1180 \dots \approx \mathbf{1.12}$$

- Coefficient of Variation

$$cv = \frac{\sigma}{\bar{x}} \times 100\% = \frac{1.11}{17.5} \times 100\% = 0.0634 \dots \approx 0.06$$

7. Louie's test scores for two semesters of mathematics are listed below. The percentage of each semester's grade represented by each score is also given

1 st Semester	2 nd Semester	% of Grade
78	87	15
68	66	15
84	81	15
86	89	15
90	88	40

- a. Compute the weighted arithmetic mean and standard deviation for each semester.

1. Weighted Arithmetic Mean

- 1st Semester

1st Semester (x)	% of Grade (f)	f(x)
78	15	1,170
68	15	1,020
84	15	1,260
86	15	1,290
90	40	3,600
	n=100	$\Sigma fx=8,340$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{8,340}{100} = 83.4$$

- 2nd Semester

2nd Semester (x)	% of Grade (f)	f(x)
87	15	1,305
66	15	990
81	15	1,215
89	15	1,335
88	40	3,520
	n=100	$\Sigma fx=8,365$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{8,365}{100} = 83.65$$

2. Standard Deviation

- 1st Semester

x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
78	81.2	-3.2	10.24
68	81.2	-13.2	174.24
84	81.2	2.8	7.84
86	81.2	4.8	23.04
90	81.2	8.8	77.44
$\Sigma = 406$			$\Sigma (x - \bar{x})^2 = 292.8$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}} = \sqrt{\frac{292.8}{5-1}} = \sqrt{\frac{292.8}{4}} = \sqrt{73.2} = 8.5556... \\ \approx \mathbf{8.56}$$

- 2nd Semester

x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
87	82.2	4.8	23.04
66	82.2	-16.2	262.44
81	82.2	-1.2	1.44
89	82.2	6.8	46.24
88	82.2	5.8	33.64
$\Sigma = 411$			$\Sigma(x - \bar{x})^2 = 366.8$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}} = \sqrt{\frac{366.8}{5-1}} = \sqrt{\frac{366.8}{4}} = \sqrt{91.7} = 9.5760... \\ \approx \mathbf{9.58}$$

b. Is Louie improving? Explain your answer.

8. The following set of data represents a simple random sample of IQ scores of 32 students at ISU.

137	141	128	135	159	122	140	118
126	133	111	125	127	116	138	203
120	121	131	122	125	126	118	119
117	139	133	124	168	135	126	131

a. Calculate the range of the test scores.

- Range
203 - 111 = 92

b. Calculate the sample variance, and standard deviation.

x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
139	131.75	7.25	53.56
126	131.75	-5.75	33.06
120	131.75	-11.75	138.06
117	131.75	-14.75	217.56
141	131.75	9.25	85.56
133	131.75	1.25	1.56
121	131.75	-10.75	115.56
139	131.75	7.25	52.56
128	131.75	-3.75	14.06
111	131.75	-20.75	430.56
131	131.75	-0.75	0.56
133	131.75	1.25	1.56
135	131.75	3.25	10.56
125	131.75	-6.75	45.56
122	131.75	-9.75	95.06
124	131.75	-7.75	60.06
159	131.75	27.250	742.56
127	131.75	-4.75	22.56
125	131.75	-6.75	45.56
168	131.75	36.25	1,314.06
122	131.75	-9.75	95.06
116	131.75	-15.75	248.06
126	131.75	-5.75	33.06
135	131.75	3.25	10.56
140	131.75	8.25	68.06
138	131.75	6.25	39.06
118	131.75	-13.75	189.06
126	131.75	-5.75	33.06
118	131.75	-13.75	189.06
203	131.75	71.25	5,076.56
119	131.75	-12.75	162.56
131	131.75	-0.75	0.56
$\Sigma = 4216$			$\Sigma(x - \bar{x})^2 = 9624.92$

- Variance

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n-1} = \frac{9624.92}{32-1} = \frac{9624.92}{31} = 310.4812... \approx 310.48$$

- Standard Deviation

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}} = \sqrt{\frac{9624.92}{32-1}} = \sqrt{\frac{9624.92}{31}} = \sqrt{310.48} = 17.6204... \approx \mathbf{17.62}$$

9. Using the following, calculate the z -score that corresponds to the raw score indicated.

$\bar{x}$	<i>s</i>	<i>x</i>	<i>z-score</i>
97	9.23	100	0.33
46	8.0	38	-1
8	0.52	9	1.92
22	4.69	24	0.43
31	7.15	24	-0.98
54	1.50	39	-10
100	6.50	110	1.54
75	3.75	72	-0.8

10. Using the following data, estimate the raw score that corresponds to the z -score indicated.

$\bar{x}$	<i>s</i>	<i>z-score</i>	<i>x</i>
28	5.2	-1.62	19.78
69	2.35	+2.58	75.06
7	0.86	+1.03	7.89

41	4.73	-2.37	29.79
72	1.05	+0.40	72.42
85	3.21	-3.20	74.73
750	9.61	-0.26	147.50
36	0.90	+3.50	39.15

MODULE 6

Data Management: Probabilities and Normal Distribution

Learning Activity 1

Direction. Find the following probabilities.

- | | |
|---------------------------------|---------------|
| 1. $P(z \leq -1.73)$ | <u>0.0418</u> |
| 2. $P(z \geq -0.67)$ | <u>0.7486</u> |
| 3. $P(-1.73 \leq z \leq -0.67)$ | <u>0.2096</u> |

Learning Activity 2

Direction. Solve the following probabilities.

The heights of 1000 students are normally distributed with a mean of 174.5 centimeters and a standard deviation of 6.9 centimeters. Assuming that the heights are recorded to the nearest half centimeters, how many of these students would you expect to have heights

- a. Less than 160.0 centimeters? **17.9**
- b. Between 171.5 and 182.0 centimeters inclusive? **0.5285**
- c. Equal to 175.0 centimeters? **8 students**
- d. Greater than or equal to 188.0 centimeters? **25 students**

6.6 Assessment Task

Direction. Solve the following problems.

1. Find the value of z if the area under a standard normal curve
 - a. to the right of z is 0.3622; - **$z = -0.35$**
 - b. to the left of z is 0.1131; - **$z = -1.21$**
 - c. between 0 and z , with $z > 0$, is 0.4838; **$z = 2.14$**
 - d. between $-z$ and z , with $z > 0$, is 0.9500. **$z = 1.96$**
2. Given a normal distribution with $\mu = 30$ and $\sigma = 6$, find
 - a. the normal curve area to the right of $x = 17$; **0.9850**
 - b. the normal curve area to the left of $x = 22$; **0.0918**
 - c. the normal curve area between $x = 32$ and $x = 41$; **0.3371**

- d. the value of x that has 80% of the normal curve area to the left; **35.04**
 - e. the two values of x that contain the middle 75% of the normal curve area. **30 and 34.02**
3. A research scientist reports that mice will live an average of 40 months when their diets are sharply restricted and then enriched with vitamins and proteins. Assuming that the lifetimes of such mice are normally distributed with a standard deviation of 6.3 months, find the probability that a given mouse will live
- a. more than 32 months; **0.8980**
 - b. less than 28 months; **0.0287**
 - c. between 37 and 49 months. **0.6080**
4. A soft-drink machine is regulated so that it discharges an average of 200 milliliters per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 milliliters,
- a. what fraction of the cups will contain more than 224 milliliters? **0.0548**
 - b. what is the probability that a cup contains between 191 and 209 milliliters? **0.4514**
 - c. how many cups will probably overflow if 230- milliliter cups are used for the next 1000 drinks? **10.8 = 11 cups will over flow**
 - d. below what value do we get the smallest 25% of the drinks? **189.95**

MODULE 7

Data Management: Regression and Correlation

Learning Activity 1

Direction. Write T if the statement is true and F if it is not.

1. A high positive correlation indicates that variable X causes a predictable change in variable Y . **T**
2. The Pearson r can only be used with data measured in ordinal scale. **F**
3. Two variables with a correlation of -0.75 have a weaker association than two variables with correlation of 0.56 . **T**
4. An inverse correlation indicates that the two variables tend to change either in the direct or in the opposite directions. **T**
5. If two variables have a low correlation, such as $r = +0.04$, the two variables cannot have a strong relationship to one another. **F**
6. Direct relationships in a set of data are always stronger than indirect or inverse relationship in a data set. **T**
7. The relative strength of the relationships in different samples is most accurately assessed by computing the r^2 values for the samples. **F**
8. A meaningful correlation coefficient can only be computed when data are measured in interval or ratio scales. **F**
9. The sign of the correlation coefficient can be used to determine the strength of the relationships between two variables. **T**
10. A correlation coefficient of -0.60 is approximately twice as strong as correlation coefficient of $+0.30$. **F**

Learning Activity 2

Direction. Solve the following.

1. The scores of five students in mathematics and chemistry classes are:

Mathematics 6 4 8 5 3. 5

Chemistry 6. 5 4. 5 7 5 4

Calculate the linear correlation coefficient.

x	y	x ²	y ²	xy
6	6.5	36	42.25	39
4	4.5	16	20.25	18
8	7	64	49	56
5	5	25	25	25
3.5	4	12.25	16	14

$$\Sigma y^2 = 152.5 \quad \Sigma xy = 26.5$$

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2]} \sqrt{[n\Sigma y^2 - (\Sigma y)^2]}}$$

$$\Sigma y = 27$$

$$r = \frac{5(152) - (26.5)(27)}{\sqrt{[5(153.25) - (26.5)^2]} \sqrt{[5(152.5) - (27)^2]}}$$

$$r = + 0.9610 \quad \Sigma x^2 = 153.25 \quad \Sigma x = 26.5$$

- From the result, we can say that Mathematics and Chemistry have a strong positive linear correlation

Learning Activity 3

Direction.

Solve the following.

An exercise instructor remembers that the data given in the following table, which shows the recommended maximum exercise heart rates for individuals of

	x	y	x ²	y ²	xy
37	20	1.75	400	3.0625	64.75
35	40	3	1600	9	105
24	60	2.35	3600	5.5225	56.4
15	20	2	400	4	80
28	40	2.50	1600	6.25	70
	60		3600		160
	Σx= 120	Σy= 459	Σx ² = 5600		Σxy= 17680
35		1.63	1225	2.6569	57.05
32		2.80	1024	7.84	89.6
38		2.52	1444	6.3504	95.76
30		2.95	900	8.7025	88.5
27		1.95	459	3.8025	52.65
29		-0.85	841	4.5369	61.77
30		2.42	900	5.8564	72.6
24	40	2.53	576	6.4009	60.72
39	153	2.89	1521	8.1225	111.15
35		2.50	1225	6.25	87.5

$$b = 153 - (-0.85)(40) = 187$$

$$y = 142$$

$$\Sigma x = 458 \quad \Sigma y = 35.88 \quad \Sigma x^2 = 14564 \quad \Sigma y^2 = 88.0565 \quad \Sigma xy = 1103.45$$

7.6AssessmentTask

a.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]} \sqrt{[n\sum y^2 - (\sum y)^2]}}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]} \sqrt{[n\sum y^2 - (\sum y)^2]}}$$

$$r = \frac{15(1103.45) - (458)(35.88)}{\sqrt{[15(14564) - (458)^2]} \sqrt{[15(88.0565) - (35.88)^2]}}$$

$$r = 0.2200$$

- From the result we can say that Community service(hours) and GWA has a positive correlation.

b.

$$a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

$$b = \bar{y} - a\bar{x}$$

$$\hat{y} = ax + b$$

$$a = \frac{15(1103.45) - (458)(35.88)}{15(14564) - (458)^2}$$

$$a = 0.0136511$$

$$\bar{x} = \sum x / n = 458/15$$

$$\bar{x} = 30.533$$

$$\bar{y} = \sum y / n = 35.88/15$$

$$\bar{y} = 2.392$$

$$b = \bar{y} - a\bar{x}$$

$$b = 2.392 - (0.0136511)(30.533)$$

$$b = 1.9751$$

$$\hat{y} = ax + b$$

$$\hat{y} = 0.014x + 2$$
