

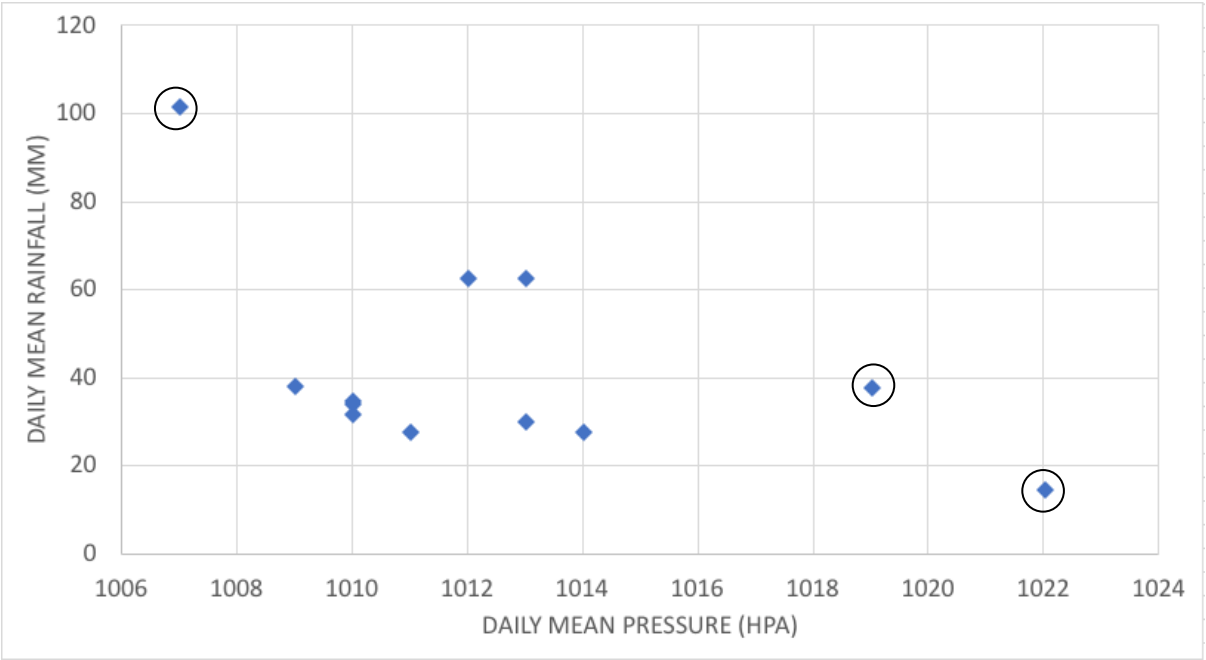
3. Pete is investigating the relationship between daily rainfall,  $w$  mm, and daily mean pressure,  $p$  hPa, in Perth during 2015. He used the large data set to take a sample of size 12.

He obtained the following results.

$p$	1007	1012	1013	1009	1019	1010	1010	1010	1013	1011	1014	1022
$w$	102.0	63.0	63.0	38.4	38.0	35.0	34.2	32.0	30.4	28.0	28.0	15

Pete drew the following scatter diagram for the values of  $w$  and  $p$  and calculated the quartiles.

	$Q_1$	$Q_2$	$Q_3$
$p$	1010	1011.5	1013.5
$w$	29.2	34.6	50.7



An outlier is a value which is more than 1.5 times the interquartile range above  $Q_3$  or more than 1.5 times the interquartile range below  $Q_1$ .

- (a) Show that the 3 points circled on the scatter diagram above are outliers. (2)
- (b) Describe the effect of removing the 3 outliers on the correlation between daily rainfall and daily mean pressure in this sample. (1)

John has also been studying the large data set and believes that the sample Pete has taken is not random.

- (c) From your knowledge of the large data set, explain why Pete’s sample is unlikely to be a random sample. (1)

John finds that the equation of the regression line of  $w$  on  $p$ , using all the data in the large data set, is

$$w = 1023 - 0.223p$$

(d) Give an interpretation of the figure  $-0.223$  in this regression line.

(1)

John decided to use the regression line to estimate the daily rainfall for a day in December when the daily mean pressure is 1011 hPa.

(e) Using your knowledge of the large data set, comment on the reliability of John's estimate.

(1)

**(Total for Question 3 is 6 marks)**