MATHEMATICAL APPLICATIONS 3

WEEK 1 NOTES & EXERCISES

Bivariate Data

A manager of a small ski resort has encountered a problem. She wants to be able to predict the number of skiers using her resort each weekend in advance, so that she can organise additional staffing catering if needed. She knows that good deep snow will attract skiers in big numbers, but scant covering is unlikely to attract a crowd. To investigate the situation further, she collects the following data over twelve consecutive weekends at her resort.

Depth of snow (m)	Number of skiers
0.5	120
0.8	250
2.1	500
3.6	780
1.4	300
1.5	280
1.8	410
2.7	320
3.2	640
2.4	540
2.6	530
1.7	200

As there are two types of data in this example, this is known as **bivariate data**. For each item (weekend), two variables are considered (depth of snow and number of skiers). When analysing bivariate data, we are interested in examining the relationship between the two variables.

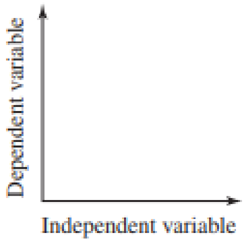
In this case, the manager might be interested in answering the following questions:

1. Are visitor numbers related to depth of snow?
2. If there is a relationship between visitor numbers and depth of snow, is it always true or is it just a guide? In other words, how strong is the relationship?
3. How much confidence could be placed in the prediction?

In a relationship involving two variables, if the values of one variable **'depend'** on the values of another variable, then the former variable is referred to as the **dependent variable** and the latter variable is referred to as the **independent variable**. When a relationship between two sets of variables is being examined, it is important to know which one of the two variables depends on the other. Most often we can make a judgement about this, although sometimes it may not be possible.

Consider the case where a study compared the heights of company employees against their annual salaries. Common sense would suggest that the height of a company employee would not depend on the person's annual salary nor would the annual salary of a company employee depend on the person's height. In this case, it is not appropriate to designate one variable as independent and one as dependent.

In the case where the ages of company employees are compared with their annual salaries, you might reasonably expect that the annual salary of an employee would depend on the person's age. In this case, the age of the employee is the independent variable and the salary of the employee is the dependent variable. It is useful to identify the independent and dependent variables where possible since it is the usual practice when displaying data on a graph to place the **independent variable on the horizontal axis** and the **dependent variable on the vertical axis**.



The Australian Bureau of Statistics conducts **real life statistics** on different aspects of our lives to provide various government departments with information about the general population. We can also use this information to make informed decisions and predict what could happen in the future.

Extension: Use the link below to view the Mathspace lesson on *Dependent and Independent variables* and *Real Life Statistics - Society*. This provides additional information on this topic. After you view the lesson, try completing the question set.

[Dependent and independent variables](https://mathspace.co/textbook/subtopic/138337/lessons) (https://mathspace.co/textbook/subtopic/138337/lessons)

[Real Life Data – Stats in Society](https://mathspace.co/textbook/subtopic/138332/lessons) (https://mathspace.co/textbook/subtopic/138332/lessons)

Example
For each of the following pairs of variables, identify the independent variable and the dependent variable. If it is not possible to identify this, then write 'not appropriate'.

1. The number of visitors at a local swimming pool and the daily temperature
2. The blood group of a person and his or her favourite TV channel

Solution
1. It is reasonable to expect the number of visitors at the swimming pool on any day will depend on the temperature on that day.
   Daily temperature is the independent variable and number of visitors at a local swimming pool is the dependent variable

2. Common sense suggests that the blood type of a person does not depend on the person's TV channel preferences or vice versa.
   Not appropriate

Exercise Set 1

1. For each of the following pairs of variables, identify the independent variable and the dependent variable. If it is not possible to identify this, then write 'not appropriate'.

   a. The age of an AFL footballer and his annual salary

   b. The growth of a plant and the amount of fertiliser it receives

   c. The number of books read in a week and the eye colour of the readers

d. The voting intentions of a woman and her weekly consumption of red meat

e. The month of the year and the electricity bill for that month

f. The mark obtained for a Maths test and the number of hours spent preparing for the test

g. The mark obtained for a Maths test and the mark obtained for an English test

h. The cost of grapes (in dollars per kilogram) and the season of the year

i. Ticket sales and revenue of show

Two-way Frequency Table

When we are examining the relationship between two categorical variables, the two-way table is an excellent tool as it allows us to have a clear breakdown of the data.

Example
At a local shopping centre, 34 females and 23 males were asked which of the two major political parties they preferred. Nineteen females and 12 males preferred Labor. Display these data in a two-way frequency table.

Solution
Raw Data:

Party preference	Female	Male	Total
Labor	19	12	31
Liberal			
Total	34	23	57

Completed table:

Party preference	Female	Male	Total
Labor	19	12	31
Liberal	15	11	26
Total	34	23	57

This shows a clear breakdown of the data in terms of numbers. It shows more females prefer Labor (19 females vs 12 males). However, only 23 males were surveyed comparted to 34 females. Using percentages helps overcome this.

The table is filled in by expressing the number in each cell as a **percentage** of the **column's total**. For example, to obtain the percentage of males who prefer Labor, divide the number of males who prefer Labor by the total number of males and multiply by 100. Percentage is 52.2% (1 decimal place).

Party preference	Female	Male
Labor	55.9	52.2
Liberal	44.1	47.8
Total	100.0	100.0

Similar percentages of females and males preferred Labor.

However, we could also find the percentages of those who preferred Labor were female or male.

Female: $\frac{19}{31} \times 100 = 61.3\%$ of those preferring Labor were females.
Male: $\frac{12}{31} \times 100 = 38.7\%$ of those preferring Labor were males.

The general rule is that the independent variable (respondent's gender) is placed in the columns of the table and the percentages should be calculated in columns.

Comparing percentages in each row of a two-way table allows us to establish whether a relationship exists between the two categorical variables that are being examined. As we can see from the table, the percentage of females who preferred Labor is about the same as males. Likewise, the percentage of females and males preferring Liberal are almost equal. This indicates that for the group of people participating in the survey, party preference is not related to gender.

Exercise Set 2

1. Members of a gym club were asked what kind of training they do. Each responder only did one kind of training. The table shows the results.

	Cardio	Weight
Female	44	18
Male	12	26

   a. How many gym members were asked altogether?

   b. How many members do weight training?

   c. What percentage of members do weight training?

2. Ben surveyed all the students in Year 12 at his school and summarised the results in the following table.

	Play sports	Do not play sports	Total
Height >170 cm	46	73	119
Height <170 cm	30	45	75
Total	76	118	194

a. What percentage of Year 12 students whose height is less than 170 cm play sports?

b. What percentage of students from Year 12 do not play sports?

3. In a study, some people were asked how many times they lie in a day. 20 responders said they life at least once a day, 5 of which were children. 13 children said they never lie, and 15 adults said they never lie.

   a. Complete the table

	0 times	1 or more times
Children		
Adults		

   b. What percentage of responders said they never lied?

   c. What percentage of adults said they had lied at least once?

4. Complete the table.

Attitude	Female	Male	Total
For	25		47
Against			
Total	51		92

5. In a survey, 139 women and 102 men were asked whether they approved or disapproved of a proposed freeway. 37 women and 79 men approved of the freeway. Display this data in a two-way table (not as percentages)

6. The following table shows the heart rate data of a group of people after exercise.

Height of step	Stepping rate	Heart rate
Short step	Slow	80
Short step	Slow	91
Short step	Medium	106
Short step	Medium	105
Short step	Fast	124
Short step	Fast	128
Tall step	Slow	100
Tall step	Slow	96
Tall step	Medium	125
Tall step	Medium	129
Tall step	Fast	132
Tall step	Fast	127

a. Complete the table. Give all answers to one decimal place.

		Stepping rate		
Height of step	**Data**	**Slow**	**Medium**	**Fast**
Short step	Average of heart rate	85.5		
Tall step	Average of heart rate			

b. Which of the following combinations of step height and stepping rate generated the higher heart rate?

i. A tall step at a slow stepping rate ii. A short step at a fast stepping rate

c. Considering a slow heart rate to be better, what category of responders was the healthiest?

7. The data show the reactions of administrative staff and technical staff to an upgrade of the computer systems at a large corporation.

Attitude	Administrative staff	Technical staff	Total
For	53	98	151
Against	37	31	68
Total	90	129	219

a. From the table, we can conclude that:
   i. 53% of administrative staff were for the upgrade
   ii. 37% of administrative staff were for the upgrade
   iii. 27% of administrative staff were against the upgrade
   iv. 59% of administrative staff were for the upgrade
   v. 54% of administrative staff were against the upgrade

b. From the table, we can conclude that:
   i. 98% of technical staff were for the upgrade
   ii. 65% of technical staff were for the upgrade
   iii. 76% of technical staff were for the upgrade
   iv. 31% of technical staff were against the upgrade
   v. 14% of technical staff were against the upgrade