**4A Correlation & Drawing Scatter Diagrams**



1. In the study of a city, the population density, in people/hectare, and the distance from the city centre, in km, was investigated by choosing sample areas. The results are as follows:



Plot a scatter graph and describe the correlation. Interpret what the correlation means.



1. Hideko was interested to see if there was a relationship between what people earn and the age at which they left education or training. She asked 14 friends to fill in an anonymous questionnaire and recorded her results in a scatter diagram.
2. Describe the type of correlation shown



Hideko says that her data supports the conclusion that more education causes people to earn a lower hourly rate of pay

1. Give one reason why Hideko’s conclusion might not be valid

**4B Linear Regression - Lines of Best Fit**

1. From the large data set, the daily mean windspeed, $w$ knots, and the daily maximum gust, $g$ knots, were recorded for the first 15 days in May in Camborne in 2015.

The data was plotted on a scatter diagram:

1. Describe the correlation between daily mean windspeed and daily maximum gust



The equation of the regression line of $g$ on $w$ for this data is:

$$g=7.23+1.82w$$

1. Give an interpretation of the value of the gradient of this regression line
2. Justify the use of a linear regression line in this case.
3. The head circumference, $y$ cm, and gestation period, $x$ weeks, for a random sample of newborn babies were recorded, and the scatter graph shows the results.

 

The equation of the regression line of $y$ on $x$ is:

$$y=8.91+0.624x$$

1. This equation is used to estimate the head circumference of a baby born after 39 weeks and a baby born after 30 weeks. Comment on the reliability of these estimates.

A nurse wants to estimate the gestation period for a baby born with a head circumference of 31.6cm.

1. Explain why the regression equation above is not suitable for this estimate