Correlation and Regression

1. In 'The Effects of Temperature of Marthon Runners' Performance' By David Martin and John Buoncristiani (Chance Magazine), high tmperatures and times (in minutes) were given for women who won the New York City marathon in recent years. The results are shown in the table below:

Temp($^\circ F$)	55	61	49	62	70	73	51	57
Time	145.28	148.72	148.30	148.10	147.62	146.40	144.67	147.53

   i. Represent the data on a scatter graph.
   ii. Calculate the correlation coefficient.
   iii. Does it appear that winning times are affected by temperature? Explain your answer.
   iv. Add the line of best fit to your scatter plot.
   v. Find the equation of the line of best fit.
   vi. In what time would you expect someone to win the race, in 50° F heat?

2. The heights (in centimetres) and weights (in kilograms) of eight adults were as follows:

Height	157	181	203	214	197	178	162	210
Weight	59	84	81	100	92	77	61	105

   i. Plot the results on a scatter graph.
   ii. Calculate the correlation coefficient.
   iii. Add the line of best fit to your scatter plot.
   iv. Find the equation of the line of best fit.
   v. What weight would you expect someone to be if they were 200 centimetres tall?

3. The following table shows the age and annual income of a sample of eight employees from a large company. The income is in thousands of euro.

Age	36	25	44	48	32	50	33	40
Income	36	29	53	55	34	62	38	46

   i. Draw a scatter diagram of the data

ii. Calculate the correlation coefficient; and comment on the correlation.

iii. Add the line of best fit to your scatter plot.

iv. Find the equation of the line of best fit.

v. What income would you expect a 35 year old to have?

4. A manager wishes to find out whether there is a relationship between the number of radio ads aired per week and the amount of sales (in thousands of dollars) of a product. The data for the sample follow.

No. of ads,x	2	5	8	8	10	12
Sales,y	\$2	\$4	\$7	\$6	\$9	\$10

i. Plot the results on a scatter graph.

ii. Calculate the correlation coefficient.

iii. Add the line of best fit to your scatter plot.

iv. Find the equation of the line of best fit.

5. A researcher wishes to determine if a person's age is related to the number of hours he or she exercises per week. The data for the sample are shown here.

Age,x	18	26	32	38	52	59
Hours,y	10	5	2	3	1.5	1

i. Plot the results on a scatter graph.

ii. Calculate the correlation coefficient.

iii. Add the line of best fit to your scatter plot.

iv. Find the equation of the line of best fit.

6. A store manager wishes to find out whether there is a relationship between the age of her employees and the number of sick days they take each year. Tha data for the sample follow.

Age,x	18	26	39	48	53	58
Days,y	16	12	9	5	6	2

i. Plot the results on a scatter graph.

ii. Calculate the correlation coefficient.

iii. Add the line of best fit to your scatter plot.

iv. Find the equation of the line of best fit.

7. An insurance company wants to determine the strength of the relationship between the number of hours a person works per week and the number of injuries or accidents that person has over a period of one week. The data follow.

Hours worked,x	40	32	36	44	41
No. of Accidents,y	1	0	3	8	5

i. Plot the results on a scatter graph.

ii. Calculate the correlation coefficient.

iii. Add the line of best fit to your scatter plot.

iv. Find the equation of the line of best fit.