

DEPARTMENT OF EDUCATION

GRADE 12 MATHEMATICS B

12.3: PROBABILITY AND STATISTICS



FODE DISTANCE LEARNING



PUBLISHED BY FLEXIBLE OPEN AND DISTANCE EDUCATION
FOR THE DEPARTMENT OF EDUCATION
PAPUA NEW GUINEA



GRADE 12

MATHEMATICS B

UNIT MODULE 3

PROBABILITY AND STATISTICS

TOPIC 1:	Fundamental Principles of Counting
TOPIC 2:	Probability
TOPIC 3:	Statistics



Acknowledgements

We acknowledge the contribution of all Secondary and Upper Primary teachers who in one way or another helped to develop this Course.

Special thanks are given to the staff of the Mathematics Department- FODE who played active role in coordinating writing workshops, outsourcing of lesson writing and editing processes involving selected teachers in NCD.

We also acknowledge the professional guidance and services provided throughout the processes of writing by the members of:

Mathematics Subject Review Committee-FODE
Academic Advisory Committee-FODE
Mathematics Department- CDAD

This book was developed with the invaluable support and co-funding of the GO-PNG and World Bank.

MR. DEMAS TONGOGO
Principal-FODE



Flexible Open and Distance Education
Papua New Guinea

Published in 2016

@ Copyright 2016, Department of Education
Papua New Guinea

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means electronic, mechanical, photocopying, recording or any other form of reproduction by any process is allowed without the prior permission of the publisher.

ISBN: 978- 9980-89-413-7

National Library Services of Papua New Guinea

Written by: Prof. Grace Tiqui

Compiled and finalised by: Mathematics Department-FODE

Printed by the Flexible, Open and Distance Education



CONTENTS		Page
Title.....		1
Acknowledgement and Copy Right.....		2
Contents.....		3
Secretary's Message.....		4
Unit Introduction.....		5
12.3.1: FUNDAMENTAL PRINCIPLES IN COUNTING.....		6
☐ 12.4.1.1: Counting Principles.....		6
☐ 12.4.1.2: Permutation.....		16
☐ 12.4.1.3: Combination.....		26
Summative Task 1.....		33
12.3.2: PROBABILITY.....		34
☐ 12.3.2.1: Fundamentals of Probability.....		34
☐ 12.3.2.2: Probabilities of Events.....		38
☐ 12.3.2.3: Independent and dependent Events.....		45
☐ 12.3.2.3: Conditional Probability.....		53
Summative Task 12.3.2.....		62
12.3.3: STATISTICS.....		65
☐ 12.3.3.1: The Scatter Diagram.....		65
☐ 12.3.3.2: Analysis Using Correlation.....		71
☐ 12.3.3.3: Regression.....		82
☐ 12.3.3.3.1: Line of best Fit.....		83
☐ 12.3.3.3.2: Interpolation and Extrapolation.....		88
Summative Task 12.4.3.....		92
SUMMARY.....		94
ANSWERS TO LEARNING ACTIVITIES 12.4.1 to 12.4.4.....		97
REFERENCES.....		105



SECRETARY'S MESSAGE

Achieving a better future by individual students, their families, communities or the nation as a whole depends on the curriculum and the way it is delivered.

This course is part and parcel of the NDOE new reformed curriculum. Its learning outcomes are student centred and written in terms that allow them to be demonstrated, assessed and measured.

It maintains the rationale, goals, aims and principles of the National Curriculum and identifies the knowledge, skills, attitudes and values that students should achieve.

This is a provision of Flexible, Open and Distance Education as an alternative pathway of formal education.

The Course promotes Papua New Guinea values and beliefs which are found in its constitution, Government policies and reports. It is developed in line with the National Education Plan (2005 – 2014) and addresses an increase in the number of school leavers which has been coupled with a limited access to secondary and higher educational institutions.

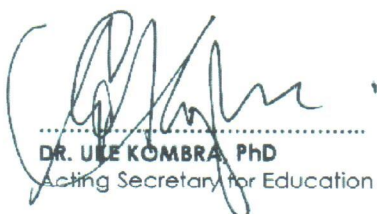
Flexible, Open and Distance Education is guided by the Department of Education's Mission which is fivefold:

- to facilitate and promote the integral development of every individual
- to develop and encourage an education system which satisfies the requirements of Papua New Guinea and its people
- to establish, preserve, and improve standards of education throughout Papua New Guinea
- to make the benefits of such education available as widely as possible to all of the people
- to make education accessible to the physically, mentally and socially handicapped as well as to those who are educationally disadvantaged

The College is enhanced to provide alternative and comparable pathways for students and adults to complete their education, through one system, many pathways and same learning outcomes.

It is our vision that Papua New Guineans harness all appropriate and affordable technologies to pursue this program.

I commend all those teachers, curriculum writers and instructional designers who have contributed so much in developing this course.



DR. ULE KOMBRA, PhD
Acting Secretary for Education



UNIT 3: PROBABILITY AND STATISTICS

Introduction

*"There are known knowns. These are things we know that we know.
There are known unknowns. That is to say, there are things that we know we don't know.
But there are also unknown unknowns.
These are things we don't know we do not know,"*
- Donald Rumsfeld

The above quotation makes us think and reflect of the known and unknowns we deal in our everyday life. Such dealing involves two important branches of mathematics, Probability and Statistics.

Probability deals with certainty or uncertainty that an event will happen while statistics deals with data collection, analysis and interpretation. These two are very important in our daily lives as they guide us in making informed and sound decisions.

This module will consider the local environment as its context for most application problems.



LEARNING OUTCOMES

On successful completion of this module, you will be able to:

- calculate simple probability of events
- classify and calculate events as independent and dependent
- investigate and calculate events as mutually exclusive and non-mutually exclusive
- plot scatter diagrams
- discuss and interpret linear and non-linear relationships
- calculate the regression and correlation coefficients
- write equation of the regression line
- interpret regression and correlation coefficient in the context of the problem



TIME FRAME

This unit should be completed within 6 weeks.

If you set an average of 3 hours per day, you should be able to complete the unit comfortably by the end of the assigned week.

Try to do all the learning activities and compare your answers with the ones provided at the end of the unit. If you do not get a particular exercise right in the first attempt, you should not get discouraged but instead, go back and attempt it again. If you still do not get it right after several attempts then you should seek help from your friend or even your tutor. Do not pass any question without solving it first.

12.3.1: Fundamental Principles in Counting

12.3.1.1: Counting Principles and Techniques

The first Mathematical skill probably taught to you by your parents when you started talking could be counting. Yes, learning to count as 1, 2, 3 and so on may be too familiar. But what do we really mean by counting?

Counting is more than saying the numbers one after the other; it is the fundamental in solving many problems in mathematics. Specifically Probability involves counting the number of ways an event can occur or, equivalently, counting the number of possible ways there are to perform a specific task.

The idea of counting may sound easy but sometimes the number of outcomes of your sample space is too large to list and count. Therefore, principles of counting and techniques will help us facilitate.

Counting the number of cars in a parking lot may be easy. How about counting the possible number of car plates for taxis in a certain area?

We can possibly list a passcode using 3 letters and 2 numbers as much as we can. But can you count the total number of passcodes to be generated?

If you toss a coin, it is easy to determine the possible outcomes because it could just be either head or tail. If you toss 4 coins altogether, how will you determine the total number of possible outcomes?

This is where the counting principles and techniques will assist in giving the exact count with speed and accuracy.



Counting Technique 1: The Listing Method

This method suggests that you list all possible outcomes and simply count them. This is the most basic of all methods and usually it is employed when the total outcomes expected are very few.

Example 1

How many possible outcomes when a fair coin is tossed?



Solution:

Any fair coin will give you Head (H) or Tail (T) when tossed. Therefore, there are 2 outcomes expected when a coin is tossed.

Example 2

How many possible outcomes when a fair dice is tossed?



Solution

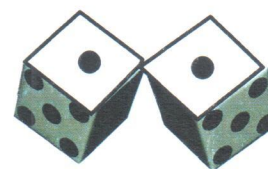
A fair dice is a cube with 6 faces where each face represents the numbers 1-6.



Therefore, there are six possible outcomes and they are 1, 2, 3, 4, 5 and 6.

Example 3

When a pair of fair dice is tossed, list all possible outcomes.



Solution:

Imagine a pair of fair dice being tossed and thinking both dice will show the same face so we can now determine that outcomes include (1,1), (2,2), (3,3), (4,4), (5,5) and (6,6) but there's more and listing all of them can be a bit confusing.

To neatly and accurately present the list, a table like the one on the next page can be used:



		Possible outcomes of the first dice					
Possible outcomes of the second dice		1	2	3	4	5	6
	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
	5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
	6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

The table shows a neat and organized listing of the **36 outcomes** when a pair of fair dice is tossed.

Counting Technique 2: The Tree Diagram

This technique does not only aid counting but helps to create a more accurate and organized listing as well. In addition, this technique is used when more data variables are included in the outcome.

Example 4

A restaurant offers a promotion where the customer can make his own set of meal for only K25 by choosing a chicken meal, a side dish, and a drink in the list of menu shown below:

Chicken	Side Dish	Drinks
Roasted	Mashed Potato	Cola
Fried	Chips	Orange juice
Grilled		Bottled Water

The manager of the restaurant would like to know the total number of meal set possible and wants to see the specific list in every set.

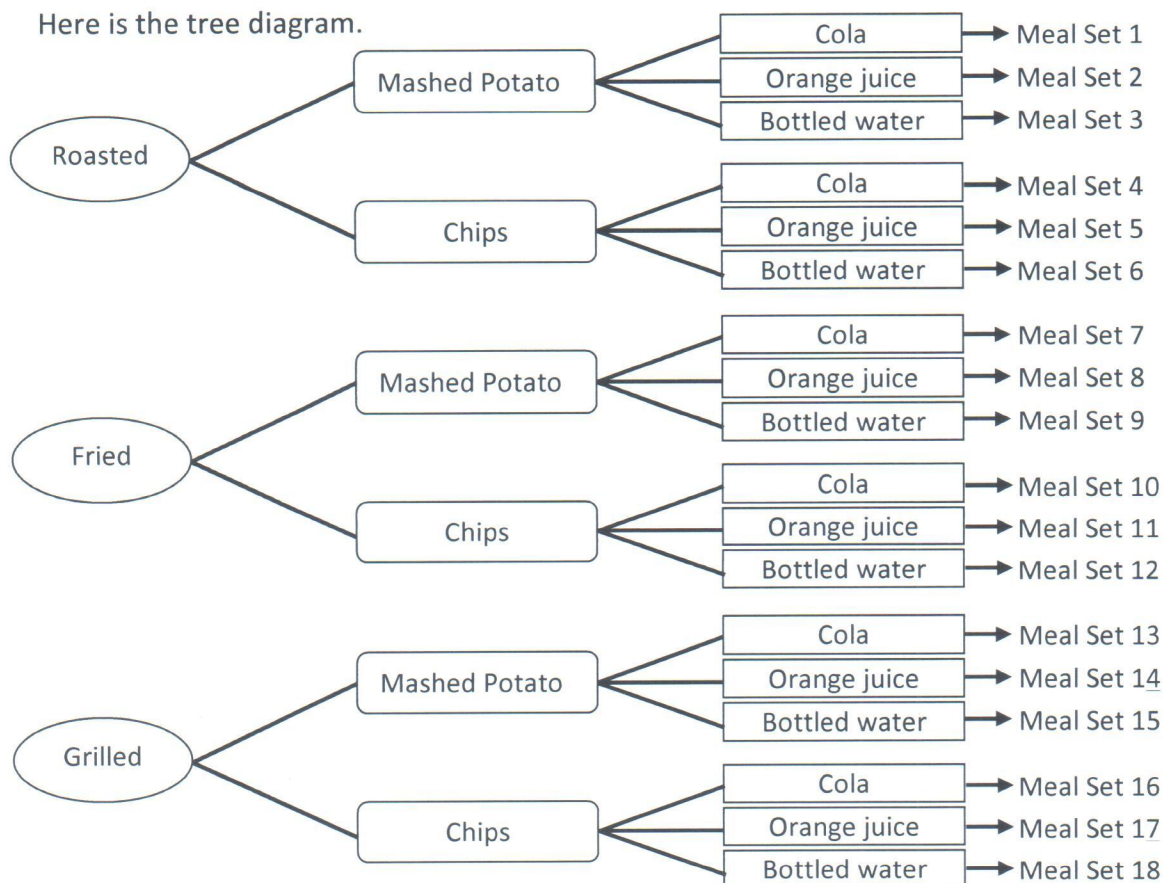
Find the number of “meal set” the restaurant offers and make a listing of the sets.

Solution:

To make an accurate list of meal sets, we will use a tree diagram. See next page.



Here is the tree diagram.



Now looking at the diagram, we can see 18 sets of meals and they can be listed as follow:

Promo Meal	Content
Meal Set 1	Roasted Chicken, Mashed Potato and Cola
Meal Set 2	Roasted Chicken, Mashed Potato and Orange Juice
Meal Set 3	Roasted Chicken, Mashed Potato and Bottled Water
Meal Set 4	Roasted Chicken, Chips and Cola
Meal Set 5	Roasted Chicken, Chips and Orange Juice
Meal Set 6	Roasted Chicken, Chips and Bottled Water
Meal Set 7	Fried Chicken, Mashed Potato and Cola
Meal Set 8	Fried Chicken, Mashed Potato and Orange Juice
Meal Set 9	Fried Chicken, Mashed Potato and Bottled Water
Meal Set 10	Fried Chicken, Chips and Cola
Meal Set 11	Fried Chicken, Chips and Orange Juice
Meal Set 12	Fried Chicken, Chips and Bottled Water
Meal Set 13	Grilled Chicken, Mashed Potato and Cola
Meal Set 14	Grilled Chicken, Mashed Potato and Orange Juice
Meal Set 15	Grilled Chicken, Mashed Potato and Bottled Water
Meal Set 16	Grilled Chicken, Chips and Cola
Meal Set 17	Grilled Chicken, Chips and Orange Juice
Meal Set 18	Grilled Chicken, Chips and Bottled Water



This method is very helpful when you need to count and list at the same time. Let us try another example.

Example 5

Determine the possible number of outcomes when a K1 coin and a fair dice are tossed together. Make a listing of the outcomes:

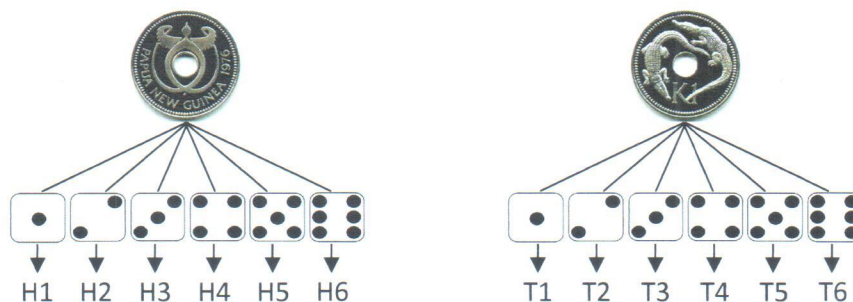


Solution:

A K1 coin has 2 faces, the head (H)  and the tail (T)  while a fair dice has 6 faces.



We can draw a tree diagram like the one below:



Now we can clearly see that the 12 possible outcomes are H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5 and T6.

Counting Technique 3: The Fundamental Counting Principle

In some cases, like in computing probability which will be discussed in the succeeding topics, the complete list of possible outcomes may no longer be necessary but the total number of outcome does. And in most cases, the number of outcomes can be really big. Imagine the number of outcomes in a lottery draw or the number of possible pass code combinations using 3 letters and two numbers.

The third technique will be very useful for these cases.

The Fundamental Counting Principle

If there are **A** ways for one event to occur, and **B** ways for another event to occur, then there are **A x B** ways for both to occur.

This principle applies to any fixed number of event denoted by **A x B x C...**

Where A, B, C... represents fixed, certain or exact values.

**Example 6**

A mobile phone application needs a passcode using two letters and 2 numbers. How many passcodes can possibly be used for this mobile application?

Solution:

In this case since we need 2 letters and 2 numbers for a passcode, there are 4 events namely:

- First event: Using 1 letter out of the 26 letters in the English alphabet.
- Second event: Using another letter (maybe the same or different from the first) out of the 26 letters in the English alphabet.
- Third event: Using 1 digit from the digits 0,1,2,3,4,5,6,7,8 or 9.
- Fourth event: Using 1 digit (maybe the same or different from the first) from the digits 0,1,2,3,4,5,6,7,8 or 9.

Applying the Fundamental Counting Principle we say:

If there are **26** ways for the first event to occur, 26 ways for the second event to occur, 10 ways for the third event to occur and another 10 ways for the fourth event to occur, then there are **$26 \times 26 \times 10 \times 10$** ways for all of them to occur at the same time which form the passcode.

Calculate: $26 \times 26 \times 10 \times 10 = 67\,600$

Therefore, there are 67 600 passcodes that can be possibly used for the mobile application.

Example 7

In a fund raising activity, a raffle is held where each raffle ticket must bear a control ID with the letter A, B or C with 2 numbers.

- a) Find the total number of tickets to sell for this fund raising activity.
- b) Find the number of tickets with NO repeated digits.

Solution:

- a) There are 3 events involved here.
 - First event: Using 1 letter out of the 3 (A, B or C)
 - Second event: Using 1 digit out of the 10 digits (0, 1, 2,3,4,5,6,7,8 or 9)
 - Third event: Using 1 digit (maybe the same or different from the first) out of the 10 digits.



Applying the Fundamental Counting Principle we say:

If there are **3** ways for the first event to occur, 10 ways for the second event to occur, and 10 ways for the third event to occur, then there are **$3 \times 10 \times 10$** ways for all of them to occur at the same time which forms the control ID for raffle tickets in a fund raising activity.

Calculate: $3 \times 10 \times 10 = 300$

Therefore, there are 300 tickets to sell for the fund raising activity

b) There are also 3 events involved here.

First event: You have 3 choices using 1 letter out of the 3 (A, B or C).

Second event: You have 10 choices using 1 digit out of the digits 0, 1, 2,3,4,5,6,7,8 or 9.

Third event: Using 1 digit **different from the first**. Since we are only counting the one with NO repeated digits, therefore, you have only 9 digits to choose from since in the second event a digit was already chosen.

Applying the Fundamental Counting Principle we say:

If there are **3** ways for the first event to occur, 10 ways for the second event to occur, and 9 ways for the third event to occur, then there are **$3 \times 10 \times 9$** ways for all of them to occur at the same time which forms the control ID for raffle tickets in a fund raising activity.

Calculate: $3 \times 10 \times 9 = 270$

Wondering what are the 30 control IDs with repeating digits?

These are the control IDs with repeating digits which were subtracted from the 300 total.

A00, A11, A22, A33, A44, A55, A66, A77, A88, A99

B00, B11, B22, B33, B44, B55, B66, B77, B88, B99

C00, C11, C22, C33, C44, C55, C66, C77, C88, C99

Therefore, there are 270 control ID tickets with NO repeated digits

Now, you are ready to try the following learning activities to challenge yourself to count more intelligently.

Now do the learning activity on the next page.

**Learning Activity 12.3.1.1**

20 minutes

- 1) Joshua and John play toss coin. As they toss 3 coins at the same time, they list the outcomes they can observe. How many outcomes will there be in their list? Show your complete list.

List:

Total number of outcomes:

- 2) A planner offers a “wedding package” where the couple can choose from 3 varieties of flowers available, 5 color motifs and 3 wedding set ups as detailed below:

Flower	Color Motif	Theme Set-up
Roses	Pearl White	Church Wedding
Birds of Paradise	Golden Yellow	Beach Wedding
Tulips	Luscious Purple	Garden Wedding
	Serene Green	
	Royal Blue	

Use a tree diagram to show ALL possible “Wedding Packages” to offer to clients and list them on the table provided:

Tree Diagram



Wedding Package	Description / Inclusion



- 3) Christen has 3 sport shirts, 2 blouses, and 3 shorts which she can “mix and match” and wear in school. How many possible matches can she have?
- 4) Lisa tossed 2 coins and 4 fair dices all together. How many possible outcomes can she draw?



12.3.1.2: Permutation

Now that you have already familiarized yourselves with the different counting techniques, let us move to another way of counting but this time we will count the possible arrangements of objects involved.

Permutation is an ordered arrangement of items that occurs when:

- a) no item is used more than once.
- b) the order of arrangement makes a difference.

Example 1

If you are asked to arrange the letters **ABC**, how many possible arrangements can you have?

Solution:

You may list them as follows:

ABC ACB BAC BCA CAB CBA

Therefore, you may conclude that there are 6 possible arrangements when the letters **A**, **B** and **C** were arranged. This may look easy if the objects to arrange are very few.

Notice that no letter was used more than once and the order of arrangement makes a difference.

The fundamental counting principle can be used to determine the number of permutations of **n** objects.

For instance, you can find the number of ways you can arrange the letters **A**, **B**, and **C** by multiplying.

There are 3 choices for the first letter, 2 choices for the second letter, and 1 choice for the third letter, so there are $3 \times 2 \times 1 = 6$ ways to arrange the letters.

In general, the number of permutations of distinct objects (**n**) is:

$n!$ read as “n factorial”

where:

$$n! = n \times (n - 1) \times (n - 2) \times \dots \times 3 \times 2 \times 1$$

0! (zero factorial), by definition, is 1. ($0! = 1$)



Example 2

In how many ways can the letters in the word LOVE be arranged?

Solution:

Since there are 4 distinct letters to arrange, we use $4! = 4 \times 3 \times 2 \times 1 = 24$.

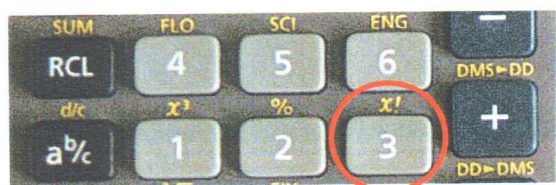
Therefore, there are 24 ways the letters in the word LOVE can be arranged.

Scientific Calculators can also be used for a bigger number such as $50!$, $100!$ and so on. Just look at the symbol $x!$ in your calculator which is usually found on the 2nd level of your calculator.

The following are samples where the $x!$ can be found.

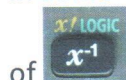
If you want to get $9!$

Using the model below where $x!$ is found on the number keys on top of

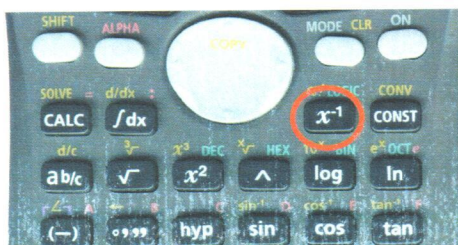


Simply press and it will show $9!$ in your calculator screen. Pressing will give you **362880**.

In some models, $x!$ can be found on the function keys. In the model below it is found on top of



of



Simply press and it will show $9!$ in your calculator screen. Pressing will give you **362880**.



Example 3

Twelve runners from different provinces are competing in the final round of a marathon entitled “best of the best runners”.

- In how many different ways can the runners finish the competition? We assume that there are no ties because time is counted by seconds.
- In how many different ways can 3 of the runners finish first, second, and third to win the gold, silver, and bronze medals?

Solution:

- Since there are 12 runners, therefore the number of ways the runners finish the competition is counted by $12!$

$$12! = 12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 479\,001\,600$$

Therefore, there are 479 001 600 ways the runners finish the competition.

- Any of the 12 runners can finish first, then any of the remaining 11 runners can finish second, and finally any of the remaining 10 runners can finish third.

So, we can count the number of ways the runners can win the medals (gold, silver, bronze) by multiplying: $12 \times 11 \times 10 = 1320$

Therefore, there are 1320 ways wherein 3 of the runners finish first, second, and third and win the gold, silver, and bronze medals.

Example 4:

In how many ways can 4 Algebra, 3 Geometry, 2 Trigonometry and 2 Calculus books be arranged in a shelf if:

- arrangement is in no particular order?
- books of the same topic must be arranged next to each other?
- Algebra books are arranged next to each other while the rest may be arranged in any order?



Solution:

- Since arrangement needs no particular order, therefore the number of ways the books may be arranged in a shelf is counted by $(4+3+2+2)! = 11!$

$$11! = 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 39\,916\,800$$



Therefore, there are 39 916 800 ways 4 Algebra, 3 Geometry, 2 Trigonometry and 2 Calculus books can be arranged in a shelf.

- b) If books of the same topics must be arranged next to each other, let us first get the number of ways we can arrange them next with each other per topic.

Algebra books:	$4! = 4 \times 3 \times 2 \times 1$	$= 24$
Geometry books: $3!$	$3! = 3 \times 2 \times 1$	$= 6$
Trigonometry books: $2!$	$2! = 2 \times 1$	$= 2$
Calculus books: $2!$	$2! = 2 \times 1$	$= 2$

Since the books can be moved and arranged per topic (Algebra, Geometry, Trigonometry, Calculus) we have $4!$

$$\begin{aligned}4! &= 4 \times 3 \times 2 \times 1 \\ &= 24\end{aligned}$$

We can now solve the number of ways books of the same topic be arranged next to each other by multiplying.

$$\begin{aligned}4! 3! 2! 2! 4! &= 24 \times 6 \times 2 \times 2 \times 24 \\ &= 13\,824\end{aligned}$$

Therefore, there are 13 824 ways books of the same topic be arranged next to each other.

- c) If Algebra books are arranged next to each other ($4!$) while the rest may be arranged in any order $(11-4)!$ Or $7!$ Then, we count by:

Algebra books:	$4! = 4 \times 3 \times 2 \times 1$
	$= 24$

Rest of the books (3 Geometry + 2 Trigonometry + 2 Calculus):
$7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$
$= 5\,040$

Now we simplify $4! 7! = 24 \times 5\,040 = 120\,960$.

Therefore, there are 120 960 ways Algebra books are arranged next to each other while the rest may be in any order.



Example 5

In how many ways can 10 people seat in a round table?

Solution:



If they are seated in row or column, we can simply say it is $10!$ But since they are in a round table, if all of them will move one seat to their right, they will still be seated next to same person and their arrangement does not change. This is an example of a case of **Circular Permutation**.

Circular permutations of distinct objects (n) is
 $(n - 1)!$

Following the formula, we have $(10-1)! = 9! = 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 362\,880$

Therefore, there are 362 880 ways 10 people can be seated in a round table.

Sometimes, we have a set of objects represented by n . And we only want to take a certain number represented by r . We can count the number of arrangements using the following formula:

$${}_nP_r = \frac{n!}{(n-r)!}$$

Where: ${}_nP_r$ is read as “the number of permutations of r items taken from n ”

n is the total number of items

r is the number of items to be arranged taken from n .

Please note that r can be equal or less than n but cannot be greater than n .

Example 6

From a class of 20 students the teacher needs to form a committee of 3 students to assist her in leading the class in the field trip. One has to take the role of the class captain, one to be the class monitor and the third one to be class secretary.

In how many ways can the teacher form the committee if all of the 20 students are equally competent to take any role?

Solution:

In this case the total number of students (n) = 20 and we will take a number to be in the committee at a time (r) = 3.



Using the formula: ${}_nP_r = \frac{n!}{(n-r)!}$

Substituting the values of n and r in the formula we have:

$${}_{20}P_3 = \frac{20!}{(20-3)!}$$

Simplify: ${}_{20}P_3 = \frac{20!}{17!}$

At this point, you can use your scientific calculator to solve for this by simple pressing **20! – 17!** Then **EXE** or equal sign (=).

Without the calculator, you can also solve for this manually by further simplifying the formula:

$$\begin{aligned} {}_{20}P_3 &= \frac{20 \times 19 \times 18 \times 17 \times 16 \times 15 \times \dots \times 1}{17 \times 16 \times 15 \times \dots \times 1} \\ &= \frac{20 \times 19 \times 18}{1} \\ &= 6\,840 \end{aligned}$$

Please note that $17 \times 16 \times 15 \times 14 \times 13 \times 12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ is equal to $17!$

$$\begin{aligned} \text{The simplified working out can also be written as } {}_{20}P_3 &= \frac{20 \times 19 \times 18 \times 17!}{17!} \\ &= 20 \times 19 \times 18 \\ &= 6\,840 \end{aligned}$$

This is a neater working out since we will cancel $17!$ And we are left with $20 \times 19 \times 18 = 6\,840$

Example 7

There are 10 books in a shelf. In how many ways can they be arranged?

- If all books will be included in the arrangement?
- If 4 books were taken out and only 6 books were left to be arranged?

Solution:

- If all books will be included, we can simply say $10!$. This is similar with the Examples 1, 2 and 3. But for you to fully understand, we will use the formula to show how it was derived.
-



In this case, the total number of books (n) = 10 and we will take and arrange (r) = 10.

Substitute the values of n and r in the formula: ${}_nP_r = \frac{n!}{(n-r)!}$

$$\begin{aligned}\text{Hence, we have } {}_{10}P_{10} &= \frac{10!}{(10-10)!} \\ &= \frac{10!}{0!}\end{aligned}$$

Remember that $0!$ is equal to 1. Therefore, when $n = r$, the denominator will become $0!$ and the numerator is the given number itself in factorial form.

$$\begin{aligned}\text{Therefore, } {}_{10}P_{10} &= \frac{10!}{0!} \\ &= 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 \\ &= 3\,268\,800\end{aligned}$$

There are 3, 268, 800 ways to arrange 10 books in a shelf.

- c) If 4 books were taken out and only 6 books were left to be arranged, we can now say that $n = 10$ while $r = 6$. Why 6 and not 4? Remember, r represents the number of items taken out to be arranged. In this example, although 4 were taken out, the 6 left were the ones to be arranged.

Substitute the values of n and r in the formula, ${}_nP_r = \frac{n!}{(n-r)!}$

$$\begin{aligned}\text{Hence, we have } {}_{10}P_6 &= \frac{10!}{(10-6)!} \\ &= \frac{10!}{4!} \\ &= \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4!}{4!} \\ &= 10 \times 9 \times 8 \times 7 \times 6 \times 5 \quad \text{Cancel out } 4! \\ &= 151\,200\end{aligned}$$

Therefore, there are 151 200 ways to arrange 6 books randomly taken from 10 books.



Another type of arrangement that you can encounter is **Permutations of duplicate items**.

In this case, the number of permutations of n items, where n_1 items are identical, n_2 items are identical, n_3 items are identical, and so on, is given by:

$${}_nP_r = \frac{n!}{(n_1!)(n_2!)(n_3!)\dots}$$

In this formula, the denominator consists of the number of identical or duplicated items multiplied by each other (if there is more than 1 duplicated item).

Example 8

In how many distinct ways can the letters of the word MISSISSIPPI be arranged?

Solution:

Since there are a total of 11 letters in the word MISSISSIPPI, therefore $n = 11$.

n_1 will represent the number of letter I in the word, therefore $n_1 = 4$

n_2 will represent the number of letter s in the word, therefore $n_2 = 4$

n_3 will represent the number of letter p in the word, therefore $n_3 = 2$

Substitute the given values in the formula: ${}_nP_r = \frac{n!}{(n_1!)(n_2!)(n_3!)\dots}$

$$\begin{aligned} \text{We have } {}_{11}P_r &= \frac{11!}{(4!)(4!)(2!)} \\ &= \frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4!}{(4!)(4!)(2!)} && \text{Expand cancel out 4!} \\ &= \frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5}{(4!)(2!)} \\ &= \frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5}{4 \times 3 \times 2 \times 1 \times 2 \times 1} \\ &= \frac{1\,663\,200}{48} && \text{Simplify} \\ &= 34\,650 \end{aligned}$$

A scientific calculator can be used to make the computation a lot easier and faster. But remember these manual computations too, will make you more competent as you learn to deal with bigger numbers manually.

Try to practice using the calculator too and familiarize yourself with both techniques.

**Learning Activity 12.3.1.2**

20 minutes

Solve the following problems. Show your working out.

- 1) In how many ways can the letters in the word JUSTICE be arranged?
- 2) In how many ways can the vowel letters be arranged?
- 3) In how many ways can the numbers 0-9 be arranged?
- 4) In how many ways can the name CHERRIE be arranged?
- 5) In how many ways can the following pots be arranged



- a) In no particular order?



b) If pots of the same flowers are placed next to each other?

6) If 5 boys and 7 girls are arranged in row, how many ways can they be lined up



a. in no particular order?

b. if the girls and the boys must be together?

c. if 3 girls must be in the first 3 spots while the rest may be arranged in no particular order?

7) In how many ways can 10 leaders be seated in a round table for a meeting?



8) There are 8 athletes in a marathon. In how many ways can the champion, first runner up and second runner up be determined if all of the athletes are equally competitive?





12.3.1.3 Combination

Combination is a counting technique wherein the number of groupings and not order of the objects is being counted. The order or arrangement of objects is not important or does not really matter.

Unlike Permutation, in Combination we count the possible groupings of objects and not the arrangement itself. We can note that a combination of items occurs when

- the items are selected from the same group.
- no item is used more than once.
- the order of items makes no difference.

Suppose a teacher forms a committee of 3 students from a class of 20 to assist her in leading the class in the field trip. This time, the members of the committee will play the same role. No one will act as the class captain, monitor or secretary.

Suppose you, John and Maria were consecutively selected on the first draw, how will the committee differ if John was selected first, Maria, second, and you, the third to be appointed by the teacher?

Does the committee vary when the order of selection of members as they were called differed?

Good observation if you say that the committee chosen on the first draw is the same as the committee drawn in the second regardless of the order the members were called or selected by the teacher.

This kind of situation involves combination. Take note that John, Maria and you were taken from the same group of 20 students in the class. That each of you is a distinct individual and the order you were called makes no difference.

The number of possible combinations if r items are taken from n items is determined by the formula:

$${}_nC_r = \frac{n!}{r!(n-r)!}$$

Where: ${}_nC_r$ is read as “the combination of r taken n at a time”

n is the total number of items

r is the number of items to be included in the group taken from n .

Note that r can be equal or less than n but cannot be greater than n . And when $n = r$, the combination is always 1, which means all the members are taken from the same group.



Example 1

Let us try to solve the previous example presented in a situation earlier.

Suppose a teacher forms a committee of 3 students from a class of 20 to assist her in leading the class in the field trip. How many possible committees can the teacher form?

Since we already discussed earlier that the members of the committee will play the same role, no one will act as the class captain, monitor or secretary. All of the selected committee members will play the same role. Therefore, order of the members is not really important in this case.

First identify the given in the problem:

The total number of students (n) = 20

The committee will consist of 3 members at a time (r) = 3.

Using the formula ${}_nC_r = \frac{n!}{r!(n-r)!}$

First we substitute the values of n and r in the formula ${}_{20}C_3 = \frac{20!}{3!(20-3)!}$

Calculate: ${}_{20}C_3 = \frac{20!}{3!(17)!}$

$$= \frac{20 \times 19 \times 18 \times 17!}{3! \times 17!}$$

$$= \frac{20 \times 19 \times 18}{3!}$$

$$= \frac{20 \times 19 \times 18}{3 \times 2 \times 1}$$

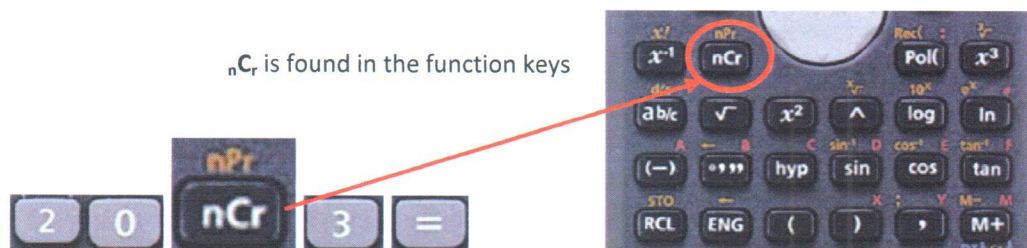
$$= \frac{6840}{6}$$

$$= 1140$$

Therefore, there are 1 140 ways a committee of three members can be formed from 20 students.



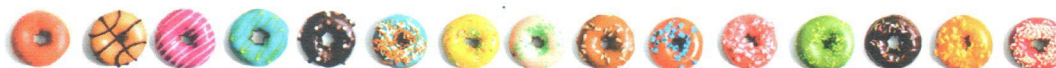
Using the calculator, is as easy as a breeze... just press the following keys



The following examples provide working out for you to fully understand the use of formula. However, it is also encouraged that you try your calculator to verify answers. Through this you will be capable of answering problems both manually and with the aid of calculators.

Example 2

A bakery sells a dozen of donuts of various kind and flavor for half the prize before its closing time. Dorothy found that there are 15 donuts left when she visited the store. How many dozens of different donuts are possible?



Solution: First identify the given in the problem.

The total number of donuts (n) = 15.

Dorothy must pick a dozen donuts at a time (r) = 12.

Using the formula ${}_nC_r = \frac{n!}{r!(n-r)!}$, substitute the values of n and r .

We have ${}_{15}C_{12} = \frac{15!}{12!(15-12)!}$

Simplify: ${}_{15}C_{12} = \frac{15!}{12!(3)!}$

$$= \frac{15 \times 14 \times 13 \times 12!}{12! \times 3!}$$

Cancel out 12!

$$= \frac{15 \times 14 \times 13}{3!}$$

$$= \frac{15 \times 14 \times 13}{3 \times 2 \times 1}$$

$$= \frac{2730}{6}$$

$$= 455$$

Therefore, there are 455 different dozens possible out of the 15 donuts left on sale.



Example 3

There are 12 females and 13 males in a Grade 12 class. A committee of 6 members is to be formed to organize their class graduation party at the end of the school year.

Calculate the number of ways the committee will be selected if:



- Any member of the class can be chosen to form part of the committee.
- The committee must be composed of 3 females and 3 males from the class.
- The committee must be composed of 2 females and 4 males from the class.

Solution:

- The total number of students $(n) = 12 + 13 = 25$
The number of members of the committee taken at a time $(r) = 6$

Using the formula ${}_nC_r = \frac{n!}{r!(n-r)!}$

First we substitute the values of n and r in the formula ${}_{25}C_6 = \frac{25!}{6!(25-6)!}$

$$\begin{aligned} \text{Simplify: } {}_{25}C_6 &= \frac{25!}{6!(19)!} \\ &= \frac{25 \times 24 \times 23 \times 22 \times 21 \times 20 \times 19!}{19! \times 6!} \end{aligned}$$

$$\begin{aligned} \text{Cancel out } 19!, &= \frac{25 \times 24 \times 23 \times 22 \times 21 \times 20}{6!} \\ &= \frac{25 \times 24 \times 23 \times 22 \times 21 \times 20}{6 \times 5 \times 4 \times 3 \times 2 \times 1} \\ &= \frac{127\,512\,000}{720} \\ &= 177\,100 \end{aligned}$$

Therefore, there are 177 100 different committees possible out of the 25 students in the class.

- If the committee must be composed of 3 females and 3 males from the class, first we have to solve them separately.

Choosing 3 out of the 12 females we have ${}_{12}C_3$. Where $n = 12$ and $r = 3$.

Using the formula ${}_nC_r = \frac{n!}{r!(n-r)!}$



First we substitute the values of n and r in the formula:

$${}_{12}C_3 = \frac{12!}{3!(12-3)!}$$

Simplify:

$$\begin{aligned} {}_{12}C_3 &= \frac{12!}{3!(9)!} \\ &= \frac{12 \times 11 \times 10 \times 9!}{3! \times 9!} \end{aligned}$$

Cancel out $9!$, we have

$$\begin{aligned} &= \frac{12 \times 11 \times 10}{3!} \\ &= \frac{12 \times 11 \times 10}{3 \times 2 \times 1} \\ &= \frac{1320}{6} \\ &= 220 \end{aligned}$$

Choosing 3 out of the 13 males we have ${}_{13}C_3$. Where $n = 13$ and $r = 3$.

Using the same formula we substitute the values of n and r .

$${}_{13}C_3 = \frac{13!}{3!(13-3)!}$$

Simplify:

$$\begin{aligned} {}_{13}C_3 &= \frac{13!}{3!(10)!} \\ &= \frac{13 \times 12 \times 11 \times 10!}{3! \times 10!} \end{aligned}$$

Cancel out $10!$, we have

$$\begin{aligned} &= \frac{13 \times 12 \times 11}{3!} \\ &= \frac{13 \times 12 \times 11}{3 \times 2 \times 1} \\ &= \frac{1716}{6} \\ &= 286 \end{aligned}$$

Since there are ${}_{12}C_3 = 220$ ways having females and ${}_{13}C_3 = 286$ ways having males, a committee consisting of having 3 females and 3 males is derived by multiplying ${}_{12}C_3$ and ${}_{13}C_3$.

$${}_{12}C_3 \cdot {}_{13}C_3 = 220 \cdot 286 = 62\,920$$

Therefore, there are 62 920 ways of forming a committee of six with 3 females and 3 males.



- c) If the committee must be composed of 2 females and 4 males from the class, we can simply solve that using example b as our guide. This time, we will try to solve it continuously so you can have another way of answering it.

Having 2 from 12 females is represented by ${}_{12}C_2$ where $n = 12$ and $r = 2$, while having 4 males out of the 13 is represented by ${}_{13}C_4$ where $n = 13$ and $r = 4$.

To get the number of ways a committee with 2 females and 4 males is formed, we simply multiply ${}_{12}C_2$ and ${}_{13}C_4$.

Using the formula $({}_{12}C_2)({}_{13}C_4)$, let us substitute the respective givens.

$$\begin{aligned} \left[\frac{12!}{2!(12-2)!} \right] \times \left[\frac{13!}{4!(13-4)!} \right] &= \left[\frac{12!}{2!(10)!} \right] \times \left[\frac{13!}{4!(9)!} \right] \\ &= \frac{12 \times 11 \times 10!}{2!(10)!} \times \frac{13 \times 12 \times 11 \times 10 \times 9!}{4!(9)!} \quad \leftarrow \begin{array}{l} \text{Cancel out } 10! \text{ and } 9!, \\ \text{then expand the} \\ \text{denominator.} \end{array} \\ &= \frac{12 \times 11}{2 \times 1} \times \frac{13 \times 12 \times 11 \times 10}{4 \times 3 \times 2 \times 1} \\ &= \frac{132}{2} \times \frac{17160}{24} \\ &= 66 \times 715 \\ &= 47\,190 \end{aligned}$$

Therefore, there are 47 190 ways of forming a committee of six with 2 females and 4 males.

Now do the learning activity.



Learning Activity 12.3.1.3



30 minutes

- Find the value of the following: (You may verify your answers using a scientific calculator.)

a) ${}_5C_2 =$ _____



b) ${}_{13}C_8 =$ _____

c) ${}_{15}C_{10} =$ _____

d) ${}_{12}C_9 =$ _____

e) ${}_{18}C_{16} =$ _____

f) ${}_{25}C_{20} =$ _____

g) ${}_5C_4 \cdot {}_8C_6 =$ _____

h) ${}_6C_2 \cdot {}_4C_3 =$ _____

i) ${}_4C_3 + {}_5C_3 =$ _____

j) ${}_8C_4 + {}_5C_5 =$ _____



2) Solve the following problems.

a) How many sub-groups with 5 members can be formed from a group of 15 dancers?

b) In an international sports festival, PNG will be represented by a total of 10 athletes. The following are the trained athletes that can represent the country.

Athletics Event	Number of Athletes Trained
Track and Field	10
Shot Put	5
Javelin Ambulant	6
Hurdles and Jumps	8
Pole Vault	5

- i. How many ways can the country be represented if there are 2 athletes from every event?
- ii. How many possible groups are there if there are 6 representatives from the track and field group and 4 from the hurdles and jumps group?
- iii. How many possible groups can be made if all members from the shot puts will join and the rest may come from any athletics group?

**SUMMATIVE TASK 12.3.1**

40 minutes

1. Find the value of the following. You may use your scientific calculator to verify the answers.

a) $8!$ = _____

b) $(9 - 6)!$ = _____

c) $2!3!5!$ = _____

d) ${}_{15}P_6$ = _____

e) ${}_{10}P_3$ = _____

f) ${}_{18}C_4$ = _____

g) ${}_{21}C_7$ = _____

h) ${}_6C_3 \cdot {}_5C_3$ = _____

i) ${}_5C_3 + {}_6C_3$ = _____

2. How many outcomes will there be if two coins and a fair dice are tossed?

3. How many 6-digit numbers can be drawn in a 0 to 42 lottery draw?



4. Liza is trying to make a passcode consisting of 4 letters from A to J and 4 even numbers. How many possible passcodes can she form if repetition of letters and numbers is not allowed?



12.3.2: Probability

Probability is the language used in mathematics to model certainty or the “chance” that something will happen.

Maybe you are wondering how we can measure chance. Is that possible?

Let us consider this situation.

Your Mum called from the office saying that she misplaced her pair of earrings in the living room. It happened that you and your brother were the only ones left in the house and you tried searching for it. What is the chance that you will be the one to find the missing earrings?

Since there are only 2 of you in the house, you might say you have “one in two” or one half chance of finding it. And your brother has an equal chance of finding it too. Mathematically speaking, each of you has 50% chance. Right? And your Mum who is no longer in the house has no chance to find it at that moment.

Let us say, when your Mum called more people are in the house. How will you describe your chances of being the person to find it? Did your chance increase or decrease?

Of course, the above situation sounds too easy to consider. But as you go along with the module, you will learn how to calculate simple probability of events and analyze data using probability.

The data and examples we will be looking at and discussing the outcomes of simple mathematics and scientific experiments. These outcomes will show many different kinds of uncertainty and randomness. These will lead you to develop a proper understanding of experimental results, and lead you to be able to understand the randomness and other underlying principles involved in probability.

12.3.2.1: Fundamentals of Probability

We begin this topic by discussing the fundamentals or basic concepts in probability. Since we already equated the word “chance” with probability, we will simplify the definitions as follows.

Probability is the measure of the chance or likelihood that an event will occur or happen.

When something is certain NOT to happen or we say it is impossible to happen, then it has zero probability.



Sample spaces, events and sets

In Mathematics when dealing with probabilities, an activity in which we perform a number of trials to enable us to measure the chance of a certain things may occur is called an **experiment**. This is not the similar experiment conducted in laboratories but somehow, this may be similar in a way that the activity aims to test possibility of random outcomes or results.

Simple activity of tossing a coin is considered as an experiment. Because this activity can lead us to outcomes such as getting Heads or Tails.

The set of all possible outcomes of the experiment is known as the **sample space**. It is usually denoted by S (capital S), and an element of the sample space or an outcome is denoted by s .

A **sample space** is the set of all possible outcomes of an experiment.

By doing an experiment of tossing a coin, the sample space $S = \{\text{Heads, Tails}\}$ and s can either be Heads or Tails only.

Example 1

Jade rolled a fair die 20 times and he listed what he got in a table as shown below.

Outcomes	Number of occurrence
	
	
	
	≡
	==
	

Identify the experiment, outcomes and sample space.



Solution:

Experiment: rolling a die

Outcomes: (may be in random order) landing with face up showing the following



Sample space: $S = \{1, 2, 3, 4, 5, 6\}$

Using the sample space in Example 1, $S = \{1, 2, 3, 4, 5, 6\}$

Subsets can also be drawn like getting even numbers or odd numbers. These subsets which are outcomes taken from the sample set s are called events.

An **event** is a collection of outcomes having a common characteristics from the sample space. It is a subset of the sample space wherein its elements were taken from the sample space. A **simple event** consists of exactly one outcome and a **compound event** consists of more than one outcomes.

If Jade rolled a fair die, the event of getting even numbers is $S_{(\text{even})} = \{2, 4, 6\}$, the event of getting odd numbers $S_{(\text{odd})} = \{1, 3, 5\}$, and the event of getting a number greater than 5 is $S_{(\text{greater than } 5)} = \{6\}$. $S_{(\text{even})}$ and $S_{(\text{odd})}$ are compound events because there are three outcomes in each event while $S_{(\text{greater than } 5)}$ is a simple event wherein the outcome is exactly one only.

Example 2

Given the sample space as the set of integers $S = \{\dots -3, -2, -1, 1, 2, 3\dots\}$, identify the following events:

- a) Getting whole numbers.
- b) Getting an integer less than 1 but greater than -1.
- c) Getting an integer less than zero.

Solution:

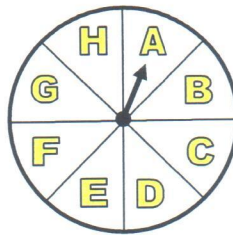
- a) The set of whole numbers is a subset of the sample space (set of Integers), therefore $S_{(\text{whole numbers})} = \{0, 1, 2, 3\dots\}$ the three dots (ellipsis) means that the set goes on to positive infinity.
- b) The integer less than 1 but greater than -1 is zero. Other values in between -1 and 1 are fractions and decimals, therefore $S_{(\text{less than } 1 \text{ but greater than } -1)} = \{0\}$.
- c) The set of integers less than zero are the set of negative numbers, therefore, $S_{(\text{less than zero})} = \{-1, -2, -3, -4\dots\}$ the three dots (ellipsis) means that the set goes on to negative infinity.

**Learning Activity 12.3.1.3**

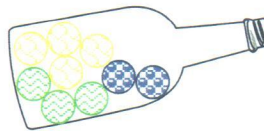
15 minutes

Identify the sample space (S) in the following by listing the complete set of outcomes:

1. Tossing two coins
2. Tossing a coin and a die
3. Spinning the spinner



4. Drawing marbles from a bottle which contains 2 blue marbles (BM), 3 green marbles (GM) and 4 yellow marbles (YM).



5. What will happen if you only brought the following clothes in your out-of-town tour?



- a) Choose a method to be able to make neat listings of possible trousers-shirt pairs. Use two letters to write the pair. Example TL means White Shirt (first) paired with Long Pants (last).
- b) List the sample space for the trouser-shirt pairs. Use two letters to write the pair. Example TL means White Shirt (first) paired with Long Pants (last).



12.3.2.2 Probability of Events

In this lesson, we will deal with two important terminologies we have discussed previously: **probability and event**.

We have defined probability as the measure of the chance or likelihood that an event will occur or happen while an event is a collection of outcomes having a common characteristic from the sample space.

We may classify events into four types.

- The **null event** is the empty subset of the sample space.
- A **simple event** is a subset consisting of a single element of the sample space.
- A **compound event** is a subset consisting of more than one element of the sample space.
- The **sample space** itself is also an event.

Understanding the two important terminologies will enable you to fully understand what probability of events means.

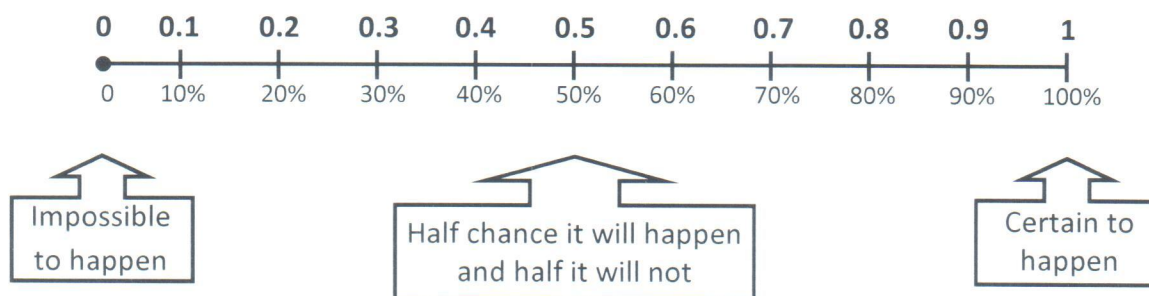
The Probability of Event **P(E)** or also known as classical probability or simple probability is concerned with carrying out probability calculations based on equal or likely outcomes. That is, we assume that each element in the sample space have the same chance.

We calculate the Probability of Event as

$$P(E) = \frac{E}{n}$$

Where: **E** is the number of ways E can occur
n is the total number of outcomes

The calculated value of **P(E)** is usually in between 0 to 1 and it may be interpreted as follows:



A PROBABILITY SCALE



If $P(E)$ is equal to 1, it means that it is 100% sure to happen, while 0 $P(E)$ means it is impossible to happen. There is no chance that the event will occur if $P(E) = 0$, that is why some reference note that probability may come close to zero but not really zero. Like a probability of 0.0000001 may seem equal or too close to zero but still it is not equal to zero itself. Therefore a cloud of chance is still seen.

Example 1

Suppose a fair die is rolled, what is the probability of

- a) getting a whole number?
- b) getting an even number?
- c) getting a number less than 1?
- d) getting a number less than 3?

Solution:

Since you are already familiar with rolling a fair die, and we have noted in our previous discussions that the total number of outcomes in rolling a die is 6 (that is getting 1, 2, 3, 4, 5 or 6), then we say that $n = 6$.

- a) The event is getting a whole number since we all know that 1, 2, 3, 4, 5 and 6 are all whole numbers, therefore $E = 6$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n , we have $P(E) = \frac{6}{6} = 1$

(Multiply it by 100 to express the probability as percent) so $1 \times 100 = 100\%$

Therefore, the probability of getting a whole number when rolling a die is 100%.

- b) The event is getting an even number. The even numbers we can derive in rolling a die are 2, 4 and 6, therefore $E = 3$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n , we have $P(E) = \frac{3}{6} = 0.5$

(Multiply it by 100 to express the probability as percent) so $0.5 \times 100 = 50\%$

Therefore, the probability of getting an even number when rolling a die is 50%.



- c) The event is getting a number less than 1. Since we all know that the least number we can derive in rolling a die is 1, therefore the event is null. $S(\text{less than } 1) = \emptyset \frac{E}{n}$ (empty set)

Therefore, it is impossible to draw a number less than 1 when rolling a die.

- d) The event is getting a number less than 3. These numbers are 2 and 1, therefore $E = 2$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n, we have $P(E) = \frac{2}{6} = 0.33333...$

(Multiply it by 100 to express the probability as percent) so $0.33333... \times 100 = 33.33...%$ or $33 \frac{1}{3} \%$.

Therefore, the probability of getting a number less than 3 when rolling a die is $33 \frac{1}{3} \%$.

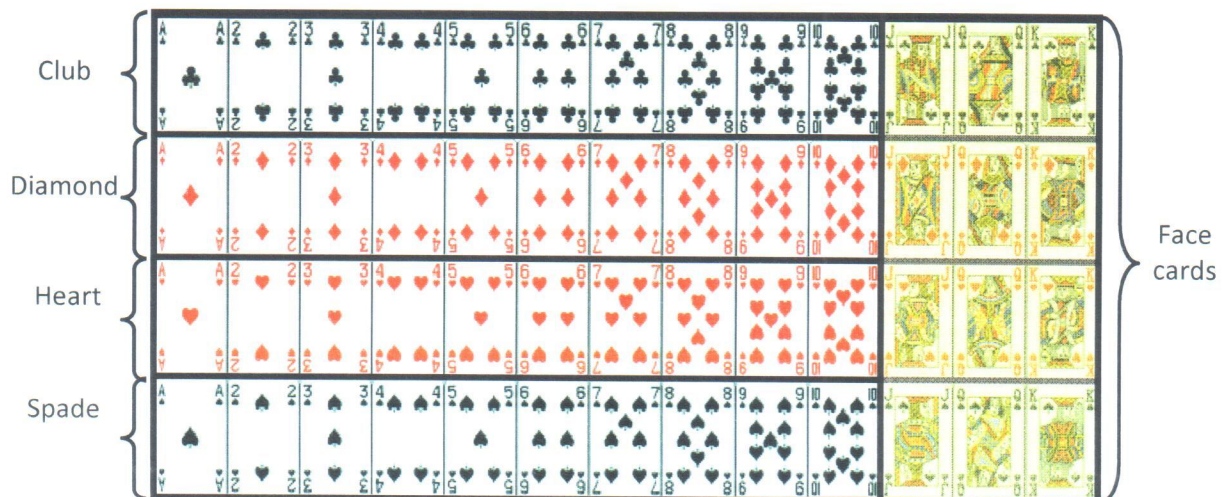
Example 2

In an ordinary deck of playing cards, compute the probability of

- drawing an Ace card
- drawing a red card.
- drawing a face card.
- drawing a hearts card.
- drawing an Ace of club.
- drawing a joker.

Solution:

An ordinary deck of cards consists of 52 cards, therefore, we can say that $n = 52$.





The set of diamonds and set of hearts are also called as “red cards” while the set of clubs and the set of spades are called “black cards”.

The “face cards” consist of all Jack, Queen and King cards.

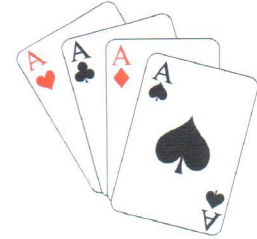
- a) The event is drawing an Ace card. Since there are 4 Aces in the deck of cards, therefore $E = 4$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n , we have $P(E) = \frac{4}{52} = 0.076923$

(Multiply it by 100 to express the probability as percent) so $0.076923 \times 100 = 7.69\%$.

Therefore, the probability of drawing an Ace card from an ordinary deck is 7.69%.



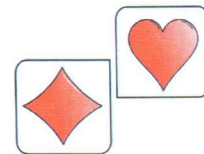
- b) The event is drawing a red card. There are 13 hearts card and 13 diamonds cards, therefore $E = 26$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n , we have $P(E) = \frac{26}{52} = 0.5$

(Multiply it by 100 to express the probability as percent) so $0.5 \times 100 = 50\%$.

Therefore, the probability of drawing a red card from an ordinary deck is 50%.



- c) The event is drawing a face card. There are 12 face cards, therefore $E = 12$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n , we have $P(E) = \frac{12}{52} = 0.230769$

(Multiply it by 100 to express the probability as percent) so $0.230769 \times 100 = 23.08\%$.

Therefore, the probability of drawing a face card from an ordinary deck is 23.08%.



- d) The event is drawing a spade card. There are 13 spade cards, therefore $E = 13$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n , we have $P(E) = \frac{13}{52} = 0.25$

(Multiply it by 100 to express the probability as percent) so $0.25 \times 100 = 25\%$.

Therefore, the probability of drawing a spade card from an ordinary deck is 25%.





- e) The event is drawing an Ace of club card, there is only one Ace of club card, therefore $E = 1$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n, we have $P(E) = \frac{1}{52} = 0.01923$

(Multiply it by 100 to express the probability as percent) so $0.01923 \times 100 = 1.92\%$.

Therefore, the probability of drawing an Ace of club card from an ordinary deck is 1.92%.



- f) The event is drawing a joker card, in an ordinary deck of cards, the joker is not included, therefore $E = \emptyset$. We can also say that $E = 0$. If we apply the formula, we will divide 0 by 52 and the result is still 0.

Therefore, there is no chance of drawing a joker card from an ordinary deck.



Example 3

In a school fund raising, the tickets being sold have a control number from 000 to 999. The solicitor (the person who sold the ticket) wins a special prize for every winning ticket sold. What is the probability of winning in the raffle draw if you have sold

- a) 50 tickets?
b) 200 tickets?
c) Only 1 ticket?



Solution:

First we have to determine the sample space or the total number of tickets (n). Knowing that the tickets were numbered using 3 digits and repetitions of digits are allowed, we use the fundamental counting principle to determine n.

Since the digits includes 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9, therefore there are **10** ways for the first digit to occur, 10 ways for the second digit to occur and **10** ways for the last digit to occur. Then there are **$10 \times 10 \times 10 = 1\ 000$** ways for all the 3 digits found in the tickets to occur.

So we now say that $n = 1\ 000$.

- a) The event is winning in the ticket draw for selling 50 tickets, therefore $E = 50$.

Using the formula $P(E) = \frac{E}{n}$



Substitute the values of E and n, we have $P(E) = \frac{50}{1000} = 0.05$

(Multiply it by 100 to express the probability as percent) so $0.05 \times 100 = 5\%$.

Therefore, the probability of winning in the raffle draw as a solicitor for selling 50 tickets is 5%.

- b) The event is winning in the ticket draw for selling 200 tickets, therefore $E = 200$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n, we have $P(E) = \frac{200}{1000} = 0.2$

(Multiply it by 100 to express the probability as percent) so $0.2 \times 100 = 20\%$.

Therefore, the probability of winning in the raffle draw as a solicitor for selling 200 tickets is 20%.

- c) The event is winning in the ticket draw for selling 1 ticket only, therefore $E = 1$.

Using the formula $P(E) = \frac{E}{n}$

Substitute the values of E and n, we have $P(E) = \frac{1}{1000} = 0.001$

(Multiply it by 100 to express the probability as percent) so $0.001 \times 100 = 0.1\%$.

Therefore, the probability of winning in the raffle draw as a solicitor for selling just 1 ticket is 0.01%.

This example shows that no matter how small the probability may seem, there is still chance of winning. Take note, in counting probability, each raffle ticket or coupon has 0.01% chance of being picked for the win.

Lesson? Never lose hope. And the cliché is true that “the more coupons you have, the more chances of winning”.

What is the difference now that you know how to compute probability?

This time, you will be more accurate in computing your chances. First you have to determine the sample space to make an accurate computation.

Now, try the following learning activities for you to challenge yourself.

**Learning Activity 12.3.2.2**

20 minutes

Solve the following problems. Express the final answer in percent %.

- 1) There are 15 boys and 20 girls in a class. The teacher will have to draw the name of 1 student to represent the class in a Mathematics challenge quiz bee. What is the probability that the representative is

a) a boy?

b) a girl?

- 2) There are 50 different cookies in a jar. Half of them are oatmeal cookies while 10 of them are chocolate cookies. The rest of the cookies are butter caramel flavored. If a child will randomly pick cookies from the jar, find the probability that the cookie is

a) an oatmeal?



b) a chocolate?



c) a butter caramel?

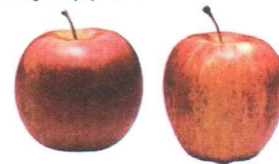


- 3) In a shop, 4 employees are female and 8 employees are male. What is the probability that the employee of the month is a female?





- 4) In a grocery shop, a box contains 50 apples, 15 of them are Gala apples and the rest are Fuji apples. What is the probability that a shopper will pick a Fuji apple?



- 5) A jar contains 8 red marbles, 5 blue marbles and 7 green marbles. What is the probability of drawing

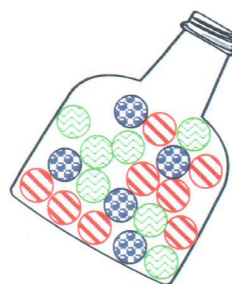
- a) a red marble?



- b) a green marble?



- c) a blue marble?

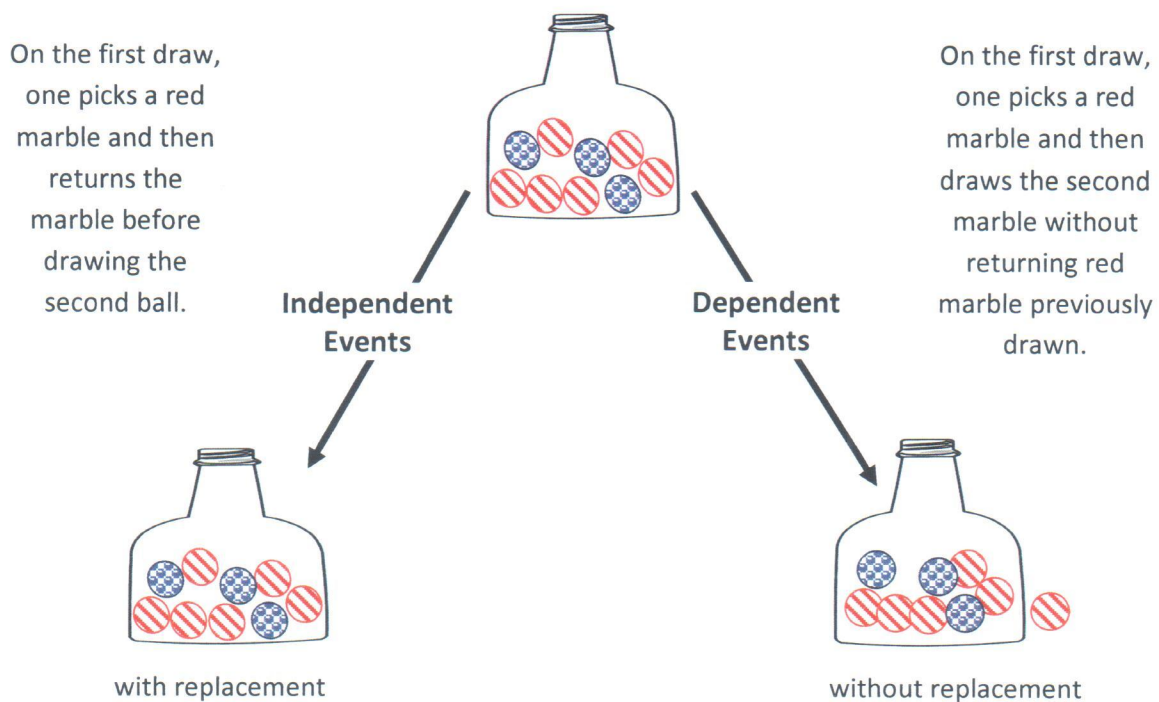


12.3.2.3 Independent and Dependent Events

In the previous lesson, we have discussed what an event means and classified them. We also computed the probability of a single event. In this lesson we will discuss compound events. **Compound events** refer to two or more events occurring simultaneously. These simultaneous events may be classified as Independent or Dependent events.

Events are **independent** when the outcome of one event **does not influence** the outcome of the next event. On the other hand, two or more events are **dependent** if the outcomes of the first affect the others.

The illustration on the next page shows the difference between dependent and independent events.



Probability of independent events

The illustration above shows that in dealing with independent events, **replacement** is being done before the next occurrence.

If A and B are independent events, then the probability that both A and B occur is

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

The above formula means that for two independent events, the probability that both events occur is the product of the probabilities of the events. The formula can be used for more than two independent events. Just multiply the probability of each event.

Example 1

The tethering function of an android mobile phone randomly generates passcodes using the digits 0 to 9, wherein each digit can be used more than once. What is the probability that the first digit is 7 and the second digit is an odd number?

Solution:

This problem involves independent events since the occurrence of the first digit does not affect the occurrence of the second digit.





$P(A)$ is the probability that the first digit is 7. Since having 7 in the first digit is a simple event wherein it will only occur once, therefore $E = 1$ and $n = 10$ since there are 10 digits from 0 to 9. We can write $P(A) = \frac{1}{10}$. (It is advisable to retain it in the form of a fraction to avoid rounding error when dealing with decimals. It is also easier to multiply fractions than decimals with a number of place values.)

$P(B)$ is the probability that the second digit is an odd number. Since 1, 3, 5, 7, and 9 are the odd digits from 0 to 9, therefore $E = 5$ and $n = 10$ since there are 10 digits from 0 to 9. We can write $P(B) = \frac{5}{10} = \frac{1}{2}$

Since $P(A)$ and $P(B)$ were already determined, we can now substitute their values in the formula (or simply multiply them) to get the probability of both events.

Formula: $P(A \text{ and } B) = P(A) \cdot P(B)$

Substitute $P(A)$ and $P(B)$: $P(A \text{ and } B) = \frac{1}{10} \cdot \frac{1}{2}$ (recall the rule in multiplying fractions)

Simplify: $P(A \text{ and } B) = \frac{1}{20}$ (Express it in percent by dividing 1 by 20 and multiplying the resulting decimals to 100.)

$$\begin{aligned} P(A \text{ and } B) &= .05 \\ &= 5\% \end{aligned}$$

Therefore, the probability of having 7 in the first digit and an odd number in the second digit of a randomly generated passcode is 5%.

Example 2

Hurray Lucky Mall gives their customers the chance to join their lucky spin promotion for every purchase of a laptop or desktop computer. Lucy bought two units of laptops for their office needs. She is then entitled to spin the wheel twice.

What is the probability that she will lose by getting a sorry in the first spin and wins the jackpot on her second spin?





Solution:

This problem also involves independent events since the first spin on the wheel does not affect the second spin.

$P(A)$ is the probability that Lucy loses by getting a “sorry”. Since there are 6 occurrence of sorry in the wheel therefore $E = 6$ and $n = 12$ since the wheel is divided into 12 equal parts.

We can write $P(A) = \frac{6}{12}$ or $\frac{1}{2}$ (Again, it is advisable to retain it in a form of a fraction to avoid rounding error when dealing with decimal. It is also easier to multiply fractions than decimals with a number of place values.)

$P(B)$ is the probability that Lucy will hit the jackpot. Since hitting the jackpot is a simple event wherein it will only occur once, therefore $E = 1$ and $n = 12$. We can write $P(B) = \frac{1}{12}$.

Since $P(A)$ and $P(B)$ were already determined, we can now substitute their values in the formula (or simply multiply them) to get the probability of both events.

Formula: $P(A \text{ and } B) = P(A) \cdot P(B)$

Substitute $P(A)$ and $P(B)$: $P(A \text{ and } B) = \frac{1}{2} \cdot \frac{1}{12}$ (apply the rule in multiplying fractions)

Simplify: $P(A \text{ and } B) = \frac{1}{24}$ (Express it in percent by dividing 1 by 24 and multiplying the resulting decimals to 100)

$$\begin{aligned} P(A \text{ and } B) &= .04167 \\ &= 4.17\% \end{aligned}$$

Therefore, the probability that Lucy will lose by getting a sorry in the first spin and wins the jackpot on her second spin is 4.17%.

Did you notice that in both $P(A)$ and $P(B)$ we have used the same value for n ?

This is what makes the problems independent. If n of $P(A)$ is equal with n of $P(B)$ it suggest that there is a replacement done in the elements of the sample space.



Probability of dependent events

Referring back on the illustration on page 48, it shows that in dealing with dependent events, **no replacement** is being done before the next occurrence.

If A and B are dependent events, then the probability that both A and B occur is

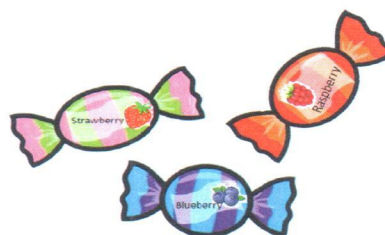
$$P(A \text{ and } B) = P(A) \cdot P(B \text{ given } A)$$

The above formula means that for two dependent events, the probability that both events occur is the product of the probability that the first event occurs and the probability that the second event occurs given that the first event has occurred. The formula can be used for more than two dependent events. Just always consider the occurrence of the previous events.

Example 3

A jar of assorted candies consist of the following:

- 25 pieces strawberry candies
- 30 pieces blueberry candies
- 35 pieces raspberry candies



Angel picked 2 candies at random. What is the probability that the first candy she picked is raspberry and the second is a blueberry?

Solution:

This problem involves dependent events since drawing the first candy Angel has a total of 90 candies to choose from and there are only 89 left to choose from when drawing the second candy.

P(A) is the probability that Angel picks a raspberry candy. Since there are 35 raspberry candies, then $E = 35$ and $n = 90$ which is the total of all the candies ($25 + 30 + 35$). We can write $P(A) = \frac{35}{90}$ or $\frac{7}{18}$.

P(B) is the probability that Angel picks a blueberry candy on her second draw. Since there are 30 blueberry candies, then $E = 30$ and $n = 89$. Why not 90? On the second draw n becomes 89 as we have to subtract the raspberry candy picked by Angel on his first draw. Since she has to pick two candies, she will not return the first candy picked, thus making n equal to 89 ($90 - 1$). We can write $P(B) = \frac{30}{89}$.

Since **P(A)** and **P(B)** were already determined, we can now substitute their values in the formula (or simply multiply them) to get the probability of both events.



Formula: $P(\text{A and B}) = [P(\text{A})][P(\text{B})]$

Substitute $P(\text{A})$ and $P(\text{B})$: $P(\text{A and B}) = \left(\frac{7}{18}\right)\left(\frac{30}{89}\right)$ (apply the rule in multiplying fractions)

Simplify: $P(\text{A and B}) = \frac{210}{1602} = \frac{35}{267}$

(Express it in percent by dividing 35 by 267 and multiplying the resulting decimals to 100.)

$$\begin{aligned}P(\text{A and B}) &= 0.13109 \\ &= 13.11\%\end{aligned}$$

Therefore, the probability of picking a raspberry candy on the first draw and a blueberry candy on the second is 13.11%.

Example 4

A purse contains three K1 coins and five 50 toea coins. Find the probability of choosing first a 50 toea coin and then, without replacing the 50 toea, choosing a K1 second.

Solution:

This problem shows dependent events since choosing a coin without replacement is done.

P(A) is the probability that a K1 coin is chosen on the first draw. Since there are 3 K1 coins, then $E = 3$ and $n = 8$ which is the sum total of all the coins in the purse ($3 + 5$). We can write

$$P(\text{A}) = \frac{3}{8}.$$

P(A) is the probability that a 50t coin is chosen on the second draw. Since there are five 50 toea coins, then $E = 5$ and $n = 7$. At this point you might know the reason why $n = 7$ and not 8. Great if you say that in the second draw n becomes 7 because we no longer count the K1 coin in the first draw. Since the problem suggests that there is no replacement after the first draw, thus making n equal to 7 ($8 - 1$). We can write $P(\text{B}) = \frac{5}{7}$

Since **P(A)** and **P(B)** were already determined, we can now substitute their values in the formula (or simply multiply them) to get the probability of both events.

Formula: $P(\text{A and B}) = [P(\text{A})][P(\text{B})]$

Substitute $P(\text{A})$ and $P(\text{B})$: $P(\text{A and B}) = \left(\frac{3}{8}\right)\left(\frac{5}{7}\right)$ (apply the rule in multiplying fractions)



Simplify: $P(A \text{ and } B) = \frac{15}{56}$

(Express it in percent by dividing 15 by 56 and multiplying the resulting decimals to 100.)

$$\begin{aligned} P(A \text{ and } B) &= 0.26786 \\ &= 26.79\% \end{aligned}$$

Therefore, the probability of picking a K1 coin on the first draw and a 50t coin on the second draw is 26.79%.



Learning Activity 12.3.2.3



20 minutes

- 1) Identify if the events described below are **dependent** or **independent events**.
 - a) Rolling a number cube and tossing a coin.
 - b) Tossing three coins simultaneously.
 - c) Choosing three cards consecutively from a standard deck without returning the cards previously drawn.
 - d) Picking a marble and then picking a second marble without replacing the first marble on the box.
 - e) Your teacher chooses students at random to solve a Mathematics problem on the board. She chooses you, and then another student from the remaining students in the class.
 - f) You have to draw 10 coupons from the drop box in a collection of 2 000 coupons. You select a coupon, put it aside, and select another until you complete the 10 coupons before announcing the names in the winning coupons.
 - g) You have a packet of assorted candies where 30 pieces are chocolate candies and 20 pieces are orange candies. You take one piece of candy at random from the packet, put it back, and then take a second piece of candy at random.



- h) Daniel has a blue, red, and green shirt. He also has a blue and green trousers. Daniel chooses a random shirt and trouser to wear.
- i) Leila plays NRL trading cards. She picks a card at random. Then without putting the first card back, she picks a second card.
- 2) Solve the following.

The table below shows the status of 500 registered grade 12 students from FODE NCD. A student is randomly selected to participate in the course evaluation program.

Status	Gender	
	Male	Female
Full-time student	90	120
Working Student	210	80

- a) What is the probability that the student is a female regardless of her status?
- b) What is the probability that a full-time student will be chosen regardless of its gender?
- 3) Ruru received a bouquet of flowers with 5 red roses and 7 red tulips. What is the probability of picking a red rose randomly from the bouquet?



12.3.2.4 Conditional Probability

In our previous discussions, we have already developed ways of understanding the probabilities of events. At this point, we will explore ways of modifying those probabilities when certain events occur and other more complex events.

In problem number 2 of your Learning Activity 12.3.2.3 on page 47, you used the following table.

Status	Gender	
	Male	Female
Full-time student	90	120
Working Student	210	80

It was easy to determine the probability that the student is a female regardless of her status by simply following the formula presented, right?

Now, what if you need to determine the probability that the student chosen is a male given that he or she is a working student? How will you solve it?

This kind of problem is more than just a case of dependent and independent events. This involves conditional probability.

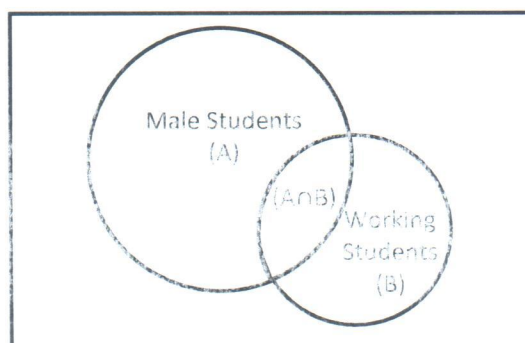
For two events A and B , the probability that A will occur given that B has already occurred is the conditional probability of A given B written as $P(A/B)$ where:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Note that $A \subseteq B$ and $A \cap B = A$. This means that event A is a special case of event B .

How do we compute $P(A/B)$? Since we are given that the student drawn is a male (event A) and he is a working student (event B), we can forget about the females. Thus, all outcomes outside event B (full time students) are also found to be irrelevant.

Let us use a Venn diagram to visualize the Male Working Students ($A \cap B$) from the group.





The big circle represents the set of male students while the small circle represents the working students. The part shaded in black is the intersection of **A** and **B** which refers to the set of male working students. Everything outside the two circles but within the rectangle represents outcomes involving female and fulltime students.

Example 1

A fair die is rolled and the outcomes were recorded. Given that the number rolled was odd, what is the probability that it was a five?

Solution:

In this case, we will solve for the probability of getting 5 given that getting an odd number has already occurred. So with this we let **A** represent the event of getting 5 and **B** represent the event of getting an odd number.

This problem involves two events: getting a number 5 and getting an odd number. The probability that the number drawn is 5 will occur given that an odd number has already been drawn suggest that this is an example of a conditional probability. The condition is, an odd number must be drawn first before it will occur that it is a 5.

First we identify $P(A \cap B)$. Since $(A \cap B) = \{5\}$ then we say that $P(A \cap B) = \frac{1}{6}$. There is only one odd number which is equivalent to 5 out of the 6 possible outcomes when a die is rolled. Now we identify $B = \{1, 3, 5\}$, there are 3 odd numbers out of the six possible outcomes in rolling a die, therefore $P(B) = \frac{3}{6}$ or $\frac{1}{2}$.

Now, using the formula

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Substitute the values of $P(A \cap B)$ and $P(B)$,

$$= \frac{\frac{1}{6}}{\frac{1}{2}}$$

Recall that in dividing fractions, we multiply the numerator by the reciprocal of the denominator:

$$= \frac{1}{6} \times \frac{2}{1} = \frac{2}{6}$$

Simplify $\frac{2}{6}$

$$= \frac{1}{3} = 0.33\bar{3}$$

Express it in percent by dividing 1 by 3 and multiplying the resulting decimals to 100.

$$= 33.33\%$$

Therefore, when a number rolled in a fair die is an odd, the probability that it was a five is 33.33%.



So now we know that in a conditional probability, the probability of something happening depends on the previous event(s) that had happened.

Example 2

Daisy and Donald have two children. What is the probability that their second child is a girl given that at least one of their children is a girl?

Solution:

In this case we use the assumption that boys (**B**) and girls (**G**) are equally likely to be born when a mother conceives. We will solve for the probability of Daisy and Donald having a girl for their second child given that at least one of their children is a girl so it means that it has occurred already that their first born could be a girl or a boy. So with this we let A represent the event of having a girl for their second child and B represent the event of having at least one of their children is a girl.

If we will determine all possibilities, Daisy and Donald can have either two boys (**BB**), a boy and a girl (**BG**), a girl and a boy (**GB**) and two girls (**GG**). So we say our sample space is 4 which refer to the total number of combination possible.

First we identify $P(A \cap B)$. Since $(A \cap B) = \{BG, GG\}$ then we say that $P(A \cap B) = \frac{2}{4}$. There are two out of four possibilities of having a girl as their second child.

Now we identify $B = \{GG, BG, GB\}$, there are 3 possible outcomes where both children are girls, the first born is a boy and the second born is a girl and the first born is a girl and the second is a boy, therefore $P(B) = \frac{3}{4}$.

Now, using the formula

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Substitute the values of $P(A \cap B)$ and $P(B)$,

$$= \frac{\frac{2}{4}}{\frac{3}{4}}$$

Recall that in dividing fractions, we multiply the numerator by the reciprocal of the denominator:

$$= \frac{2}{4} \times \frac{4}{3} = \frac{8}{12}$$

Simplify $\frac{8}{12}$

$$= \frac{2}{3} = 0.66\bar{6}$$

Express it in percent by dividing 2 by 3 and multiplying the resulting decimals to 100.

$$= 66.67\%$$

The probability that Daisy and Donald will have a girl on their second child is 66.67%.



Example 1 and 2 show the use of formula, however, other problems involving the conditional probability can be solved using analysis and the use of diagrams just like the following worked example.

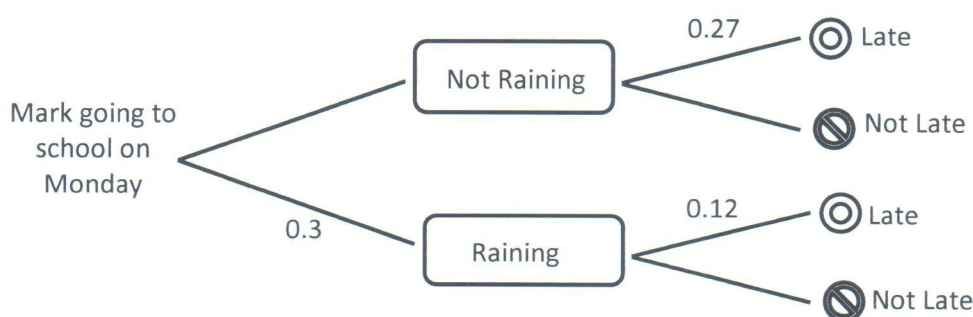
Example 3

The probability that Mark will be late for school on Monday is 0.27 based on his usual Monday experiences due to heavy traffic. But if it will rain, less vehicles will travel and the probability goes down to 0.12. According to the news, the probability of rain on Monday is 0.3.

What is the probability that Mark will be late on Monday?

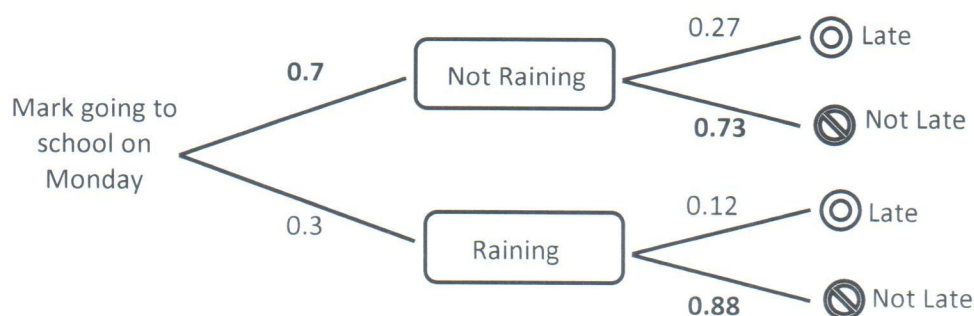
Solution:

We can use a tree diagram to carefully analyze the problem.



We know that the probability that a thing is 100% to happen is 1. So if the chance of raining on Monday is 0.3, the probability that it will not rain is 0.7 ($1 - 0.3$). Do you agree? In a similar way, we can say that if it will not rain, Mark has a 0.27 chance of being late and 0.73 ($1 - 0.27$) chance of not being late for school. However, if it rains, his chance of being late is 0.12 and not being late is 0.88 ($1 - 0.12$).

Now we add those data in the diagram for further analysis.





Now, we solve Mark's probability of being late by adding the probability of being late when it is not raining and the probability of him being late when it is raining.

$$\begin{aligned} P(\text{being late}) &= P(\text{late when not raining}) + P(\text{late when raining}) \\ &= (0.7 \times 0.27) + (0.3 \times 0.12) \\ &= 0.189 + 0.036 \\ &= 0.225 \text{ (express it into per cent by multiplying with 100)} \\ &= 22.5\% \end{aligned}$$

Therefore, the probability that Mark will be late on Monday is 22.5%.

Mutually exclusive events

As we deal with conditional probabilities, we may also encounter events which may or may not happen at the same time.

Events are **mutually exclusive** if they have no outcomes in common. This is the same as saying that these two events cannot happen together at the same time. It is impossible that both could occur in a single trial of an experiment. Their intersection is empty.

On the other hand, two events that can happen at the same time are called **non-mutually exclusive events**.

In the die-toss example, given $A = \{2\}$, $B = \{1, 3, 5\}$ and $C = \{2, 4, 6\}$ A and B are mutually exclusive events since $A \cap B = \emptyset$ while A and C are non-mutually exclusive events $A \cap C = \{2\}$.

Mutually exclusive events have nothing in common or they cannot occur simultaneously. For instance, you cannot roll a 5 and a 6 on a single roll of a die, or land on both heads and tails in a single toss of a coin.

Addition vs. Multiplication

When looking for the union of mutually exclusive events, the addition of the individual probabilities is used.

If events A and B are mutually exclusive, then

$$P(A \text{ or } B) = P(A) + P(B).$$

With more formal notation it also mean $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$



It is important that you recognize its other notation because other references use different forms of notations.

Example 4

What is the probability of drawing an ace out of an ordinary deck of cards?

Solution:

Drawing aces is an example of mutually exclusive event. You cannot draw an ace of hearts together with an ace of diamond in a single draw. There are four aces, one for each suit, and each has a probability of $\frac{1}{52}$.

In this case if we let **A** = ace of spades, **B** = ace of diamonds, **C** = ace of hearts and **D** = ace of clubs,

$$\begin{aligned}\text{We say that: } P(\mathbf{A} \text{ or } \mathbf{B} \text{ or } \mathbf{C} \text{ or } \mathbf{D}) &= P(\mathbf{A}) + P(\mathbf{B}) + P(\mathbf{C}) + P(\mathbf{D}) \\ &= \frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} \\ &= \frac{4}{52} \text{ or } \frac{1}{13}\end{aligned}$$

Therefore, the probability of drawing an ace out of an ordinary deck of cards is $\frac{1}{13}$ or 7.69%

The general probability formula in the previous lesson can also be used to deal with example 4. However, not all cases of mutually exclusive events can be solved using the general formula.

Example 5

What is the probability of drawing a red card or a club card from an ordinary deck of cards?

Solution:

Drawing a red card or a club card is an example of mutually exclusive events. Clubs are black and they cannot be red so it is impossible that a club card drawn is a red card at the same time.

We let **P(A)** as the probability of drawing a red card. Since there are 13 heart cards and 13 diamond cards, therefore $P(\mathbf{A}) = \frac{1}{26}$.



Now, we let $P(B)$ as the probability of drawing a club card. Since there are also 13 club cards, then we say $P(B) = \frac{1}{13}$.

$$\begin{aligned}\text{We say that: } P(A \text{ or } B) &= P(A) + P(B) \\ &= \frac{1}{26} + \frac{1}{13} \quad (\text{recall addition of fractions}) \\ &= \frac{3}{26}\end{aligned}$$

Therefore, the probability of drawing a red card or a club card from an ordinary deck of cards is $\frac{3}{26}$ or 11.54%

Take note that you will use the addition of probabilities when dealing with union of mutually exclusive events. The key word that you will have to remember is “or”.

Multiplication of probabilities is used when dealing with intersection of non-mutually exclusive events. Remember that these events happen at the same time.

It follows from the formula for conditional probability that for any events A and B,

$$\begin{aligned}P(A \text{ and } B) \text{ or } P(A \cap B) &= P(B/A) P(A) \\ &= P(A/B) P(B)\end{aligned}$$

In general, the formula suggests that in order to find the probabilities of successive events, multiply conditional probabilities given the previous events.

Example 6

Two cards are chosen at random without replacement from a well-shuffled deck of ordinary cards. What is the probability that they are both queens?

Solution:

Choosing two queens on a consecutive draw is possible. It is simply saying drawing a queen and a queen.

We let $P(A)$ as the probability of drawing a queen on the first draw. Since there are 4 queens out of the 52 cards, therefore $P(A) = \frac{4}{52}$



Now, we let $P(B)$ as the probability of drawing another queen. Since a queen was already drawn there are 3 remaining queens out of the 51 remaining cards, then we say $P(B) = \frac{3}{51}$.

We say that: $P(A \text{ and } B) = \left(\frac{4}{52}\right)\left(\frac{3}{51}\right)$ (recall multiplication of fractions)

$$= \frac{12}{2652}$$

Simplify $= \frac{1}{221}$

Therefore, the probability of drawing two queens is $\frac{1}{221}$ or 0.45%.

Take note that you will use the multiplication of probabilities when dealing with intersection of non-mutually exclusive events. The key word that you will have to remember is “and”.

Always remember the key words.

Example, if the problem says “a queen **and** a king” then it suggests that the multiplication rule is to be used since you will be dealing with the intersection of non-mutually exclusive events. In the case that the problem says “a queen **or** a king” then it suggests that the addition rule is to be used for most likely, you will be dealing with the union of mutually exclusive events.

Revise what you have learnt in this lesson and be ready to answer the following learning exercises.



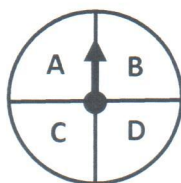
Learning Activity 12.3.2.4



20 minutes

Solve the following problems involving conditional probabilities.

1. On the spinner, what is the probability of spinning an A or a B?





2. Two cards are removed simultaneously from a deck of 48 cards. Calculate the probability of drawing:
 - a. two diamonds
 - b. at least one diamond
 - c. one diamond and one heart
 3. In your pocket, you have three 50 toea coins and two 20 toea coins.
 - a. What is the probability of selecting a 50 toea coin?
 - b. What is the probability of selecting a 20 toea coin?
 - c. If you select a 50 toea coin and place it on a table, what is the probability that the coin selected is a dime?
 4. If two dice are rolled, what is the probability that the first die shows the number 5?
 5. What is the probability of drawing a queen out of an ordinary deck of playing cards?
 6. What is the probability of drawing two Aces from a standard deck of playing cards, given that the first card is an Ace?
-

Congratulations!

You have reached the end of topic 12.3.2. Spend time revising what you have learnt and proceed to the summative task to further enhance your skill.

NOW DO THE SUMMATIVE TASK 12.3.2

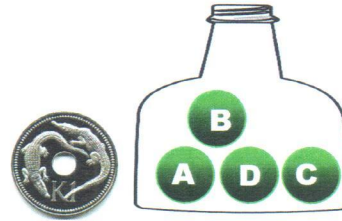
**SUMMATIVE TASK 12.3.2**

30 minutes

Circle the letter of the correct answer.

- 1) Which of the following shows the sample space (S) when a coin is tossed and a ball is drawn from the jar?

- A. $S = \{HT, AB, CD\}$
- B. $S = \{H, T, A, B, C, D\}$
- C. $S = \{HTA, HTB, HTC, HTD\}$
- D. $S = \{HA, HB, HC, HD, TA, TB, TC, TD\}$



- 2) Juliana tossed a coin and she got heads on the first toss and tails on the second. What is the experiment in the situation?

- A. Tossing a coin
- B. Getting heads and tails
- C. Getting a heads on the first toss
- D. Chance of getting heads on the first toss

- 3) It refers to a collection of outcomes from the sample space which has common characteristics.

- A. Experiment
- B. Sample space
- C. Event
- D. Probability

- 4) Which of the following is more likely to happen?

- A. $P(A) = 0.23$
- B. $P(B) = 0.72$
- C. 0.072
- D. 0.032

- 5) A box contains 30 light bulbs. During shipment and mishandling, 8 of them broke and became defective. What is the probability that person who opens the box will pick a non-defective bulb?

- A. $\frac{8}{30}$
- B. $\frac{22}{8}$
- C. $\frac{22}{30}$
- D. $\frac{8}{22}$



- 6) The following are examples of dependent events **except** _____.
- | | |
|-------------------------|---|
| A. Tossing a coin twice | C. Drawing two cards from a single deck |
| B. Rolling two dice | D. Tossing a coin and spinning a wheel |
- 7) A machine generates random reference numbers using 3 digits from 0 to 9, wherein the digits cannot be used more than once. This scenario is an example of
- | | |
|-----------------------|-----------------|
| A. Dependent events | C. Null event |
| B. Independent events | D. Simple event |
- 8) Usher has 12 coins in his pocket. Nine of them are K1 coins and the rest are 10 toea coins. If he picks one from his pocket, what is the chance that he will pick a K1 coin?
- | | |
|--------|--------|
| A. 75% | C. 39% |
| B. 55% | D. 25% |
- 9) Kattie has 15 books inside her bag. Five of them are story books, 3 Mathematics and the rest are Arts books. If she randomly picks one, what is the probability that she will not pick a Mathematics book?
- | | |
|--------|--------|
| A. 20% | C. 60% |
| B. 40% | D. 80% |
- 10) It refers to events that are possible to happen at the same time.
- | | |
|-----------------------|----------------------------------|
| A. Dependent events | C. Mutually exclusive events |
| B. Independent events | D. Non-mutually exclusive events |
-

Solve the following problems.

- 11) A bag contains three green marbles and four black marbles. Two marbles are drawn at random without replacement, If the first marble drawn is green, what is the probability that the second is black?



12. You roll two dice. The first die shows a One and the other rolls under the table and you cannot see it. What is the probability that both dice show One?

13. Two cards are chosen at random without replacement from a pack of 52 playing cards. If the first card chosen is an Ace, what is the probability that the second card chosen is a King?

14. Suppose you simultaneously roll a standard die and spin a spinner that is divided into 10 equal sectors, numbered 1 to 10, What is the probability of getting a 4 on both the die and the spinner?

15. In a library box, there are 8 novels, 8 biographies and 8 war history books. If Joseph selects two books at random, what is the probability of selecting two different kinds of books in a row?

16. In a Mathematics class, 17 are boys and 13 are girls. On a unit test, 4 boys and 5 girls made an A grade. If a student is chosen at random from the class, what is the probability of choosing a girl or an A student?



12.3.3: Statistics

Statistics is the branch of Mathematics which involves collection and organization of information or data.

Imagine dealing with a very large number of data might be prone to miscalculations and errors if they will not be properly handled. In addition, the information collected in a very large number might lose its meaning if they are not properly analyzed and interpreted.

A very good example is **survey**. Usually, an election survey is done to determine the population's bet in an upcoming election. Researchers will have to interview several people as a sample to represent the population in order to arrive at a sound result.

Companies also use statistical analysis for their products. Words such as correlations and inferences are often heard from them. Do not worry, because this topic will explore ways of analyzing large quantities of information and make some predictions or inferences based on sound information. In addition we will explore and plot scatter diagrams and learn to interpret linear and non-linear relationships. Regression and correlation coefficient in the context of the real life problems will also be discussed.

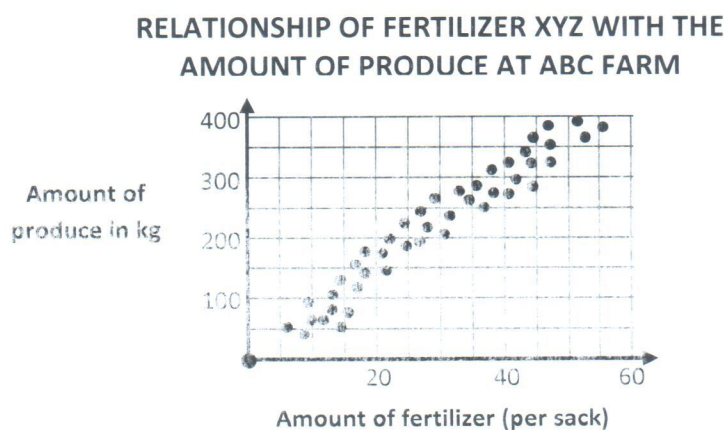
12.3.3.1 The Scatter Diagram

We begin this topic by discussing what a scatter diagram is.

A **scatter diagram** is a tool used for analyzing relationships between two variables. Usually one variable (independent) is plotted on the horizontal axis and the other (dependent) is plotted on the vertical axis. The outline of their intersecting points can show relationship patterns through a graph.

Most often a scatter diagram is used to prove or disprove relationships. It also tells the strength of relationship of two variables.

The figure below is an example of a scatter diagram showing the relationship of the amount of fertilizer used as compared with the number of produce in a farm.





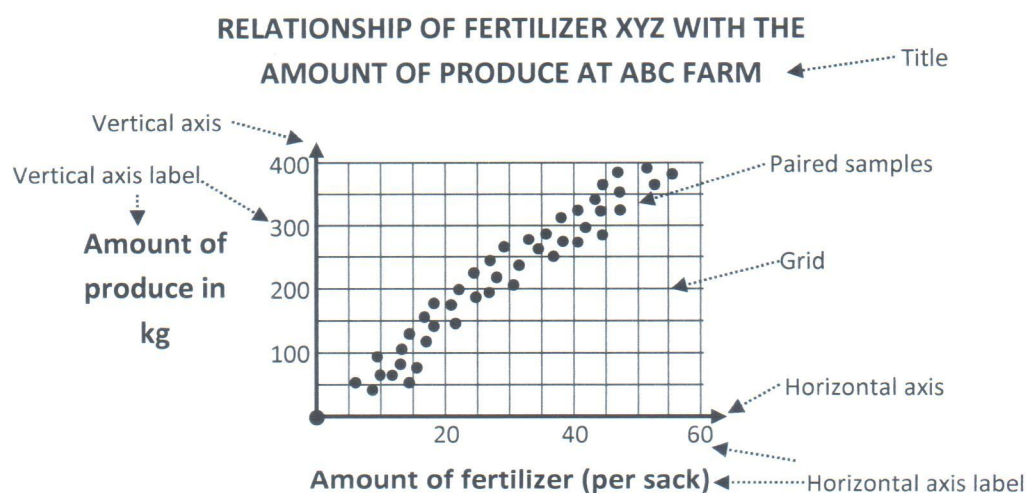
While the diagram shows relationships, it does not by itself prove that one variable causes the other.

Scatter diagrams are often used to examine cause-and-effect relationships and to search for root causes of problem identified. It is also used to design a control system to ensure that gains from quality improvement efforts are maintained.

Important features of a scatter diagram

1. Use horizontal and vertical axes. The use of grid is also helpful although others prefer to use the axes alone.
2. Plot paired samples of data (usually in number) using circles or points.
3. Label the axes in increasing order. Left to right on the horizontal axis and from bottom to top on the vertical axis.
4. Put a title and label the diagram.

Let us identify the parts of a scatter diagram in the sample below.



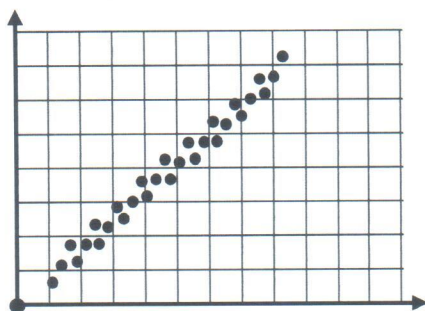
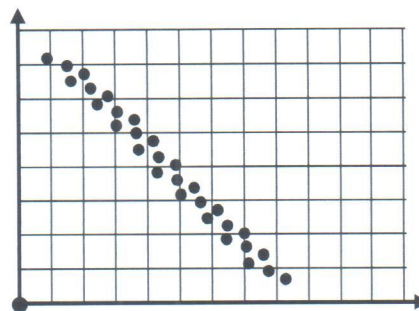
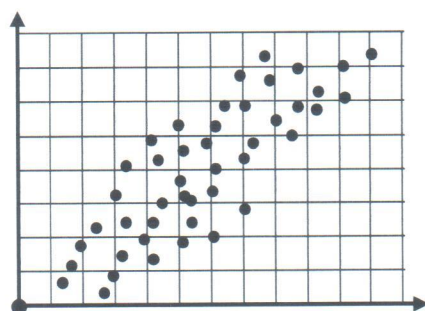
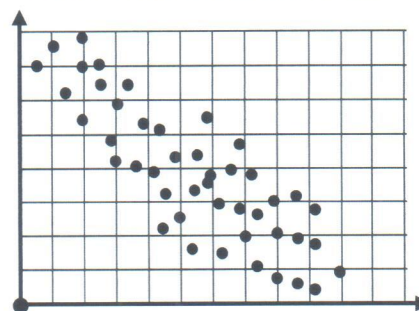
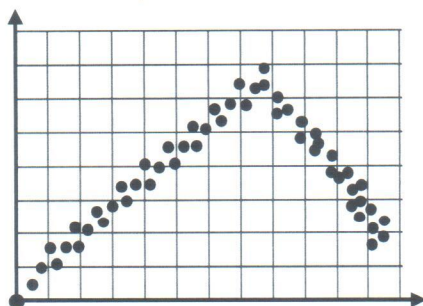
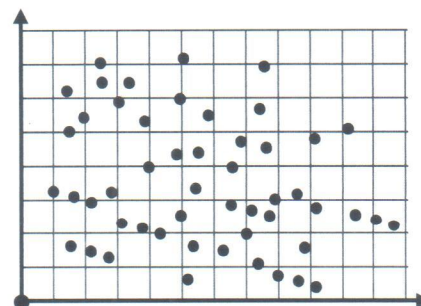
Although the use of graphs is familiar, it is important that each part of the graph must be present in order to produce a more comprehensive one.

Data in the Scatter Diagram can be interpreted using the following:

Strong positive correlation	The value of Y consistently increases as the value of X increases.
Weak positive correlation	The value of Y slightly increases as the value of X increases.
Strong negative correlation	The value of Y consistently decreases as the value of X decreases.
Weak negative correlation	The value of Y slightly decreases as the value of X decreases.
No Correlation	There is no clear pattern formed. The paired samples are scattered.



Examples of scatter diagrams

Strong Positive Correlation**Strong Negative Correlation****Weak Positive Correlation****Weak Negative Correlation****Complex Correlation****No Correlation**

Did you notice the patterns formed by the dots representing the paired samples?

In strong correlations either positive or negative, the paired samples form either increasing or decreasing lines. While correlations may show a slightly same pattern, notice that as they increase or decrease, the paired samples are a little bit more scattered as compared to the strong correlations.

When the diagram shows literally scattered paired samples all over the grids, then it suggests that there is no correlation found among the data. However, complex correlation may arise which needs further and intensive analysis.



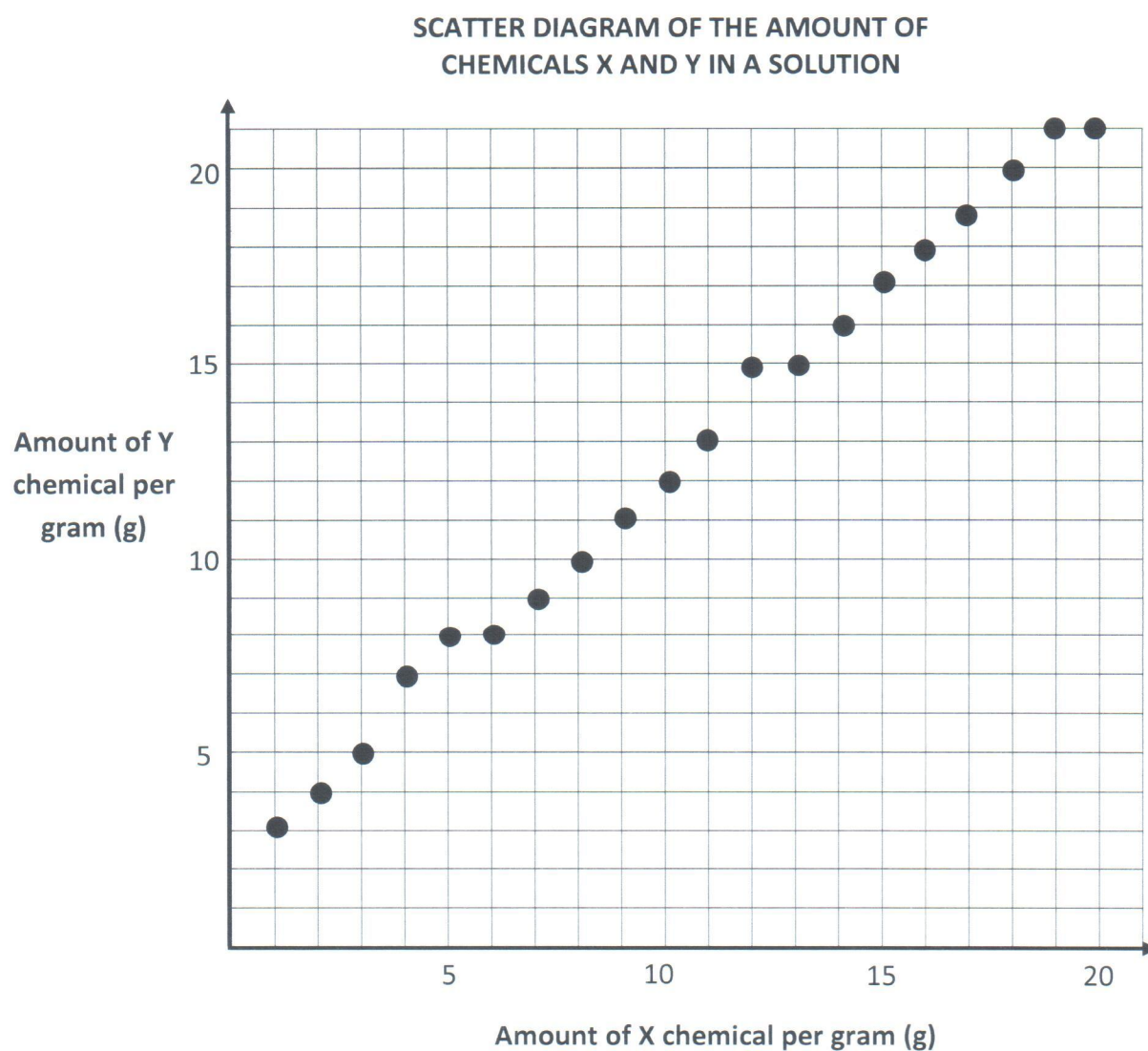
Example 1

Plot a scatter diagram to represent the following set of data. Then identify the correlation shown.

Let x be the amount of X chemical and Y be the amount of Y chemical in a solution

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
y	3	4	5	7	8	8	9	10	11	12	13	15	15	16	17	17	19	20	21	21

Solution:



The scatter diagram shows a **strong positive correlation**.

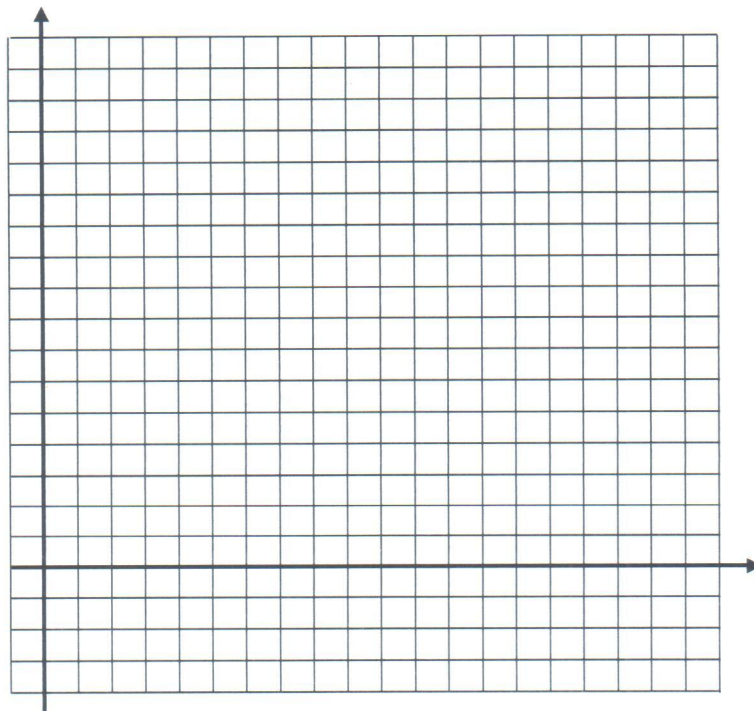
**Learning Activity 12.3.3.1**

20 minutes

Use a scatter diagram to represent the following set of data. Then identify the correlation shown.

- 1) Let x be the amount of chocolate (in grams) and Y be the amount of sugar (in grams) in a Choco bar.

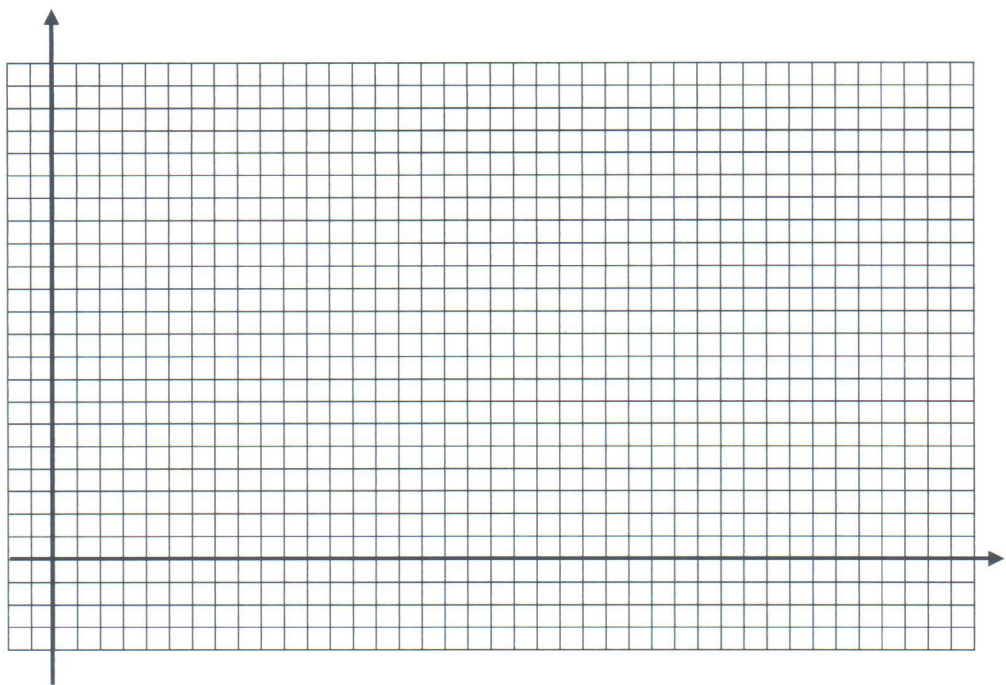
x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
y	-3	-2	-1	0	1	2	3	3	5	6	7	9	9	10	11	11	13	14	14	16





2) Let x be the number of chocolate balls (per packet) and Y be the amount of each packet of chocolate balls (in Kina).

x	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40
y	2	3	5	6	4	7	8	4	3	5	8	9	6	7	12	3	5	6	3	4





12.3.3.2: Analysis Using Correlation

The **linear correlation coefficient** is a measure of the strength and direction of linear relation or association between two quantitative variables. We use the Greek letter (ρ) to represent the population correlation coefficient and (r) to represent the sample correlation coefficient.

The formula for the sample correlation is presented below:

$$r = \frac{\sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)}{n - 1}$$
$$r = \frac{\sum yx - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n} \right) \left(\sum y^2 - \frac{(\sum y)^2}{n} \right)}}$$

Where:

- $\bar{x}$ is the sample mean of the predictor variable
- s_x is the sample standard deviation of the predictor variable
- $\bar{y}$ is the sample mean of the response variable
- s_y is the sample standard deviation of the response variable
- n is the number of individuals in the sample

Properties of the Linear Correlation Coefficient:

1. The linear correlation coefficient is always between -1 and +1, inclusive.
That is, $-1 \leq r \leq +1$.
2. If $r = +1$, there is **perfect positive linear relation** between the two variables.
3. If $r = -1$, there is **perfect negative linear relation** between the two variables.
4. The closer r is to +1, the stronger the positive association between the two variables.
5. The closer r is to -1, the stronger the negative association between the two variables.
6. If r is close to 0, there is evidence of no linear relation between the two variables. Because r is a measure of linear relation, a correlation coefficient close to 0 does not imply no relation, just no linear relation.
7. The linear correlation coefficient is a unit less measure of association. So the units of measure for x and y play no role in the interpretation of r .



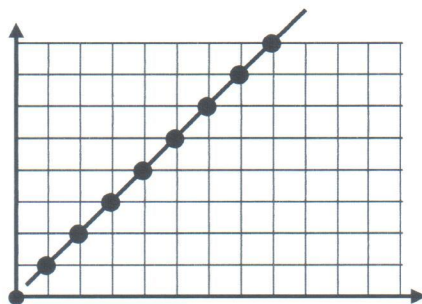
Correlation can be classified into three basic categories.

➤ **Linear Correlations**

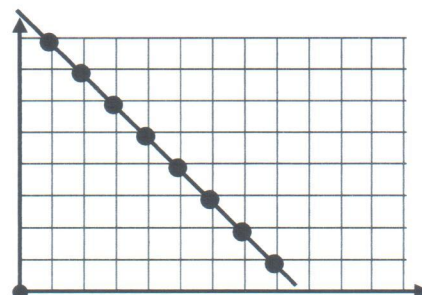
This category of correlation is derived from the term “linear” which means line. Therefore, the correlation can be compared to a straight line.

The following are examples of perfect linear correlations.

Perfect Positive Linear Correlation



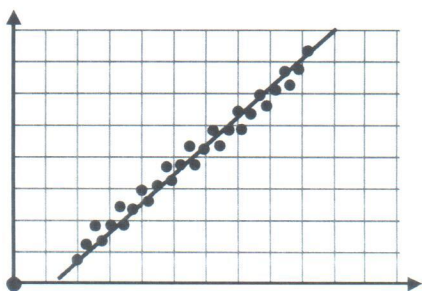
Perfect Negative Linear Correlation



Notice that they are called perfect linear correlations because the paired samples form an almost perfect straight line.

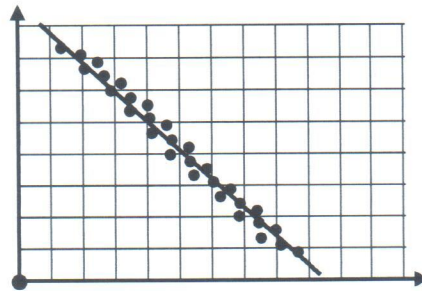
The following are also samples of linear correlations, however they may not seem perfect but they suggest strong correlations

Strong Positive Linear Correlation





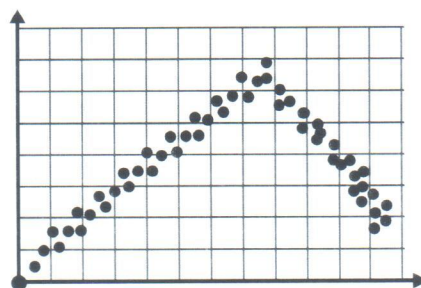
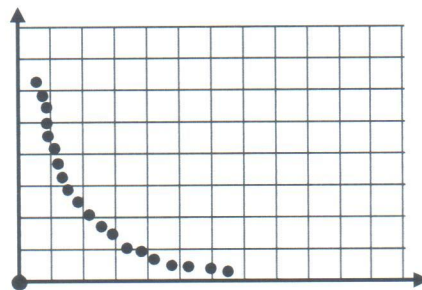
Strong Negative Linear Correlation



Strong linear correlations show that the paired samples are close to the straight line. They show homogeneity in terms of their closeness to the line.

➤ Nonlinear Correlations

This category of correlation is the opposite of linear. Non-linear means “not linear” therefore the correlation exists, however it is not in the form of a line.

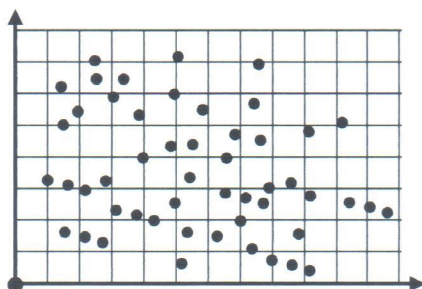


The diagrams show that the variables may be correlated but not through a linear model. That is why both are categorized as non-linear correlations.



➤ **No correlation**

This category suggests that there is no relationship existing between the variables. The following is a sample of no correlation graph.



Notice that there is no pattern formed be it linear or non-linear. The paired samples are scattered and they show no relationship with each other.

It is also important to note that correlation may be strong, moderate, or weak. You can estimate the strength by observing the variation of the points around the line.

- Large variation is weak correlation.
- When the data is distributed quite close to the line the correlation is said to be strong.

The correlation type is independent of the strength. The strength of a linear relationship is measured by the correlation coefficient. The sample correlation coefficient is given the symbol " r ". The population correlation coefficient has the symbol " ρ ".

The sign of the correlation coefficient tells us the direction of the linear relationship.

- If r is negative (<0) the correlation is negative. The line slopes down or it is said to be decreasing.
- If r is positive (>0) the correlation is positive. The line slopes up or it is said to be increasing.

The size (magnitude) of the correlation coefficient tells us the strength of a linear relationship.

- If $|r| > 0.90$ implies a strong linear association.
- For $0.65 < |r| < 0.90$ implies a moderate linear association.
- For $|r| < 0.65$ this is a weak linear association.

Please note that the correlation coefficient only gives us an indication about the strength of a linear relationship. Two variables may have a strong curvilinear relationship, but they could have a "weak" value for r .



Correlation does **NOT** imply causation. Just because two variables are highly correlated does not mean that the explanatory variable “causes” the response variable.. Correlation implies that the two variables are related either in a positive or negative ways, thus the strength of association is the one measured by correlation.

Computing for r

Pearson product moment correlation coefficient or popularly known as the Pearson r is the most commonly used formula to solve for correlation.

The following Pearson r formula is used to compute for r.

$$r = \frac{\frac{1}{n} \sum xy - \bar{x}\bar{y}}{s_x s_y}$$

where: $s_x = \sqrt{\frac{1}{n} \sum x^2 - (\bar{x})^2}$ and $s_y = \sqrt{\frac{1}{n} \sum y^2 - (\bar{y})^2}$

The following criteria are adopted to interpret r:

If r = +.70 or higher	Very strong positive relationship
+.40 to +.69	Strong positive relationship
+.30 to +.39	Moderate positive relationship
+.20 to +.29	Weak positive relationship
+.01 to +.19	No or negligible relationship
-.01 to -.19	No or negligible relationship
-.20 to -.29	Weak negative relationship
-.30 to -.39	Moderate negative relationship
-.40 to -.69	Strong negative relationship
-.70 or higher	Very strong negative relationship

Example 1

The following data are marks obtained by 10 randomly selected Grade 12 students in their Mathematics and Physics Examinations. Determine the level of correlation between the two sets of scores.

	1	2	3	4	5	6	7	8	9	10
Mathematics Exam (30 marks)	20	23	3	29	14	11	11	20	17	17
Physics Exam (40 marks)	30	35	21	33	33	26	22	31	33	36



Solution:

We can use a table to solve for the required entries.

Let x = scores in mathematics exam

y = scores in physics exam

x^2 = square of x score

y^2 = square of y score

xy = product of x and y scores

Below is the summation or total of the entries in each column.

Student	x	x^2	y	y^2	xy
1	20	400	30	900	600
2	23	529	35	1225	805
3	8	64	21	441	168
4	29	841	33	1089	957
5	14	196	33	1089	462
6	11	121	26	676	286
7	11	121	22	484	242
8	20	400	31	961	620
9	17	289	33	1089	561
10	17	289	36	1296	612
Summation	$\Sigma x = 170$	$\Sigma x^2 = 3250$	$\Sigma y = 300$	$\Sigma y^2 = 9250$	$\Sigma xy = 5313$

First Let us identify the given.

$n = 10$ (the number of scores) therefore, $\frac{1}{n} = \frac{1}{10} = 0.1$

$\Sigma xy = 5313$ (found on the last row of the last column)

$\bar{xy} = 17 \times 30 = 510$ the product of the mean of x ($\bar{x}$) and the mean of y ($\bar{y}$)

To get $\bar{x} = 170 \div 10 = 17$

To get $\bar{y} = 300 \div 10 = 30$

$$s_x = \sqrt{\frac{1}{n} \Sigma x^2 - (\bar{x})^2} = \sqrt{\frac{1}{10} (3250) - 17^2} = \sqrt{(0.1)(3250) - 289} = \sqrt{36} = 6$$

$$s_y = \sqrt{\frac{1}{n} \Sigma y^2 - (\bar{y})^2} = \sqrt{\frac{1}{10} (9250) - 30^2} = \sqrt{(0.1)(9250) - 900} = \sqrt{25} = 5$$



Did you notice that what we had in the previous page are just revisions of what you have learnt in your statistic module? Your previous knowledge and mastery of the basic skills will definitely help you use Pearson r more efficiently and accurately.

To solve for r using the Pearson r formula:
$$r = \frac{\frac{1}{n} \sum xy - \bar{x}\bar{y}}{s_x s_y}$$

Substitute all values of the given we identified earlier.

$$r = \frac{(0.1)(5313) - (17)(30)}{(6)(5)}$$

$$r = \frac{(531.3) - (510)}{30}$$

$$r = \frac{21.3}{30}$$

$$r = 0.71$$

Since $r = 0.71$, it indicates a **very strong positive relationship** between the scores of the students in Mathematics and Physics examinations.

This means that the students' scores in Mathematics are related with the scores in Physics. We can further say that this result may suggest that if a student does very well in Mathematics, it is more likely that he will also excel in Physics or vice versa.

Example 2

The following are the recorded ages and glucose levels of patients who went through glucose test. The medical technician would like to determine if the age of their patients are related to their recorded glucose level. Determine the level of correlation between the age and glucose level of patients.

	1	2	3	4	5	6	7	8
Age	22	41	35	58	42	28	38	78
Glucose level	65	99	81	82	70	105	100	97

Solution:

We will use a table to solve for the required entries.

We let x be the age of the patients and let y be their respective glucose level as per record.

x^2 and y^2 are the squares of the scores and xy is the product of the scores.



Below is the summation or total of the entries in each column.

Student	x	x ²	y	y ²	xy
1	22	484	65	4225	1430
2	41	1681	99	9801	4059
3	35	1225	81	6561	2835
4	58	3364	82	6724	4756
5	42	1764	70	4900	2940
6	28	784	105	11025	2940
7	38	1444	100	10000	3800
8	78	6084	97	9409	7566
Summation	$\Sigma x = 342$	$\Sigma x^2 = 16830$	$\Sigma y = 699$	$\Sigma y^2 = 62645$	$\Sigma xy = 30326$

First, let us identify the given.

$n = 8$ (the number of scores) therefore, $\frac{1}{n} = \frac{1}{8} = 0.125$

$\Sigma xy = 30\,326$ (found on the last row of the last column)

$\bar{xy} = 42.75 \times 87.38 = 3\,735.50$ the product of the mean of x ($\bar{x}$) and the mean of y ($\bar{y}$)

To get $\bar{x} = 342 \div 8 = 42.75$

To get $\bar{y} = 699 \div 8 = 87.38$

$$s_x = \sqrt{\frac{1}{n} \Sigma x^2 - (\bar{x})^2} = \sqrt{\frac{1}{8}(16830) - 42.75^2} = \sqrt{(0.125)(16830) - 1827.5} = \sqrt{276.19} = \mathbf{16.62}$$

$$s_y = \sqrt{\frac{1}{n} \Sigma y^2 - (\bar{y})^2} = \sqrt{\frac{1}{8}(62645) - 87.38^2} = \sqrt{(0.125)(62645) - 7635.26} = \sqrt{195.37} = \mathbf{13.98}$$

$$\text{To solve for } r \text{ using the Pearson } r \text{ formula: } r = \frac{\frac{1}{n} \Sigma xy - \bar{xy}}{s_x s_y}$$

Substitute all values of the given we identified earlier.

$$r = \frac{(0.125)(30326) - (42.75)(87.38)}{(16.62)(13.98)}$$

$$r = \frac{(3790.75) - (3735.50)}{232.35}$$

$$r = \frac{55.25}{232.35}$$

$$\mathbf{r = 0.24}$$

Referring back on the criteria adopted to interpret r on page 77, the computed $r = \mathbf{0.24}$ indicates a **weak positive relationship**.



This means that the glucose level of patients does not depend on their ages. It can further be elaborated through this test that a younger person does not necessarily mean that he is not prone to having a high glucose level. Diabetes is one of the most deadly diseases nowadays that cannot be correlated with age, therefore, we should really watch our sugar intake and make our lifestyle active and healthy.

As you noticed in the examples presented, the closer the value of r to 1, the stronger the relationship is ; on the other hand, the closer it is to zero, the weaker the relationship is.



Learning Activity 12.3.3.2

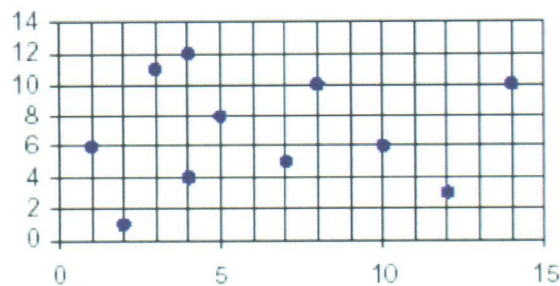


10 minutes

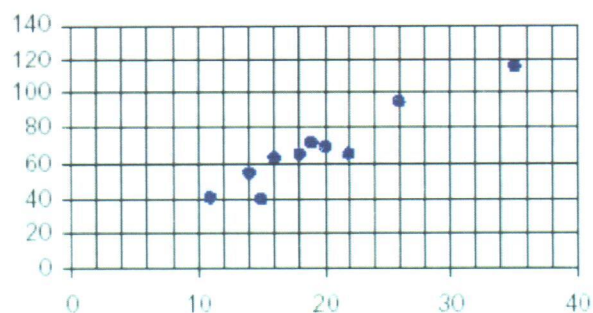
The following data shows the alcoholic consumption as compared to the death rates due to liver disease in the following hospitals.

- 1) Identify the correlation in each graph below. Label each graph as positive, negative or no correlation.

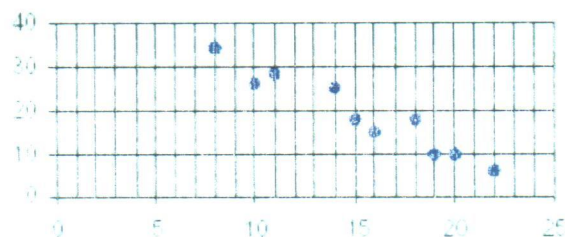
a)



b)

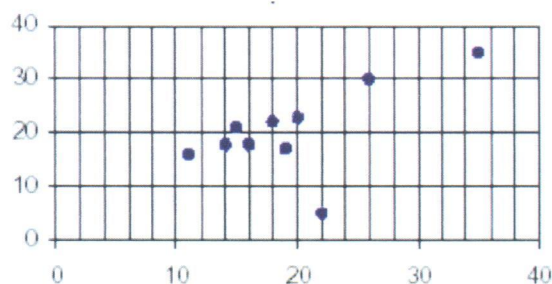


c)





d)



2) Determine the level of correlation between the two given sets of data using Pearson r .

Hospital	A	B	C	D	E	F	G	H	I	J
Alcohol consumption (X)	6.8	5.3	8.2	3.1	3.5	4.2	3.9	10.3	24.5	10.8
Percentage of Death Rate (Y)	3.5	3.1	6.8	1.1	1.01	3.9	2.05	8.6	22	8

a) Complete the table.

Hospital	x	x^2	y	y^2	xy
A					
B					
C					
D					
E					
F					
G					
H					
I					
J					
Summation	$\Sigma x =$	$\Sigma x^2 =$	$\Sigma y =$	$\Sigma y^2 =$	$\Sigma xy =$

b) Identify the following:

- $n =$ _____
- $\frac{1}{n} =$ _____
- $\Sigma xy =$ _____
- $\overline{xy} =$ _____
- $\overline{x} =$ _____
- $\overline{y} =$ _____



- c. Calculate the sample standard deviation (S_x) of the predictable variable.
- d. Calculate the sample standard deviation (S_y) of the response variable.
- e. Using the Pearson r formula, work out and interpret the correlation coefficient (r).

$r =$ _____

Interpretation: _____



12.3.3.3: Regression

Regression analysis is the art and science of fitting straight lines to patterns of data. **Regression** means finding a linear model (using a straight line) that best fits the set of data. This regression line provides a relationship that can be described algebraically which allows us to make predictions, forecasts or estimates.

The **art of regression modeling** is to collect data that is relevant and informative with respect to your decision or inference problem, and then define your variables and construct your model in such a way that the assumptions are plausible.

Choosing a good regression model requires the following:

- gathering useful data and making sure you know where it came from and how it was measured,
- performing descriptive analysis on it to understand its general patterns and to spot data-quality problems,
- applying appropriate data transformations if you see strong evidence of relationships that are nonlinear,
- fitting and refining and comparing models,
- checking to see whether a given model's assumptions are reasonably well satisfied or whether an alternative model is suggested,
- choosing among reasonable models based on the appropriate bottom-line accuracy measure, and
- deriving some useful insights from the whole process.

A regression model does not assume that **Y** variable merely “depends in some way” on the **X** variable.

If you have a variable **Y** that you wish to predict, and you have some other variables X_1 , X_2 , that you believe have “some sort of effect” on **Y** or “some sort of predictive value” with respect to future values of **Y**, this does NOT suffice to justify using a linear regression model to predict **Y** from the **X**'s.

The regression model makes very strong assumptions about the way in which **Y** depends on the **X**'s, namely that the causal or predictive effects of the **X**'s with respect to **Y** are linear and additive and non-interactive and that any variations in **Y** that are not explained by the **X**'s are statistically independent of each other and identically normally distributed under all conditions.

There are plenty of models to be used for regression analysis. Some of them are really sophisticated and require the use of software and other statistical tools to be used for analysis.

A simple way of doing regression analysis is through the use of the line of best fit.



This method may look very familiar to you because of your background knowledge about the line as discussed in your algebra related modules.

12.3.3.3.1: Line of Best Fit

The **line of best fit** is a regression line drawn to describe the trend of the data in reference to it. The equation of this line can be used to predict data values.

The following steps are used to analyze data using the line of best fit:

1. Plot the points in a coordinate plane. The scatter plot shows a negative relationship.
2. Draw a line that is close to the data points. Try to have as many points above the line as below it.
3. Using two points on the line, identify the slope of the line using rise over run or the

$$\text{formula for deriving the slope (m) given as } m = \frac{y_2 - y_1}{x_2 - x_1} \text{ or } \frac{y_1 - y_2}{x_1 - x_2}$$

4. Identify the equation of the line to be used for further analysis and prediction using the slope intercept form $y = mx + b$.

Example 1

The table shows the weekly sales of a DVD and the number of weeks since its release.

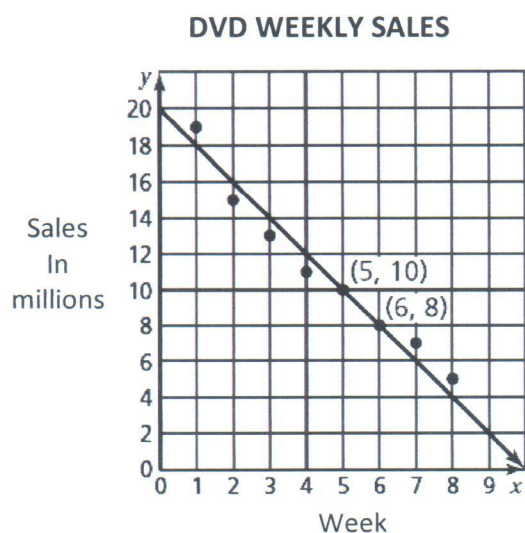
- (a) Make a scatter plot of the data.
- (b) Draw a line of best fit.
- (c) Write an equation of the line of best fit.
- (d) Predict the sales in week 9.

DVD WEEKLY SALES

Week (x)	1	2	3	4	5	6	7	8
Sales in millions (y)	K19	K15	K13	K11	K10	K8	K7	K5

Solution:

- a) Plot the points in a coordinate plane. This is the same way we plotted and formed the scatter diagram.
- b) Draw a line close to the data points. As much as possible, the lines below and above it are approximately equal. The line drawn below is the line of best fit.



Notice that the trend of the line of best fit is decreasing. Therefore it suggests that the slope is negative.

Does the idea of a slope of a line sound familiar?

Great! These were already discussed in your lower mathematics units.

- c) Before we will be able to make an equation of the line, we have to identify two points on the line first. Notice that the line passes through the points (5, 10) and (6, 8). This will help us to get the slope of the line.

In your previous Mathematics lessons, we have defined slope as the change in y over the change in x or in symbols: $\text{slope} = \frac{y_2 - y_1}{x_2 - x_1}$

It does not matter which Point is Point 1 or 2. Anyway, the same result will come out.

Suppose $P_1 (5, 10)$ and $P_2 (6, 8)$ $\text{slope} = \frac{8 - 10}{6 - 5} = \frac{-2}{1}$ or -2.

Now another important point we have to consider is y-intercept. Notice that the line passes through (0, 20) therefore, we can say that the y-intercept is 20.

Recall that if the slope and the y-intercepts were already determined, we can solve for the equation of the line using the slope intercept form represented by $y = mx + b$, where m is the slope and b is the y-intercept.

In this case $m = -2$ and $b = 20$. To find the equation of the line, we will substitute the values in the formula $y = mx + b$, then we have **$y = -2x + 20$** .

- d) The equation of the line $y = -2x + 20$ will help us predict the sales in week 9. Since the weeks were represented by x variable, we let $x = 9$ and substitute it in the formula.

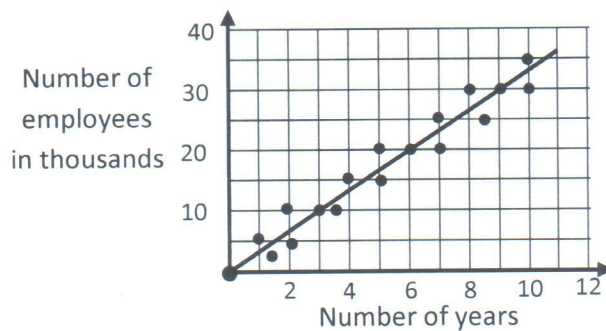


Since $y = -2x + 20$ and $x = 9$
We have, $y = -2(9) + 20$
Simplify: $y = -18 + 20$
 $= 2$

Therefore, since the trend in sales is decreasing, by week 9 the predicted sales are K2 million.

Example 2

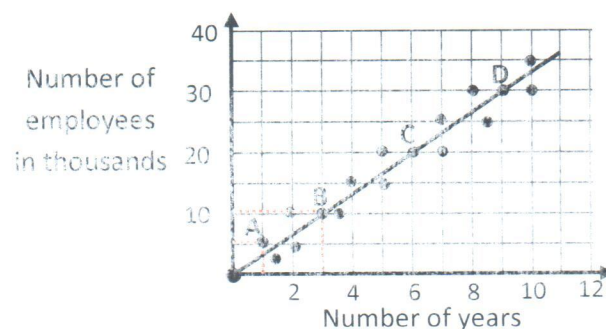
The graph below shows the number of years (x) a franchising company has been running and the number of employees (Y) they hire as their franchising business expands in the country.



- About how many employees does the company have during its first year?
- How many years did it take for the company to have 10 000 employees?
- Using the line of best fit, determine the following:
 - The slope of the line of best fit.
 - The equation of the line of best fit.
 - Estimate the number of employees the company will have on their 15th year.
 - How many years will it take for the company to have 100 000 employees?

Solution:

Analyzing the graph and labeling some points like A, B, C and D will help us facilitate a clear discussion.





- a) To determine the number of employees the company has during its first year, refer to A which is the intersection of Year 1 and 5000, or we can write it in ordered pair as (1, 5000).

Therefore, the company had 5 000 employees during its first year of franchise operation.

- b) To determine the number of years it took the company to have 10 000 employees, refer to B. Notice that B is the intersection of year 3 and 10,000 or we can write it in ordered pair as (3, 10 000).

Therefore, it only took the company 3 years of franchise operation to have 10 000 employees.

If you are in a business, it is really good to know what the company can become in the future to make further decisions.

- c) Using the line of best fit we can determine the following:

- i. Slope of the line.

There are different ways we can do this. Looking at the graph, use two points that are on the line. Did you notice that there are three points lying on the line?

Good if you saw points B(3, 10 000), C(6, 20 000) and D (9, 30 000). You may use any two points.

Suppose you want to use B and C, the slope (m) can now be solved using the

formula $m = \frac{y_2 - y_1}{x_2 - x_1}$. Substitute the values of x and y in the formula we have:

$$m = \frac{20\,000 - 10\,000}{6 - 3} = \frac{10\,000}{3}.$$

Suppose you want to use C and D, the slope (m) can now be solved using the

formula: $m = \frac{y_2 - y_1}{x_2 - x_1}$. Substitute the values of x and y we have:

$$m = \frac{30\,000 - 20\,000}{9 - 6} = \frac{10\,000}{3}.$$

Since all these points lie on the same line, we will get the same slope $\frac{10\,000}{3}$ no matter which pair from the three points is used.



- ii. The equation of the line of best fit

Using the slope (gradient) $m = \frac{10\,000}{3}$ and a point where the line passes through the x axis (0, 0) where $b = 0$, we will solve for the equation of the line using the formula $y = mx + b$.

$$y = \frac{10\,000}{3}x + 0$$

- iii. To make an estimate of the number of employees the company will have on their 15th year, we will use the equation of the line derived on ii.

$$\text{Which is } y = \frac{10\,000}{3}x + 0.$$

Since the number of year (x) is 15, we substitute x for 15 in the equation.

$$\text{So, } y = \frac{10\,000}{3}(15) + 0. \text{ Simplify the equation and the value of } y = 50\,000.$$

Therefore, the estimated number of employees in the company after 15 years will be 50 000.

- iv. To solve for the number of years, it will take for the company to have 100 000 employees, we will use the same equation of the line.

$$y = \frac{10\,000}{3}x + 0.$$

Since the number of employees (y) is 100,000, we will substitute y for 100 000 in the equation.

$$100\,000 = \frac{10\,000}{3}x + 0$$

$$100\,000 = \frac{10\,000}{3}x$$

$$3(100\,000) = \cancel{3} \left(\frac{10\,000}{\cancel{3}} \right) x$$

$$300\,000 = 10\,000x$$

$$\frac{300\,000}{10\,000} = \frac{10\,000}{10\,000}x$$

$$30 = x$$

multiply both sides by 3

cancel 3 on the right side

divide both sides by 10 000

Therefore, it will take 30 years for the company to have 100 000 employees.



12.3.3.3.2 Interpolation and Extrapolation

Interpolation and extrapolation are both used to estimate imaginary or assumed values for a variable (either x or y) based on observations of other data. There are different methods of interpolation and extrapolation based on the complete trend of data observed.

To emphasize the difference between extrapolation and interpolation, we will look at their prefixes. The prefix “extra” means “outside” or “in addition to” while the prefix “inter” means “in between” or “among”.

Interpolation is all about estimating data points inside or within the range of the data observed while **extrapolation** is all about estimating data points outside the range of the data.

By just remembering their meanings will guide you through as to what kind of estimate you are performing.

For both methods we do some assumptions or predictions. But before we can interpolate and extrapolate, we should have identified our variables (usually x and y) and plot the paired samples in a diagram. The diagram serves as our model for our data. The diagram may also show the line of best fit which we have discussed previously.

But notice in our example 2 where we are able to predict the increase in the number of employees, our estimation is solely based on the trend of the line. Have you ever thought that the company will have a problem and along the way somewhere after 12 years, the number of employees may begin to decrease? This is where extrapolation comes in.

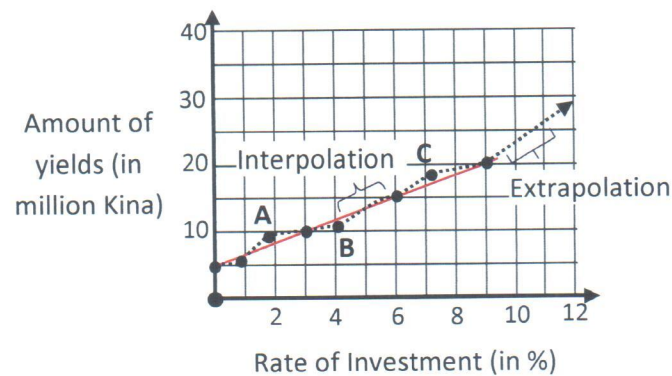
The goal of interpolation and extrapolation is to use the model for prediction. More specifically, given an independent variable, what will the predicted value of the corresponding dependent variable be? The value that we enter for our independent variable will determine whether we are working with extrapolation or interpolation.

A function (linear, quadratic or polynomial) can be used to predict the value of the dependent variable for an independent variable that is in the midst or within the given set of data. In this case interpolation is being employed.

Example 3

A multi-million company records the rate of investments (x) they engage with in comparison with the amount of yields (y) they earn investing in different real estates. It was found that they invested on real estates which offer a rate below 10%.

The record shows the following graph on the next page.



- a) Make an estimate amount of yield if the company invests on an estate which gives 5% yield?
- b) If the company invests on a 12% rate, how much yield will they get?

Solution:

In this case our data with x is in between 0 and 9 which produce a regression line represented by the straight line.

Now using the points A (3%, K10 million) and B (6%, 15 million), we solve for the slope.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{15 - 10}{6 - 3} = \frac{5}{3}$$

Since the line of best fit passes through the point (0, 5), then $b = 5$ we can now use the equation of the line $y = \frac{5}{3}x + 5$.

- a) To make an estimate amount of yield when the company invests on an estate which gives 5% yield, we will simply use the equation $y = \frac{5}{3}x + 5$ and replace x with 5 to predict yield on 5% investment.

$$y = \frac{5}{3}(5) + 5 = 13.33.$$

Therefore, if the company invests on real estate offering 5% rate, it will earn an estimate of K13.33 million.

Since the value of x (5%) is among the range of values used to make the line of best fit, this is an example of interpolation.



- b) To solve for the amount of yield the company will get upon investing on a 12% rate, we will use the equation $y = \frac{5}{3}x + 5$ and replace x with 12 to predict yield on 12% investment.

$$y = \frac{5}{3}(12) + 5 = 25.$$

Therefore, if the company invests on real estate offering 12% rate, it will earn an estimate of K25 million.

Since the value of x (12%) is not among the range or it is outside the range of values used to make the line of best fit, this is an example of extrapolation.

Note:

Of the two methods, interpolation is preferred. This is because we have a greater likelihood of obtaining a valid estimate. Extrapolation is generally less accurate than interpolation because when we use extrapolation, we are making the assumption that our observed trend continues for values of x outside the range we used to form our model. This may not be the case, and so we must be very careful when using extrapolation techniques.

**Learning Activity 12.3.3.3**

40 minutes

- 1) a) Graph a scatter plot of the data relating time and distance using the data on the table below. Label the graph appropriately. Then draw the line of best fit.

BIKING HOME

Hours (x)	1	2	3	4	5	6	7	8
Miles (y)	35	29	26	20	16	9	6	0

Graph



b) Identify:

i. the slope

ii. the equation of the line of best fit.

- 2) The table below shows the predicted annual cost for a middle income family in Australia to raise a child from birth until adulthood. Draw a scatter plot and describe what relationship exists within the data.

COST OF RAISING A CHILD BORN IN 2003					
Child's Age	3	6	9	12	15
AnnualCost (\$)	10 700	11 700	12 600	15 000	16 700

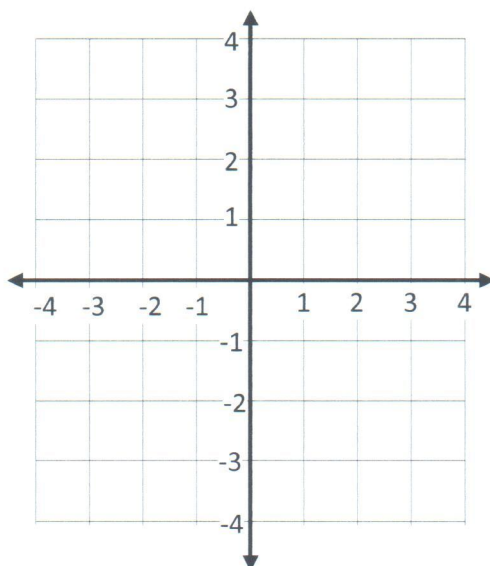
NOW DO THE SUMMATIVE TASK 3

**SUMMATIVE TASK 12.3.3**

30 minutes

- 1) Make a scatter plot of the data in the table. Draw a line of best fit. What is the equation of the line of best fit?

x	-2	-2	-1	0	1	1	1	2	2	3
y	2	3	2	1	0	1	-1	-1	-2	-2



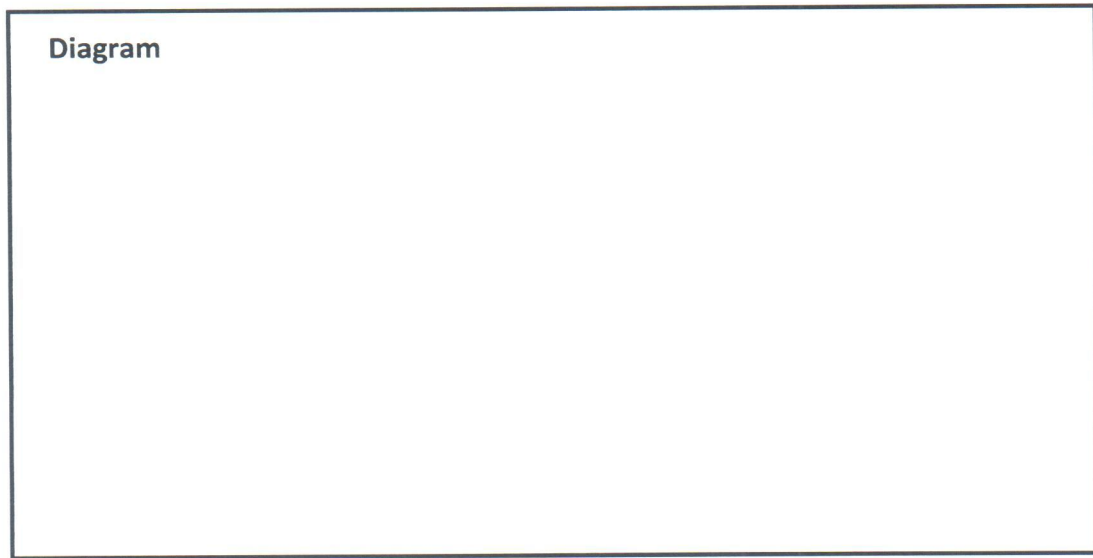
- 2) The table below gives the number of hours spent studying for an exam and the final exam grade.

Study Hours	3	2	5	1	0	4	3
Grade	84	77	92	70	60	90	75

- Draw a scatter plot of the data and draw the line of best fit.
- What is the equation for the line of best fit?
- Predict the grade of a student who studied for 6 hours.
- Could this line go on forever? Why or why not?



a)



b)

c)

d)



UNIT SUMMARY

This summary outlines the key ideas and concepts to be remembered.

- Counting Techniques are methods which help in “counting” and organizing large quantities of numbers. Common techniques include the listing method, the tree diagram, and the use of the fundamental principles in counting.
- The Fundamental Counting Principle. If there are A ways for one event to occur, and B ways for another event to occur, then there are $A \times B$ ways for both to occur. This principle applies to any fixed number of event denoted by $A \times B \times C \dots$ where A, B, C... represents fixed, certain or exact values.
- Permutation is an ordered arrangement of items that occurs when no item is used more than once and the order of arrangement makes a difference.
- The number of permutations of distinct objects (n) is $n!$
- Circular permutation refers to the arrangement of objects in a circular or oval form. The circular permutation of distinct objects (n) is $(n-1)!$
- Permutations of duplicate items involve arrangement of objects with identical or duplicate objects wherein the movements of these duplicated objects do not create another set of arrangement. It is solved using the formula $\frac{n!}{n_1! \cdot n_2! \cdot n_3! \dots}$ where n_1 items are identical, n_2 items are identical, n_3 items are identical, and so on.
- Combination is a counting technique wherein the number of groupings and not order of the objects is being counted. The order or arrangement of objects is not important or does not really matter and it can be solved using the formula ${}_nC_r = \frac{n!}{r!(n-r)!}$.
- Probability is used to model certainty or the “chance” that something will happen. It is the measure of the chance or likelihood that an event will occur or happen.
- A sample space is the set of all possible outcomes of an experiment.
- An event is a collection of outcomes having common characteristics from the sample space. It is a subset of the sample space wherein its elements were taken from the sample space. A simple event consists of exactly one outcome and a compound event consists of more than one outcomes.
- We calculate the Probability of Event as $P(E) = \frac{E}{n}$ where: E is the number of ways E can occur and n is the total number of outcomes.
- Events are independent when the outcome of one event does not influence the outcome of the next event. On the other hand, two or more events are dependent if the outcomes of the first, affects the others.
- If A and B are independent events, then the probability that both A and B occur is $P(A \text{ and } B) = P(A) \cdot P(B)$. This formula means that for two independent events, the probability that both events occur is the product of the probabilities of the events. The formula can be used for more than two independent events. Just multiply the probability of each event.



- If A and B are dependent events, then the probability that both A and B occur is $P(A \text{ and } B) = P(A) P(B \text{ given } A)$. This formula means that for two dependent events, the probability that both events occur is the product of the probability that the first event occur and the probability that the second event occurs given that the first event has occurred. The formula can be used for more than two dependent events. Just always consider the occurrence of the previous events.
- For two events A and B, the probability that A will occur given that B has already occurred is the conditional probability of A given B written as $P(A/B)$ where: $P(A/B) = \frac{P(A \cap B)}{P(B)}$.
- Events are mutually exclusive if they have no outcomes in common. This is the same as saying that these two events cannot happen together at the same time. It is impossible that both could occur in a single trial of an experiment. Their intersection is empty. On the other hand, two events that can happen at the same time are called non-mutually exclusive events.
- If events A and B are mutually exclusive, then $P(A \text{ or } B) = P(A) + P(B)$. With more formal notation, it also mean $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$.
- Use the addition of probabilities when dealing with union of mutually exclusive events. The key word that you will have to remember is “or”.
- Use the multiplication of probabilities when dealing with intersection of non-mutually exclusive events. The key word that you will have to remember is “and”. That for any events A and B, $P(A \text{ and } B)$ or $P(A \cap B) = P(B/A) P(A)$ or $P(A/B) P(B)$.
- A scatter diagram is a tool used for analysing relationships between two variables. Usually one variable is plotted on the horizontal axis and the other is plotted on the vertical axis. The outline of their intersecting points can show relationship patterns through a graph.
- Properties of the Linear Correlation Coefficient:
 - The linear correlation coefficient is always between -1 and +1, inclusive. That is, $-1 < r < +1$
 - If $r = +1$, there is perfect positive linear relation between the two variables.
 - If $r = -1$, there is perfect negative linear relation between the two variables.
 - The closer r is to +1, the stronger the positive association between the two variables.
 - The closer r is to -1, the stronger the negative association between the two variables.
 - If r is close to 0, there is evidence of no linear relation between the two variables. Because r is a measure of linear relation, a correlation coefficient close to 0 does not imply no relation, just no linear relation.
 - The linear correlation coefficient is a unit less measure of association. So the units of measure for x and y play no role in the interpretation of r.
- Linear correlation is derived from the term “linear” which means line. Therefore, the correlation can be compared to a straight line.
- Perfect linear correlation of paired samples forms an almost perfect straight line.



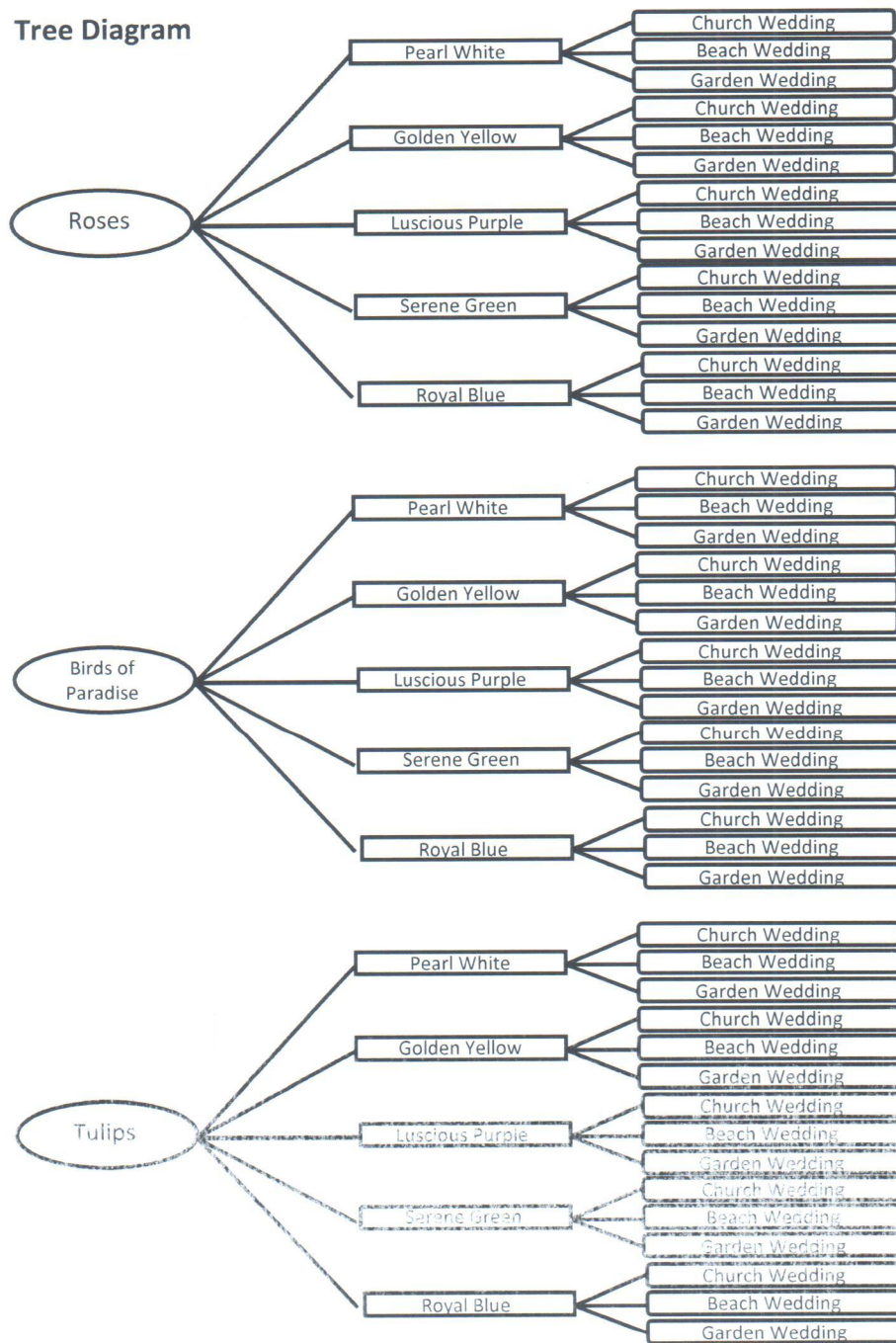
- Non-Linear correlation is the opposite of linear. Non-linear means “not linear” therefore the correlation exists, however it is not in the form of a line.
 - Pearson product moment correlation coefficient or popularly known as the Pearson r is the most commonly used formula to solve for correlation.
 - The closer the value of r to 1, the stronger the relationship is; on the other hand, the closer it is to zero, the weaker the relationship is.
 - Regression means finding a linear model (using a straight line) that best fits the set of data.
 - The line of best fit is a regression line drawn to describe the trend of the data in reference to it. The equation of this line can be used to predict data values.
 - Interpolation is all about estimating data points inside or within the range of the data observed while extrapolation is all about estimating data points outside the range of the data.
-

REVISE SECTIONS 12.3.1 TO 12.3.3 THEN DO MODULE ASSESSMENT 3

**ANSWERS TO LEARNING ACTIVITIES****Learning Activity 12.3.1.1**

- 1) List: let H = heads T = Tails
Outcomes: HHH, TTT, HTH, HHT, THH, THT, HTT, TTH
Total number of outcomes: 8

2)

Tree Diagram



Wedding Package	Description / Inclusion
1	Roses in a Pearl White Motif Church Wedding
2	Roses in a Pearl White Motif Beach Wedding
3	Roses in a Pearl White Motif Garden Wedding
4	Roses in a Golden Yellow Motif Church Wedding
5	Roses in a Golden Yellow Motif Beach Wedding
6	Roses in a Golden Yellow Motif Garden Wedding
7	Roses in a Luscious Purple Motif Church Wedding
8	Roses in a Luscious Purple Motif Beach Wedding
9	Roses in a Luscious Purple Motif Garden Wedding
10	Roses in a Serene Green Motif Church Wedding
11	Roses in a Serene Green Motif Beach Wedding
12	Roses in a Serene Green Motif Garden Wedding
13	Roses in a Royal Blue Motif Church Wedding
14	Roses in a Royal Blue Motif Beach Wedding
15	Roses in a Royal Blue Motif Garden Wedding
16	Birds of Paradise in a Pearl White Motif Church Wedding
17	Birds of Paradise in a Pearl White Motif Beach Wedding
18	Birds of Paradise in a Pearl White Motif Garden Wedding
19	Birds of Paradise in a Golden Yellow Motif Church Wedding
20	Birds of Paradise in a Golden Yellow Motif Beach Wedding
21	Birds of Paradise in a Golden Yellow Motif Garden Wedding
22	Birds of Paradise in a Luscious Purple Motif Church Wedding
23	Birds of Paradise in a Luscious Purple Motif Beach Wedding
24	Birds of Paradise in a Luscious Purple Motif Garden Wedding
25	Birds of Paradise in a Serene Green Motif Church Wedding
26	Birds of Paradise in a Serene Green Motif Beach Wedding
27	Birds of Paradise in a Serene Green Motif Garden Wedding
28	Birds of Paradise in a Royal Blue Motif Church Wedding
29	Birds of Paradise in a Royal Blue Motif Beach Wedding
30	Birds of Paradise in a Royal Blue Motif Garden Wedding
31	Tulips in a Pearl White Motif Church Wedding
32	Tulips in a Pearl White Motif Beach Wedding
33	Tulips in a Pearl White Motif Garden Wedding
34	Tulips in a Golden Yellow Motif Church Wedding
35	Tulips in a Golden Yellow Motif Beach Wedding
36	Tulips in a Golden Yellow Motif Garden Wedding
37	Tulips in a Luscious Purple Motif Church Wedding
38	Tulips in a Luscious Purple Motif Beach Wedding
39	Tulips in a Luscious Purple Motif Garden Wedding



40	Tulips in a Serene Green Motif Church Wedding
41	Tulips in a Serene Green Motif Beach Wedding
42	Tulips in a Serene Green Motif Garden Wedding
43	Tulips in a Royal Blue Motif Church Wedding
44	Tulips in a Royal Blue Motif Beach Wedding
45	Tulips in a Royal Blue Motif Garden Wedding

- 3) 18 matches
4) 5184 outcomes

Learning Activity 12.3.1.2

- 1) 5040 ways
2) 120 ways
3) 3 628 800 ways
4) 1260 ways
5) a) $13! = 6,227,020,800$ b) $3!5!4!1!4! = 414,720$
6) a) $12! = 79,001,600$ b) $5!7!2! = 1,209,600$ c) $3!9! = 2,177,280$
7) $(10-1)! = 9! = 362,880$
8) ${}_8P_3 = 336$

Learning Activity 12.3.1.3

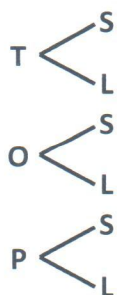
- 1) a) 10
b) 1,287
c) 3,003
d) 220
e) 153
f) 53,130
g) 140
h) 60
i) 14
j) 71
- 2) a) 3,003
b) i. 1,890,00 ii. 14,700 iii. 118,755

**SUMMATIVE TASK 12.3.1**

- 1)
 - a) 40 320
 - b) 6
 - c) 2880
 - d) 3, 603,600
 - e) 720
 - f) 3060
 - g) 116, 280
 - i) 200
 - j) 30
- 2) 24
- 3) 6,096,454
- 4) 2,560,000

Learning Activity 12.3.2.1

- 1) $S = \{ H, T \}$
- 2) $S = \{ H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6 \}$
- 3) $S = \{ A, B, C, D, E, F, G, H \}$
- 4) $S = \{ BM, BM, GM, GM, GM, YM, YM, YM, YM \}$
- 5) a) Tree Diagram



- b) $S = \{ TS, TL, OS, OL, PS, PL \}$

Learning Activity 12.3.2.2

- 1)
 - a) $\frac{3}{7} = 42.86\%$
 - b) $\frac{4}{7} = 57.14\%$
- 2)
 - a) $\frac{25}{50} = 50\%$
 - b) $\frac{10}{50} = 20\%$
 - c) $\frac{15}{50} = 30\%$
- 3) $\frac{4}{12} = 33.33\%$ or $33\frac{1}{3}\%$



4) $\frac{7}{10} = 70\%$

5) a) $\frac{2}{5} = 40\%$ b) $\frac{7}{20} = 35\%$ c) $\frac{5}{20} = 25\%$

Learning Activity 12.3.2.3

- 1) a) Independent events
b) Independent events
c) Dependent events
d) Dependent events
e) Dependent events
f) Dependent events
g) Independent events
h) Independent events
i) Dependent events

2) a) $\frac{200}{500} = 40\%$ b) $\frac{210}{510} = 42\%$

3) $\frac{5}{12} = 41.67\%$

Learning Activity 12.3.2.4

- 1) $\frac{1}{2}$
2) a. 0.059 b. 0.441 c. 0.128
3) a. $\frac{3}{5}$ b. $\frac{2}{5}$ c. $\frac{1}{2}$
4) $\frac{1}{6}$ 5) $\frac{1}{13}$ 6) $\frac{3}{51}$

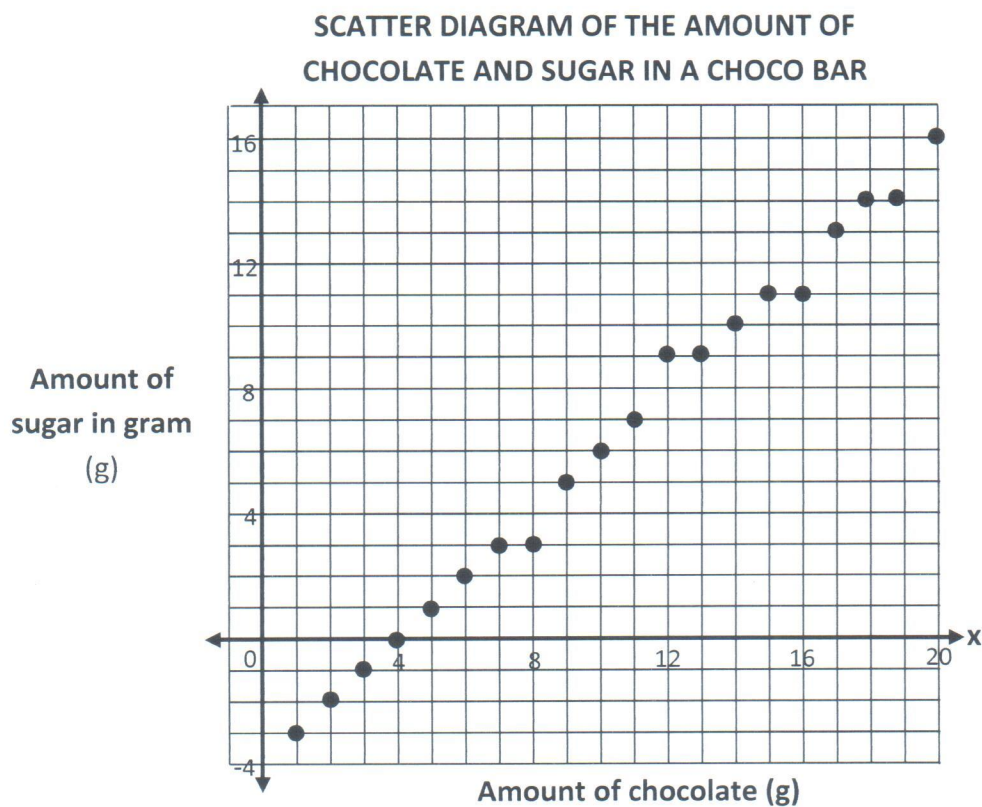
SUMMATIVE TASK 12.3.2

- | | | | |
|------|-------|--------------------|---------------------|
| 1) D | 6) C | 11. $\frac{2}{3}$ | 14. $\frac{1}{60}$ |
| 2) A | 7) D | | |
| 3) C | 8) A | 12. $\frac{1}{6}$ | 15. $\frac{16}{23}$ |
| 4) B | 9) D | | |
| 5) C | 10) D | 13. $\frac{4}{51}$ | 16. $\frac{17}{30}$ |

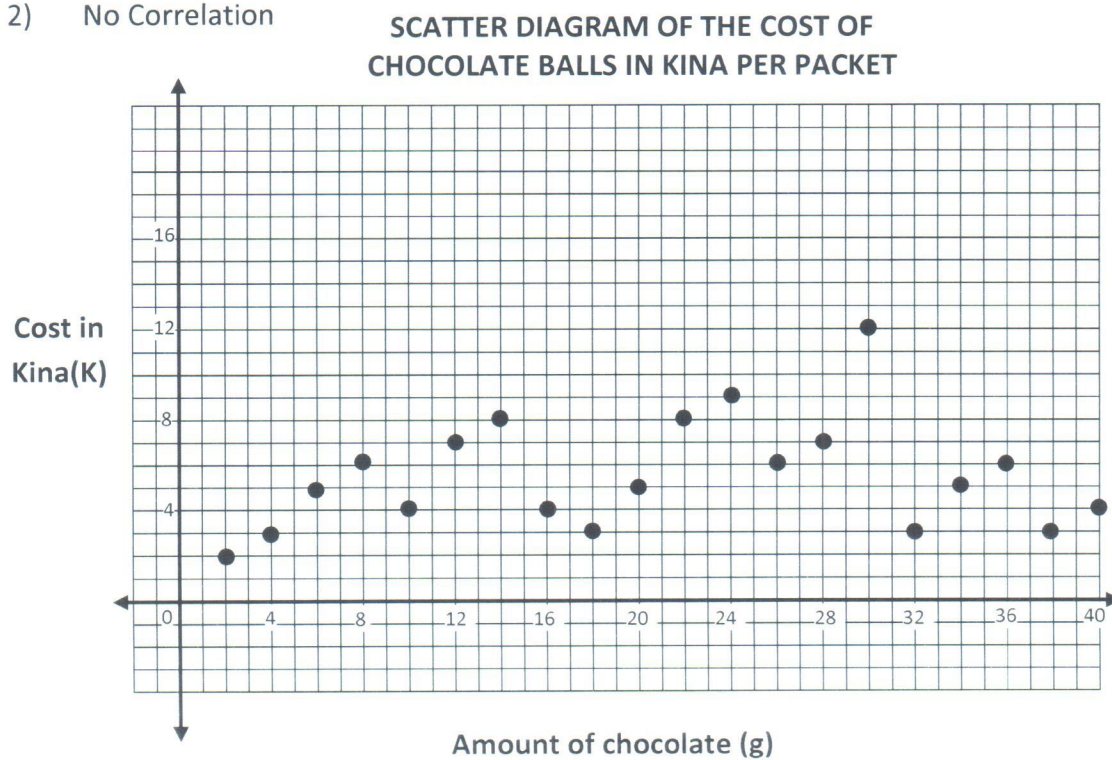


Learning Activity 12.3.3.1

1) Strong Positive Correlation



2) No Correlation



**Learning Activity 12.3.3.2**

- 1) a) No Correlation
 b) Positive correlation
 c) Negative correlation
 d) No correlation

2) a)

Hospital	x	x^2	y	y^2	xy
A	6.8	46.24	3.5	12.25	23.8
B	5.3	28.09	3.1	9.61	16.43
C	8.2	67.24	6.8	46.24	55.76
D	3.1	9.61	1.1	1.21	3.41
E	3.5	12.25	1.01	1.0201	3.535
F	4.2	17.64	3.9	15.21	16.38
G	3.9	15.21	2.05	4.2025	7.995
H	10.3	106.09	8.6	73.96	88.58
I	24.5	600.25	22	484	539
J	10.8	116.64	8	64	86.4
Summation	$\Sigma x = 80.6$	$\Sigma x^2 = 1019.26$	$\Sigma y = 60.06$	$\Sigma y^2 = 711.70$	$\Sigma xy = 841.29$

- b) Identify the following

i. $n = 10$

ii. $\frac{1}{n} = \frac{1}{10} = 0.1$

iii. $\Sigma xy = 711.70$

iv. $\overline{xy} = 8 \times 6 = 48$

v. $\overline{x} = 80.06 \div 10 = 8.00$

vi. $\overline{y} = 60.06 \div 10 = 6.00$

c) $S_x = 6.16$

d) $S_y = 5.93$

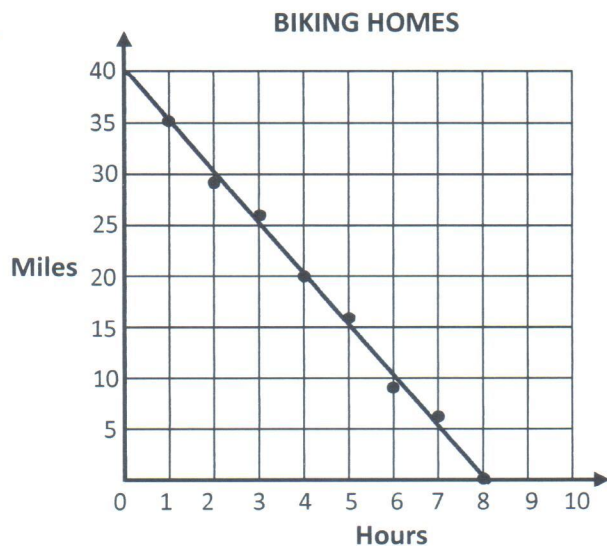
e) $r = \frac{(0.1)(841.29) - (8.00)(6.00)}{(6.16)(5.93)} = \frac{(84.13) - (48)}{36.53} = \frac{36.13}{36.53} = 0.99$

Interpretation: **Very Strong Positive Correlation**



Learning Activity 12.3.3.3

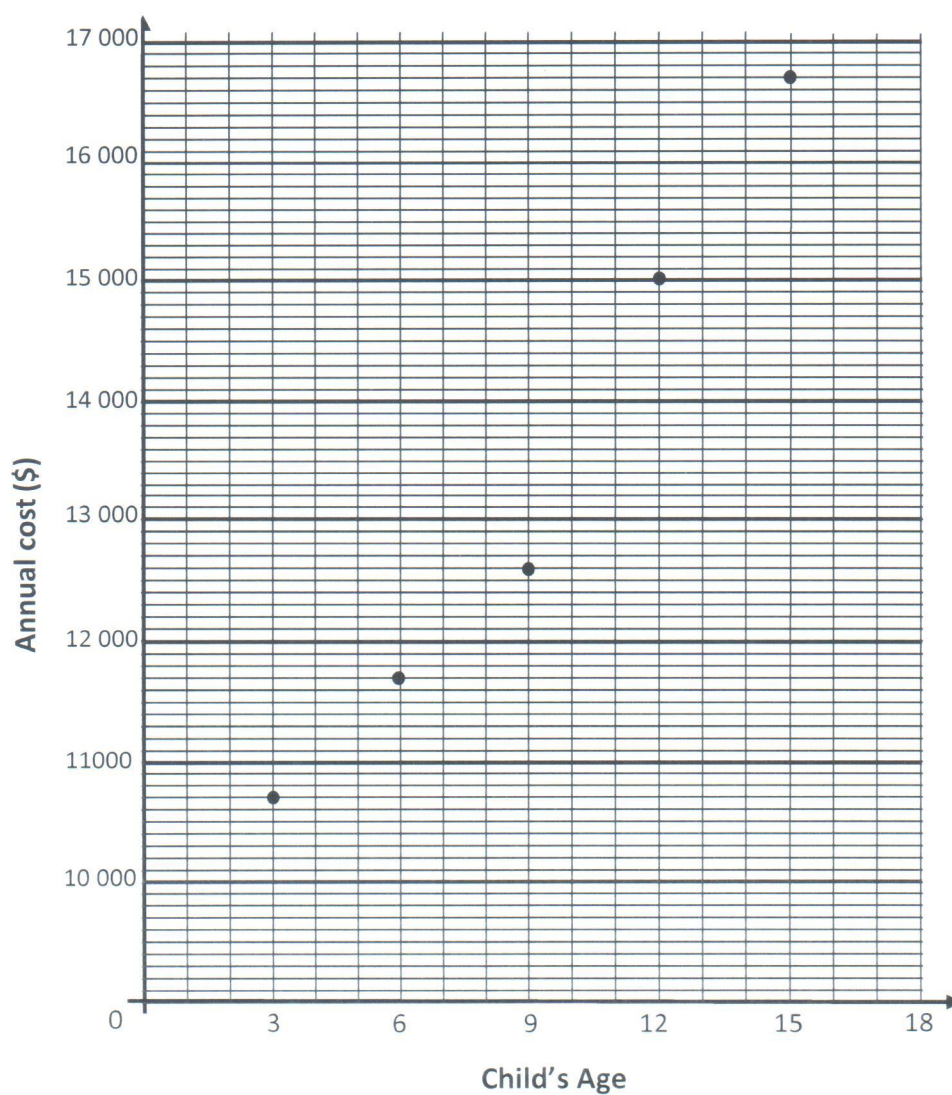
1.



$$\text{Slope} = \frac{35}{8}$$

$$\text{Equation of the line: } y = \frac{35}{8}x + 40$$

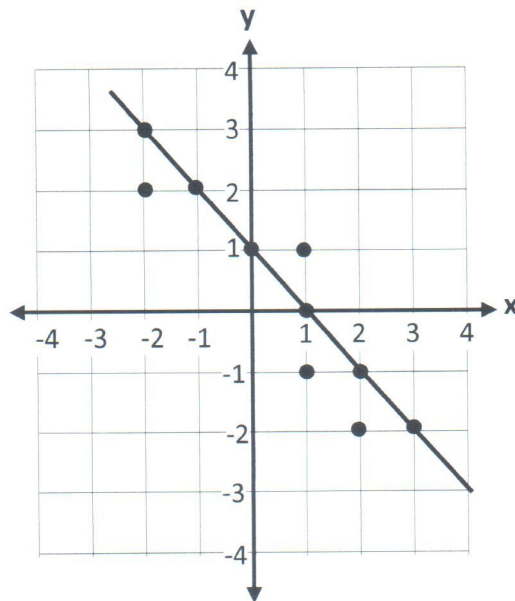
2.





Summative Task 12.3.3

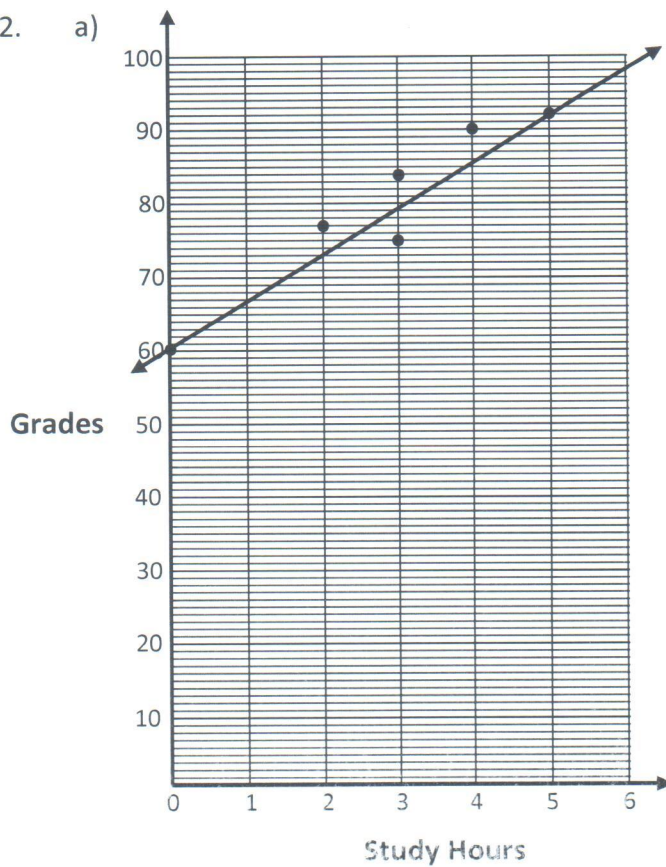
1)



Equation of the line of best fit:

$$y = -x + 1$$

2. a)



$$b) \quad y = \frac{32}{5}x + 60$$

$$c) \quad 98$$

d) The line may go forever or infinitely in increasing direction as shown in the graph. But since the data in this example are about grades, they are usually limited to 100.

END OF UNIT MODULE 12.3



REFERENCES

- http://www.triblocal.com/libertyville/files/cache/2012/05/conference-room-table12.jpg/460_345_resize.jpg
 - <http://pixshark.com/wait-in-line-clipart.htm>
 - <http://datascience.berkeley.edu/wp-content/uploads/2014/03/Colorful-Books.jpg>
 - <https://s-media-cache-ak0.pinimg.com/originals/23/67/84/2367849527ac778a3b0bbaf3390570ec.jpg>
 - <https://ilovewesthartfordct.files.wordpress.com/2014/04/runner1.gif>
 - http://gofitu.com/wp-content/uploads/2015/06/tumblr_static_wf-donuts-1.jpg
 - http://i633.photobucket.com/albums/uu52/euphemia_dco/gambling/JokerCard.jpg
 - <http://www.classzone.com/eservices/home/pdf/student/LA212EAD.pdf>
 - [http://www.lemars.k12.ia.us/webfiles/mboyd/Pre-Algebra%20Textbook%20\(e-edition\)/Source/J9D11IBD.pdf](http://www.lemars.k12.ia.us/webfiles/mboyd/Pre-Algebra%20Textbook%20(e-edition)/Source/J9D11IBD.pdf)
 - <http://www.northlandprep.org/wp-content/uploads/2014/08/Compound-Probability-WS.pdf>
 - http://anakamura.weebly.com/uploads/4/9/1/4/4914438/alg2y_10-8_lecture_notes.pdf
 - <http://www.unco.edu/nhs/mathsci/facstaff/Roberson/CourseDocs/MATH%20182/Activities/Independent%20and%20Dependent%20Probability.pdf>
 - https://www.bigideasmath.com/protected/content/ipe/grade%207/09/g7_09_04.pdf
 - <http://www.glencoe.com/sec/math/prealg/mathnet/pr01/pdf/1009a.pdf>
 - <http://www.northlandprep.org/wp-content/uploads/2014/08/Compound-Probability-WS.pdf>
 - <http://www2.isye.gatech.edu/~brani/isyebayes/bank/handout1.pdf>
 - <http://www.ams.sunysb.edu/~jsbm/courses/311/conditioning.pdf>
 - <http://illuminations.nctm.org/uploadedImages/Content/Lessons/Images/preK-2/YARNBOYTEACHERsm.jpg>
 - http://www.math-mate.com/chapter31_2.shtml
-

FODE PROVINCIAL CENTRES CONTACTS

FODE PROVINCIAL CENTRES

PC N O.	FODE PROVINCIAL CENTRE	ADDRESS	CONTACTS/ CUG PHONES	WIRELESS PHONES	SENIOR CLERK	CUG PHONE
1	DARU	P. O. Box 68, Daru	72228146	77522841	Mr Kevin Sere	72229047
2	KEREMA	P. O. Box 86, Kerema	72228124	77522842	Mr David Saria	72229049
3	CENTRAL	C/- FODE HQ	72228110	77522843	Mr Aubi Elodo	72229050
4	ALOTAU	P. O. Box 822, Alotau	72228130	77522844	Mr Albi Bapera	72229051
5	POPONDETTA	P. O. Box 71, Popondetta	72228138	77522845	Mr Stansen Sevese	72229052
6	MENDI	P. O. Box 237, Mendi	72228142	77522846	Mr Wari Tange	72229053
7	GOROKA	P. O. Box 990, Goroka	72228116	77522847	Ms Ovin Tuki	72229054
8	KUNDIAWA	P. O. Box 95, Kundiawa	72228144	77522848	Mr Denge Gundu	72229056
9	MT HAGEN	P. O. Box 418, Mt. Hagen	72228148	77522849	Mr Robert Maki	72229057
10	VANIMO	P. O. Box 38, Vanimo	72228140	77522850	Mrs Isabella Danti	72229060
11	WEWAK	P. O. Box 583, Wewak	72228122	77522851	Mr David Wombui	72229062
12	MADANG	P. O. Box 2071, Madang	72228126	77522852	Mrs Applonia Bogg	72229063
13	LAE	P. O. Box 4969, Lae	72228132	77522853	Ms Cathrine Kila	72229064
14	KIMBE	P. O. Box 328, Kimbe	72228150	77522854	Mrs Bernadette Litom	72229065
15	RABAU	P. O. Box 83, Kokopo	72228118	77522855	Mrs Verlyn Vavai	72229067
16	KAVIENG	P. O. Box 284, Kavieng	72228136	77522856	Mr John Lasisi	72229069
17	BUKA	P. O. Box 154, Buka	72228108	77522857	Mr Marlyne Meiskamel	72229073
18	MANUS	P. O. Box 41, Lorengau	72228128	77522858	Ms Roslyn Keket	72229080
19	NCD	C/- FODE HQ	72228134	77522859	Mrs Marina Tomiyavau	72229081
20	WABAG	P. O. Box 259, Wabag	72228120	77522860	Mr Salas Kamberan	72229082
21	HELA	P. O. Box 63, Tari	72228141	77522861	Mr Ogai John	72229083
22	JIWAKA	c/- FODE Hagen	72228143	77522862	Joseph Walep	72229085

SUBJECT AND GRADE TO STUDY

Grade Levels	Subjects
Grades 7 and 8	1. English
	2. Mathematics
	3. Science
	4. Social Science
	5. Making a Living
	6. Personal Development
	7. English
Grades 9 and 10	1. English
	2. Formal Mathematics
	3. Practical Mathematics
	4. Science
	5. Social Science
	6. Commerce
	7. Design and Technology- Computing
	8. Personal Development
Grades 11 and 12	1. English
	• Applied English
	• Language and Literature
	2. Mathematics
	• Mathematics A
	• Mathematics B
	3. Science
	4. Social Science
	• History
	• Geography
	• Economics
	5. Business Studies
	6. Personal Development
	7. ICT

REMEMBER:

- For Grades 7 and 8, you are required to do all six (6) courses.
- For Grades 9 and 10, you must study English, Mathematics, Science, Personal Development, Social Science and Commerce. Design and Technology-Computing is optional.
- For Grades 11 and 12, you are required to complete seven (7) out of thirteen (13) courses to be certified. Your Provincial Coordinator or Supervisor will give you more information regarding each subject.

Certificate in Matriculation

CORE COURSES

Basic English
English 1
English 2
Basic Maths
Maths 1

Maths 2

History of Science & Technology

OPTIONAL COURSES

Science Streams: Biology
Chemistry, Physics and Social Science Streams:
Geography, Introduction to Economics and Asia and the Modern World

REMEMBER:

You must successfully complete 8 courses; 5 compulsory and 3 optional