1. The following data set represents the percentage cloud cover at midday on 1 July for 20 consecutive years:

Percentage range	Number of times observed
0-9	6
10-24	4
25-49	4
50-100	6

a) Use a graphical method to display this data set, explaining how your method copes with the uneven interval widths.

b) One student produces the graph below in answer to above section. Describe what the student has done, why you think they might have done it, and it has been done correctly.
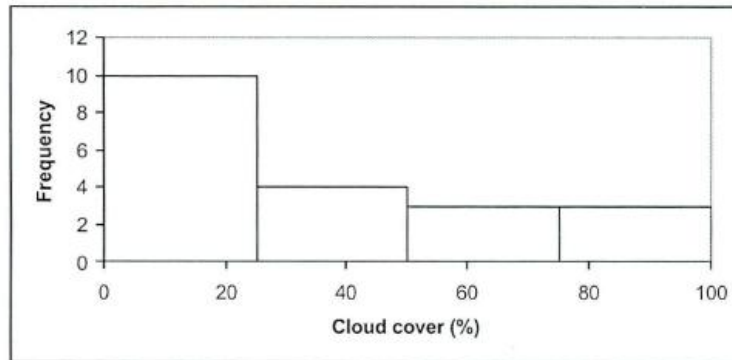
2. For ten consecutive days I aim to catch the 8:25 train to work in the morning and the 18:48 train home in the evening. On each day I note down how late (in minutes) each train leaves, as shown in the below

Day	1	2	3	4	5	6	7	8	9	10
Morning	5	7	3	14	0	2	6	1	23	7
Evening	0	2	0	7	0	0	1	6	4	0

a) Use a back-to-back stem-and-leaf diagram to compare the two data sets
b) Calculate the median and quartiles of each data set and determine whether there are any outliers present.
c) Compare the location, spread and skewness of the two data sets.

3.
a) Prove that if $y_i = a + bx_i$, $i = 1,2, \dots, n$, then
   i.   the sample mean of the y_i is given by $\quad \bar{y} = a + b\bar{x}$
   ii. the sample variance of the y_i is given by $\quad s_y^2 = b^2 s_x^2$,

   where $\bar{x}$ and s_x^2 are the sample mean and sample variance of the x_i, respectively.

b) The maximum temperature during July recorded at a weather station in London for the past 12 years is given in the table below

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
°C	24.3	22.3	24.2	22.7	23.3	28.2	21.4	22.8	23.0	25.0	21.7	21.3
°F	75.7	72.1	75.6	72.9	73.9	82.8	70.5	73.0	73.4	77.0	71.1	70.3

   i.   Calculate the sample mean and standard deviation of the temperature in degree Celsius (°C)
   ii. Calculate the sample mean and standard deviation of the temperature in degree Fahrenheit (°F)
   (Hint: $F = \frac{9}{5}C + 32$)