

MATHEMATICAL APPLICATIONS 3

WEEK 2 NOTES & EXERCISES

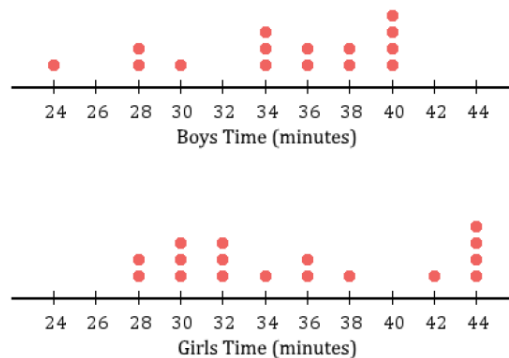
Parallel Dot Plots

Dot plots are a quick way to represent data visually. Dot plots are usually used for small data sets because they can help you to quickly see the spread of data and identify any outliers.

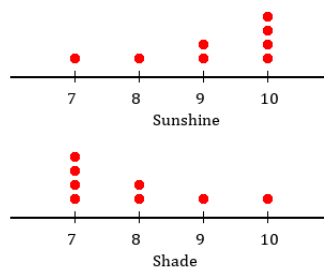
Parallel dot plots help you compare two or more sets of data, which must be plotted against the same scale. This makes the comparison between the data sets easy and ensures it isn't misleading. Parallel dot plots must be in the same unit and same scale.

Exercise Set 1

1. A class completed 40 questions for homework. The time needed for boys and girls to finish them was collected, and the data was presented as a parallel dot plot.



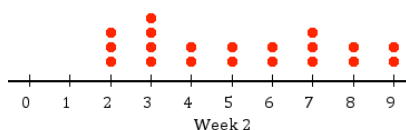
- a. Comparing boys and girls, which gender had the highest median time?
 - b. Which gender had the largest range?
 - c. Which group has the highest valued mode?
2. Elizabeth did an experiment to see how well plants grow in different conditions. She had 8 plants grow in the sunshine, and 8 plants grow in the shade. She measured how tall they grew in centimetres after 2 months and recorded the information as a parallel dot plot.



- a. Which group of plants had a higher range of heights?

- b. Which dot plot shows a positive skew?
- c. How much higher is the median height of plants that grew in the sunshine than the median height of plants that grew in the shade?

3. Twenty people joined a group fitness class and over two weeks, they were tested on the number of chin ups they can do. The dot plots show the number of chin ups each person could do.



- a. From the first to the second week, what was the increase in the median number of chin ups someone could do?
- b. In week 1, the average number of chin ups was 2.25. Did the average increase or decrease in the second week?
4. Mae and Amelia are goal shooters for their netball teams. The table shows the number of goals they score in each game of the season.

	Game 1	Game 2	Game 3	Game 4	Game 5	Game 6	Game 7	Game 8
Mae	2	5	5	4	3	2	2	4
Amelia	2	1	1	3	4	2	3	1

- a. Construct a parallel dot plots for this data.
- b. How does the range of each set of scores compare?

Divided Bar Charts

When comparing two categorical variables, it can be useful to represent the results from a two-way table (in percentage form) graphically. We can do this using divided (segmented) bar charts. A segmented bar chart consists of two or more columns, each of which matches one column in the two-way table. Each column is subdivided into segments, corresponding to each cell in that column. The segmented bar chart is a powerful visual aid for comparing and examining the relationship between two categorical variables.

Example

Sixty-seven primary and 47 secondary school students were asked about their attitude to the number of school holidays which should be given. They were asked whether there should be fewer, the same number, or more school holidays. 5 primary school students and 2 secondary students wanted fewer holidays, 29 primary and 9 secondary students thought they had enough holidays (that is, they chose the same number) and the rest thought they needed to be given more holidays. The data in percentage form is shown in two-way frequency tables and a segmented bar chart.

Solution

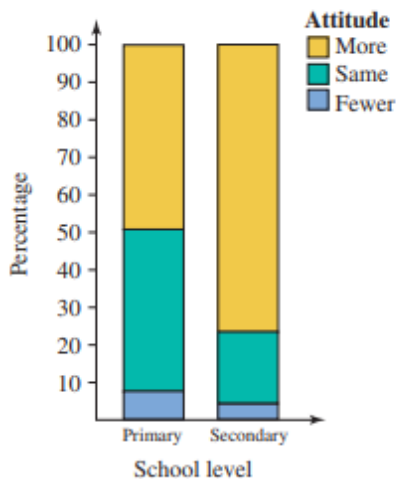
Raw Data:

Attitude	Primary	Secondary	Total
Fewer	5	2	7
Same	29	9	38
More	33	36	69
Total	67	47	114

Percentages:

Attitude	Primary	Secondary
Fewer	7.5	4.3
Same	43.3	19.1
More	49.2	76.6
Total	100	100

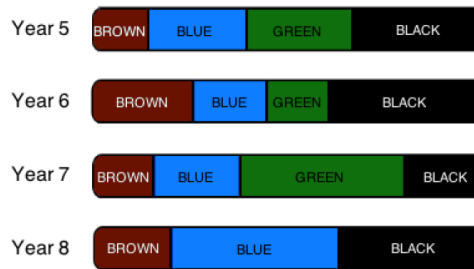
Divided Bar Chart:



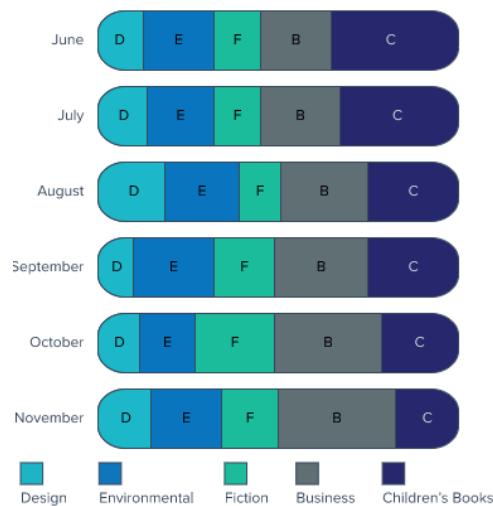
Secondary students were much keener on having more holidays than were primary students.

Exercise Set 2

1. A survey was conducted asking students from Year 5, 6, 7 and 8 what eye colour they have. Study the four divided bar graphs shown here and answer the following questions.

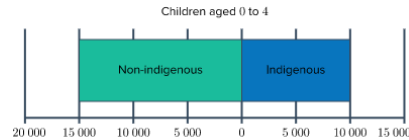


- Which grade has the largest percentage of brown-eyed children?
 - Which grade has the largest percentage of blue-eyed children?
 - Which grade has the largest percentage of black-eyed children?
 - Which grade has no green-eyed children?
 - Which grade has the same percentage of green-eyed children as blue-eyed children?
2. The local library constructed divided bar graphs showing the last six months of book loans by category. Assume that the total number of loans was the same for each month.



- In which month did subscribers borrow the most children's books?
- In which month did people borrow the same number of design books as environment books?
- Which book category grew constantly during the 6 months?

3. The divided bar graph shows the population of Indigenous and non-indigenous children aged between 0 to 4.



- a. What percentage of children aged between 0 and 4 are Indigenous?
- b. If the graph is 45 centimetres long in total, how long is the section representing non-indigenous children?
4. Sixty single men and women were asked whether they prefer to rent by themselves, or to share accommodation with friends. The results are shown below.

Preference	Men	Women	Total
Rent by themselves	12	23	35
Share with friends	9	16	25
Total	21	39	60

- a. Convert these numbers to percentages and show the results in a table.
- b. Display this information in a divided bar chart.

Scatter Plots

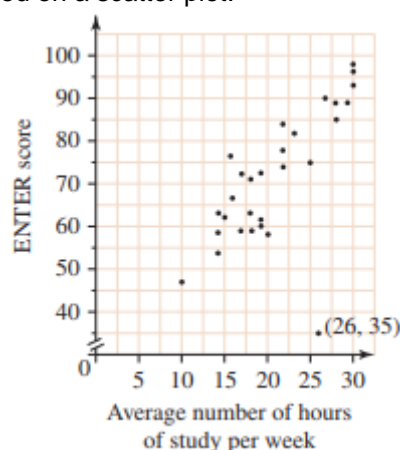
We often want to know if there is a relationship between two numerical variables. A scatter plot, which gives a visual display of the relationship between two variables, provides a good starting point. Consider the data obtained from last year's class at Northbank College. Each student in the class of 29 students was asked to give an estimate of the average number of hours they studied per week during Year 12. They were also asked for the ENTER score they obtained.

Average hours of study	ENTER score
18	59
16	67
22	74
27	90
15	62
28	89
18	71
19	60
22	84
30	98

Average hours of study	ENTER score
14	54
17	72
14	63
19	72
20	58
10	47
28	85
25	75
18	63
19	61

Average hours of study	ENTER score
17	59
16	76
14	59
29	89
30	93
30	96
23	82
26	35
22	78

The figure below shows the data plotted on a scatter plot.



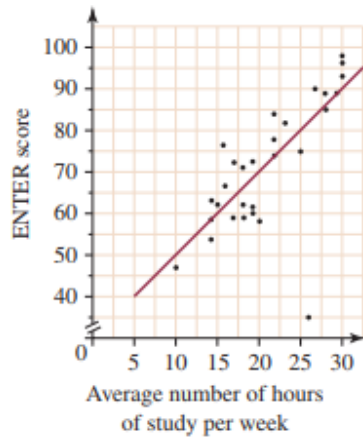
It is reasonable to think that the number of hours of study put in each week by students would affect their ENTER scores and so the number of hours of study per week is the independent variable and appears on the horizontal axis. The ENTER score is the dependent variable and appears on the vertical axis.

In analysing the scatter plot, we look for a pattern in the way the points lie. Certain patterns tell us that certain relationships exist between the two variables. This is referred to as **correlation**. We look at what type of correlation exists and how strong it is.

In the figure above, we see some sort of pattern: the points are spread in a rough corridor from bottom left to top right. We refer to this data following such a direction as having a **positive** relationship. This tells us that as the average number of hours studied per week increases, the ENTER score also increases.

The point (26,35) is an **outlier**. It stands out because it is well away from the other points and clearly is not part of the 'corridor' referred to previously. This outlier may have occurred because a student exaggerated the number of hours he or she worked in a week or perhaps there was a recording error. This needs to be checked.

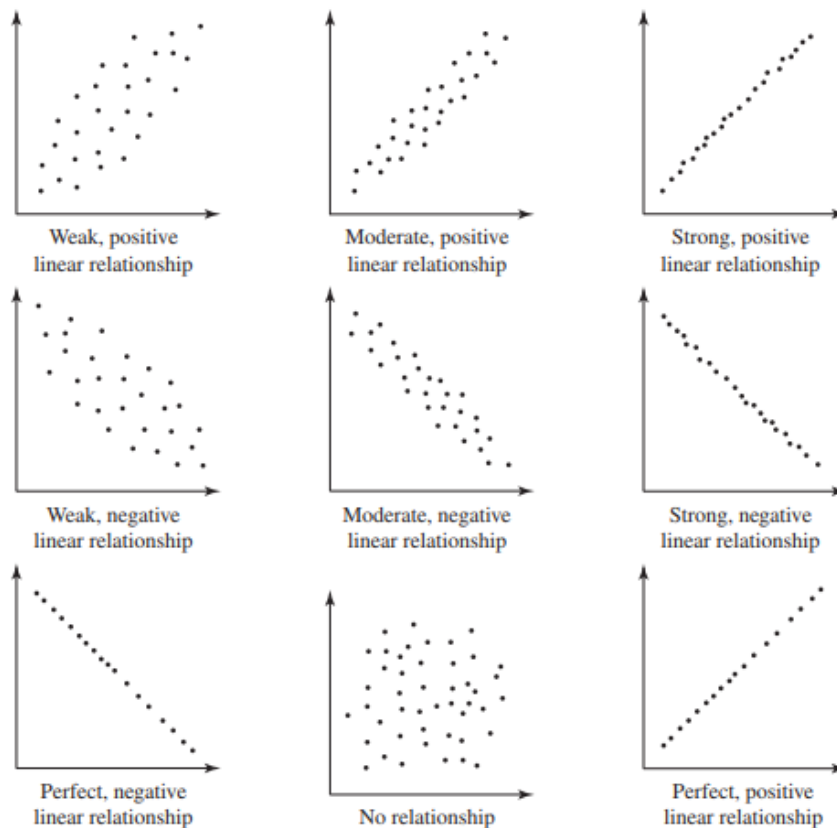
We could describe the rest of the data as having a linear form as the straight line in the diagram below indicates.



When describing the relationship between two variables displayed on a scatter plot, we need to comment on:

- The direction – whether it is positive or negative
- The form – whether it is linear or non-linear
- The strength – whether it is strong, moderate or weak
- Possible outliers.

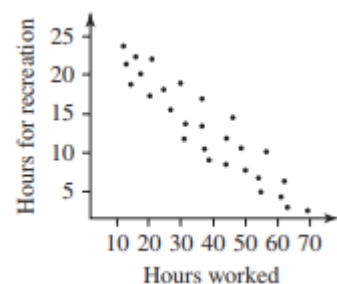
Below is a gallery of scatter plots showing the various patterns we look for:



Example

The scatter plot on the right shows the number of hours people spend at work each week and the number of hours people get to spend on recreational activities during the week.

Decide whether a relationship exists between the variables and comment on whether it is positive or negative; weak, moderate or strong; and whether it has a linear form.



Solution

- The points on the scatter plot are spread in a certain pattern, namely in a rough corridor from the top left to the bottom right corner. This tells us that as the work hours increase, the recreation hours decrease.
- The corridor is straight (that is, it would be reasonable to fit a straight line into it).
- The points are neither too tight nor too dispersed.
- The pattern resembles the central diagram in the gallery of scatter plots show previously.

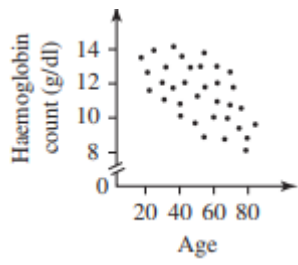
There is a moderate, negative relationship between the two variables.

Exercise Set 3

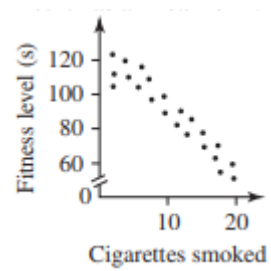
1. For each of the following pairs of variables, write down whether or not you would reasonably expect a relationship to exist between the pair and, if so, comment on whether it would be a positive or negative association.
 - a. Time spent in a supermarket and money spent
 - b. Income and value of car driven
 - c. Number of children living in a house and time spent cleaning the house
 - d. Age and number of hours of competitive sports played per week
 - e. Amount spent on petrol each week and distance travelled by car each week
 - f. Number of hours spent in front of a computer each week and time spent playing the piano each week
 - g. Amount spent on weekly groceries and time spent gardening each week
 - h. Amount of time spent studying and grade on the test

2. Use the wording in the gallery above to describe the correlation between the variables shown in the graphs below

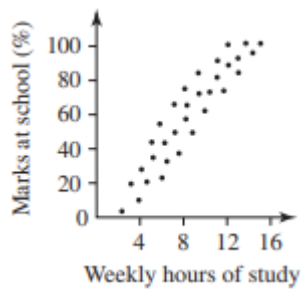
a.



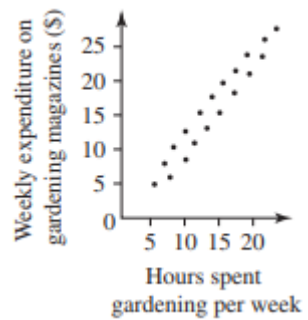
b.



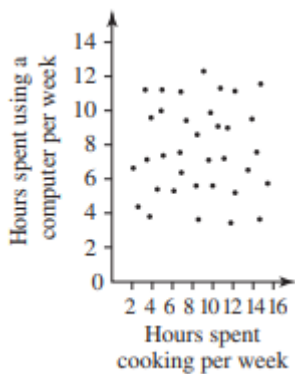
c.



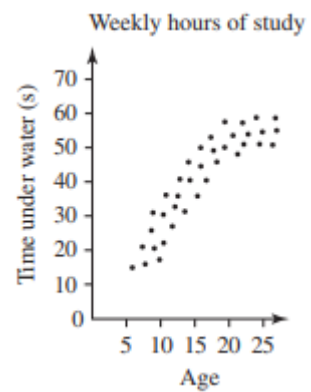
d.



e.



f.



3. The population of a municipality (to the nearest ten thousand) together with the number of primary schools in that particular municipality is given below for 11 municipalities.

Population (x1000)	110	130	130	140	150	160	170	170	180	180	190
Number of primary schools	4	4	6	5	6	8	6	7	8	9	8

Construct a scatter plot for the data and use it to comment on the direction, form and strength of any relationship between the population and the number of primary students.

Note: The population is the independent variable and thus becomes the x axis.

4. The table below contains data for the time taken to do a paying job and the cost of the job. Determine the independent variable. Construct a scatter plot for the data. Comment on whether a relationship exists between the time taken and the cost. If there is a relationship, describe it.

Time taken (hours)	Cost of job (\$)
5	1000
7	1000
5	1500
8	1200
10	2000
13	2500
15	2800
20	3200
18	2800
25	4000
33	3000