SMM634 Exercises - Simple Linear Regression

1. A manager of a city hospital investigated how the number, y, of unauthorized days which employees were absent during the last year was affected by the distance, x, in kilometres between home and work for the employees. The following data were collected on a random sample of 10 employees.

Distance (km)	1.3	2.9	4.4	5.8	8.2	9.9	12.1	13.8	14.2	17.7
Days absent	8	5	7	5	6	3	5	4	2	2

   (a) Explain why 'number of days absent' is chosen as the y variable.

   (b) Using R draw a scatterplot of number of days absent against distance to work.

   (c) Using R find the regression line for the numbers of days absent against distance to work.

   (d) The slope is negative. Explain what this tells you about the distance travelled to work and numbers of days absent.

   (e) Predict the number of days absent for an employee who lives a distance of 10 km from the hospitals.

   (f) Would you expect the straight line relationship to continue to hold for employees living more that 25 kilometres from the hospital? Discuss.

   (g) Using residual analysis, check that the model assumptions hold.