Bivariate analysis is concerned with the relationships between pairs of variables (x, y) in a data set.

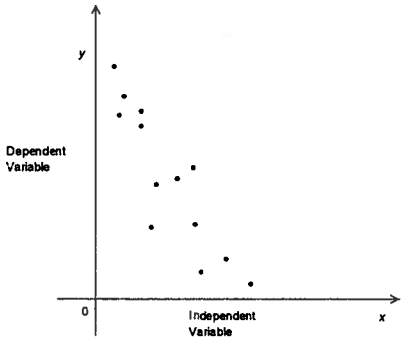
10.1 Scatter diagrams

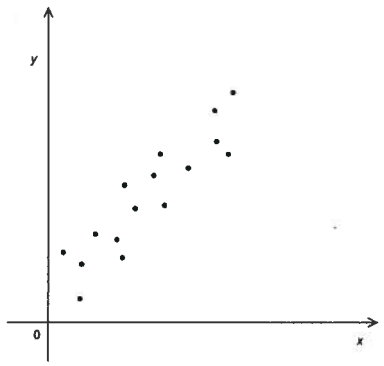
One way to view data is by showing it on a scatter diagram

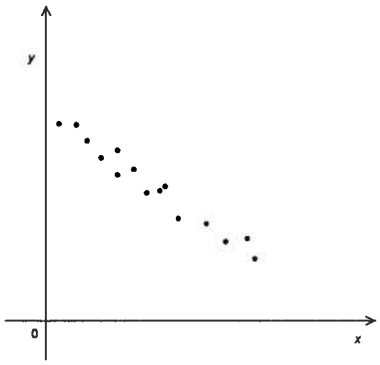
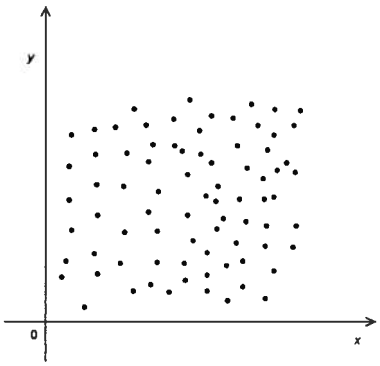
- **Scatter diagrams** (also called scatter plots) are used to investigate the possible relationship between two variables that both relate to the same "event".
- Scatter diagram have a specific purpose to show how much one variable affects another.
- The relationship between two variables is called their correlation
  - Correlations is a way to measure how associated or related two variables are.
  - The purpose of correlations is to allow us to make a prediction about one variable based on what we know about another variable.

To draw a scatter diagram plot the (x, y) value from the data table as dots on a graph. The pattern formed by the dots can give us some indication of the correlation.

- The independent variable should be on the horizontal axis
- The dependent variable is on the vertical axis.

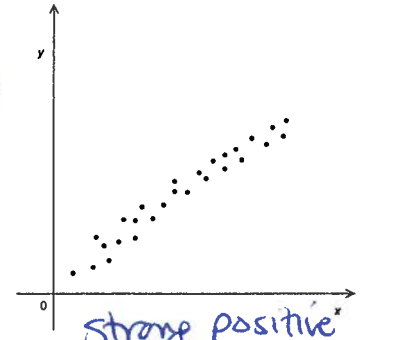


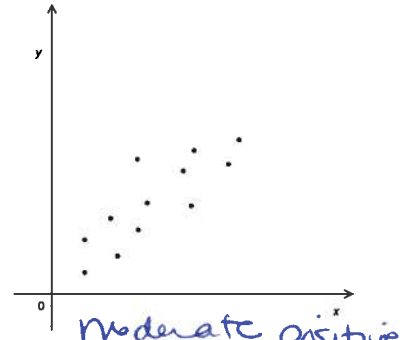
- A general upward trend in the pattern of dots shows positive correlation. - The value of the dependent variable increases as the value of the independent variable increases.	

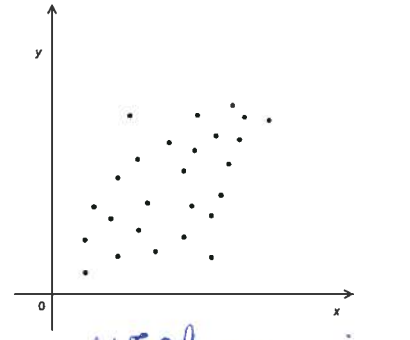
- A general downward trend in the pattern of dots shows negative correlation. - The dependent variable decreases as the independent variable increases.	
- Scattered points with no trend may indicate correlation close to zero.	

Scatter diagrams allow us to assess the strength of a correlation.

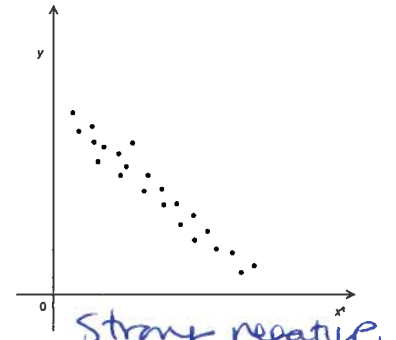
y ↑ as x ↑

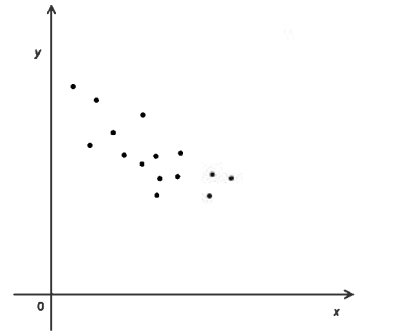


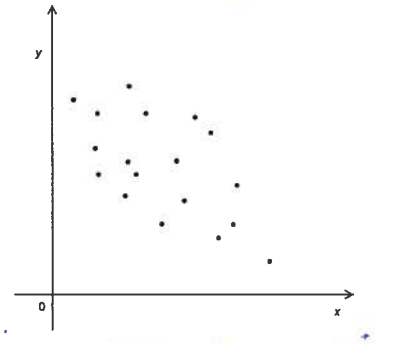




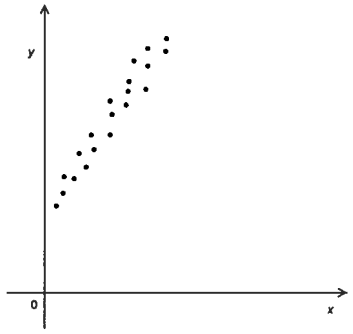
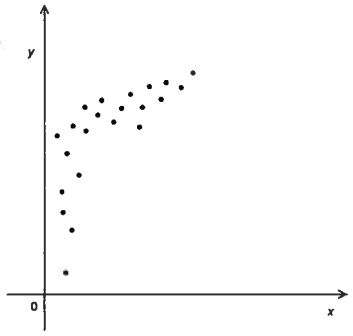
y ↓ as x ↑







Not all relationships are linear.

- The points on this graph are approximately linear.	
- The points on this graph would be represented by a curve. - There is a non-linear relationship between the variables.	

Causation v. Correlation * possible exploration idea?

- A correlation between two data sets does not necessarily mean that one causes the other.
- EXAMPLE: The shoe sizes of grade school students and the students' vocabulary have a strong positive correlation. Meaning the larger the shoe size the larger a students' vocabulary. Does one cause the other? No –
- There is a confounding factor. As a student gets older their shoe size increases & so does their vocabulary!

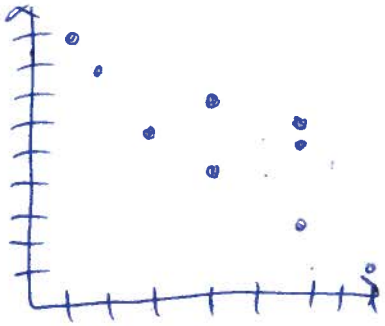
Example:

a. Represent the data on a scatter diagram.

x	1	2	3	4	4	6	6	6	8
y	9	8	6	5	7	6	5	3	1

b. Is the relationship linear or non-linear?
linear

c. Describe the type and strength of the relationship.
strong negative



Exercise 10A