Maths
Leaving Certificate
Ordinary Level

Past Exam Questions on
Statistics and Data

Q7 2013 Paper 2 Project Maths Section B

Question 7 **(75 mar**

The table below shows the rates of births, marriages and deaths in Ireland from 1990 to 2010. Th rates are per 10 000 of the estimated population.

Number of Births, Marriages and Deaths in Ireland (per 10 000 of the estimated population)			
Year	**Births**	**Marriages**	**Deaths**
1990	151	51	90
1991	150	49	89
1992	144	47	87
1993	138	47	90
1994	135	46	86
1995	135	43	90
1996	140	45	87
1997	144	43	86
1998	146	45	85
1999	144	50	87
2000	145	51	83
2001	150	50	79
2002	155	52	76
2003	155	51	73
2004	153	52	71
2005	148	52	68
2006	154	52	67
2007	163	52	64
2008	168	50	63
2009	167	48	63
2010	165	46	61

(Source: Central Statistics Office, http://www.cso.ie)

(a) Complete the back to back stem and leaf plot below to show the marriage rate and death rate in Ireland during the period covered in the table above.

Marriage rate		Death rate
	4	
	5	
	6	
	7	
	8	
	9	
		Key:

(b) State one difference that can be observed between the distributions of the marriage rate and the death rate in your plot.

(c) Find the median and interquartile range of the yearly marriage rates in Ireland from 1990 to 2010.

Median

Interquartile range

(d) **(i)** Find the mean of the death rate in Ireland from 1990 to 2010. Give your answer correct to one decimal place.

(ii) The standard deviation of the death rates in the table over is 10·3. List all of the death rates that are within 1 standard deviation of the mean.

(e) In 2010, the number of children born in Ireland was 75 174. Use this number to estimate the total population of Ireland in 2010.

(f) Use your answer to **(e)** to estimate the number of people who died in Ireland in 2010.

(g) "More children were born in Ireland in 1990 than in 2000." Give a reason, based on the data, why this statement is not necessarily true.

(h) Find the ratio, Birth rate : Death rate, for the two years 1990 and 2010. Based on your answers for the two years, what would you predict about the population of Ireland in future years. Give a reason for your answer.

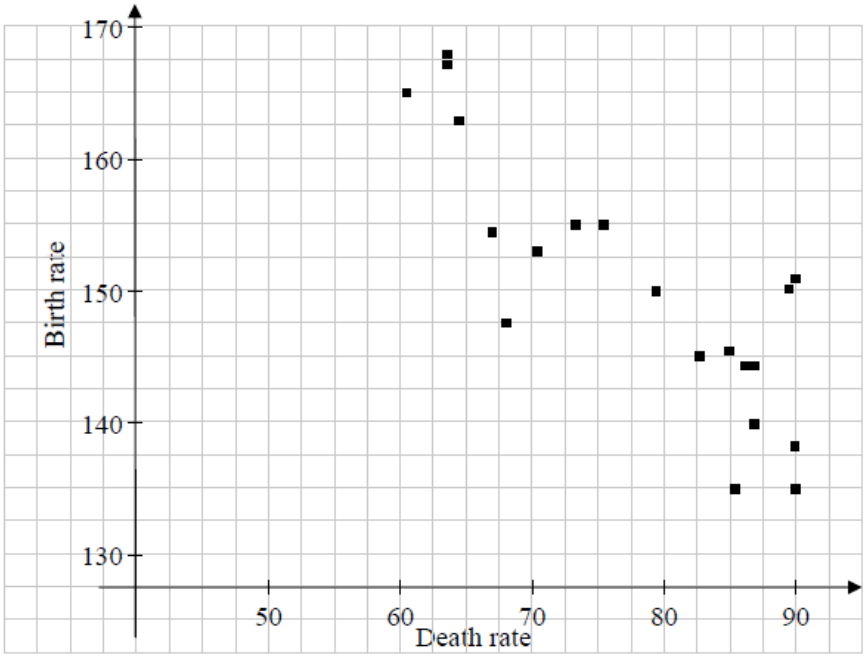
1990 Ratio

2010 Ratio

Prediction

Reason

(i) The birth rate and death rate over the 21 years are plotted against each other in the scatter plot below. The correlation coefficient between the two sets of data is −0·85. Describe the relationship between the two sets of data and suggest a reason why this might be the case.



Q1 2012 Project Maths Paper Two Sample Paper Section A

Answer **all five** questions from this section.

Question 1 **(25 marks)**

The size, mean and standard deviation of four sets of data A, B, C and D are given in this table:

	A	B	C	D
size (N)	1000	100	100	10
mean (μ)	10	100	1000	100
standard deviation (σ)	20	30	20	10

Complete the sentences below by inserting the relevant letter in each space:

(a) The set that contains more numbers than any other is ____ and the set that contains fewer numbers than any other is ____.

(b) On average, the data in set ____ are the biggest numbers and the data in set ____ are the smallest numbers.

(c) The data in set ____ are more spread out than the data in the other sets.

(d) The set that **must** contain some negative numbers is set ____.

(e) If the four sets are combined, the median is most likely to be a value in set ____.

Q6 **2012 Project Maths Paper Two Sample Paper Section B**

Question 6 **(75 marks)**

(a) The students in a 2012 Leaving Certificate class decided to investigate their heights. They measured the height of each student, in centimetres, and the results were as follows:

173	167	180	168	180	175
171	161	164	187	176	160
170	171	167	178	174	149
157	161	176	166	167	172

(i) Construct a stem and leaf plot of the above data.

Key: 14|9 = 149 cm

(ii) Describe the distribution of the data, by making **one** statement about **each** of the three characteristics indicated below.

shape of distribution:

location of data (central tendency / average):

spread of data (dispersion):

(iii) State **one** additional piece of information that you would need in order to decide whether these students are unusually tall?

Q6 2012 Project Maths Paper Two Sample Paper Section B

(b) The students decide to look at the heights of the males and the females in the class separately. The heights are given below:

Males			Females		
173	180	174	167	161	160
175	178	176	157	164	172
180	171	170	168	149	161
187	176	166	167	167	171

(i) Construct a back-to-back stem and leaf plot of the above data.

(ii) State **one difference** and **one similarity** between the two distributions.

Difference:

Similarity:

(c) The heights in 2011 of Irish males born in 1992 are normally distributed with mean 178·8 cm and standard deviation 7·9 cm.

(i) Use the *empirical rule* to complete the following sentence:

"95% of nineteen-year-old Irish men are between _____ and _____ in height."

(ii) Use the empirical rule to make one other statement about the heights of nineteen-year-old Irish men.

(d) The male students in part **(b)** are a *sample*. The males in part **(c)** are a *population*. Is this is a suitable population to compare the sample to? Give a reason for your answer.

Answer: _____

Reason:

(e) Would you say that the males in the class are taller, smaller, or about the same as the population? Use the data to justify your answer.

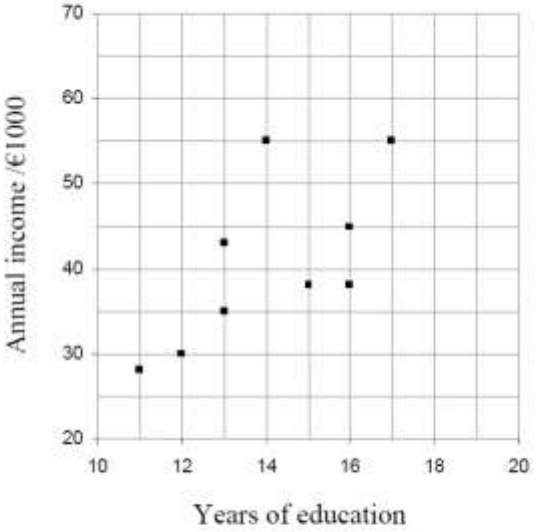
Answer: _____

Justification:

Q6 Part (f) 2012

(f) An economics student wants to find out whether the length of time people spend in education affects how much they earn. The student carries out a small study. She asks twelve adults to state their annual income and the number of years they spent in full-time education. The data are given in the table below, and a partially completed scatter plot is given.

Years of education	Income /€1,000
11	28
12	30
13	35
13	43
14	55
15	38
16	45
16	38
17	55
17	60
17	30
19	58



(i) The last three rows of data have not been included on the scatter plot. Insert them now.

(ii) What can you conclude from the scatter plot?

(iii) The student collected the data using a telephone survey. Numbers were randomly chosen from the Dublin area telephone directory. The calls were made in the evenings, between 7 and 9 pm. If there was no answer, or if the person who answered did not agree to participate, then another number was chosen at random.

Give **one** possible problem that might make the results of the investigation unreliable. State clearly why the issue you mention could cause a problem.

Q 7 2012 Paper Two Ordinary level Section B

Question 7 **(75 marks)**

The following table gives data on new private cars sold in Ireland in each quarter of each year from 2006 to 2011.

New private cars sales								
Number of cars sold						Engine type of cars sold		
Year	January to March	April to June	July to Sept.	October to Dec.	Annual Total	Petrol	Diesel	Other
2006	75 769	54 572	32 873	10 059	173 273	128 634	44 010	629
2007	81 750	57 124	32 418	9 462	180 754	128 346	50 560	1 848
2008	77 441	37 128	27 361	4 540	146 470	92 298	50 283	3 889
2009	27 140	15 225	9 049	3 018	54 432	22 802	30 645	985
2010	34 555	26 806	17 011	6 535	84 907	27 124	53 998	3 785
2011	39 484	29 770	13 467	4 211	86 932	23 246	61 730	1 956

(Source: Central Statistics Office, http://www.cso.ie)

(a) **(i)** Show the *annual total* sales of cars over the six years, using a suitable chart.

(ii) Find the mean number of cars sold per year over the six years.

(iii) Calculate the percentage increase in annual car sales between 2009 and 2011.

(iv) Aoife says that this increase shows car sales are currently going well. Paul says that car sales are currently going badly. He says that sales have fallen by 52% since 2007 and that they are well below average.
Complete the sentences below to give a criticism of each argument.

Aoife's argument does not recognise that…

Paul's argument does not recognise that…

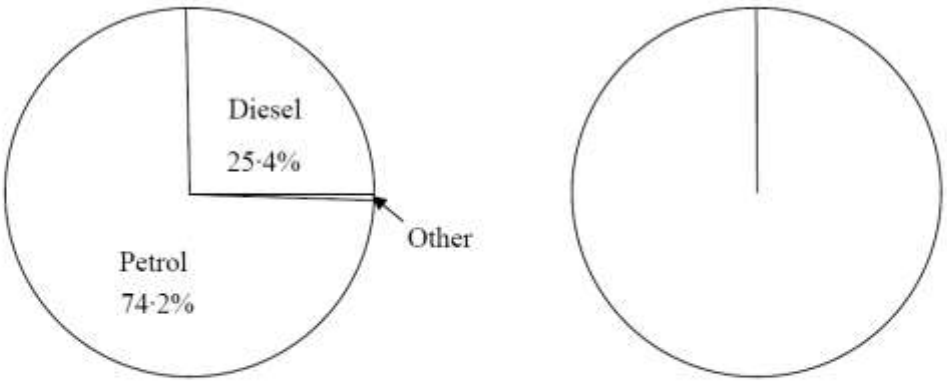
(v) Give a more balanced description of the pattern of car sales over the six years.

(b) **(i)** Describe how the sales of the cars are distributed over the four quarters of each year.

(ii) Suggest a reason for this pattern of sales.

(iii) The sales for the first quarter of 2012 are 36 081.
Find, with justification, an estimate for the total annual sales for 2012.

(c) **(i)** Two pie charts are being used to show the change from 2006 to 2011 in the popularity of petrol and diesel cars. Complete the second pie chart.



2006 2011

(ii) Which of the following statements best describes the change over time in the popularity of diesel cars as a percentage of the total?

A. Diesel cars have suddenly become very popular in the last year or two.

B. Diesel cars have increased very steadily in popularity over the last six years.

C. Diesel cars have become very popular since car sales started to improve.

D. Diesel cars got more popular each year, with an especially big increase in 2009.

E. Diesel cars became popular as car sales fell but have been getting less popular as they rise again.

Write the letter corresponding to the correct answer in the box. ☐

(d) A survey of some of the most popular models of private cars sold in 2011 examined the CO_2 emissions in g/km from diesel engines and petrol engines. The data are as follows:

Diesel engines	Petrol engines
117, 125, 120, 125, 134, 110, 118, 114, 119, 119, 116, 107.	139, 133, 150, 157, 138, 159, 129, 138, 134, 129, 129, 136.

(i) Construct a back-to-back stem-and-leaf plot of the above data.

(ii) Does the information suggest that diesel engines produce lower CO_2 emissions than petrol engines? In your answer you should refer to the stem-and-leaf plot and to an appropriate measure of central tendency.

(iii) Does the information suggest that there is a greater variation in the CO_2 emissions of diesel engines than petrol engines? In your answer you should refer to the stem-and-leaf plot and an appropriate measure of variability.

Question 7 **(75 marks)**

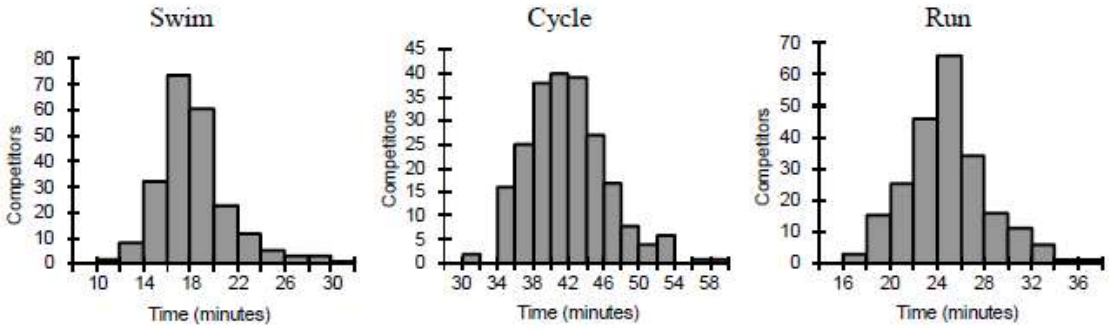
The *King of the Hill* triathlon race in Kinsale consists of a 750 metre swim, followed by a 20 kilometre cycle, followed by a 5 kilometre run.

The questions below are based on data from 224 athletes who completed this triathlon in 2010.

Máire is analysing data from the race, using statistical software. She has a data file with each competitor's time for each part of the race, along with various other details of the competitors.

Lizzie Lee, winner of the women's event

Máire produces histograms of the times for the three events. Here are the three histograms.



(a) Use the histograms to complete the following sentences:

(i) The event that, on average, takes longest to complete is the _____.

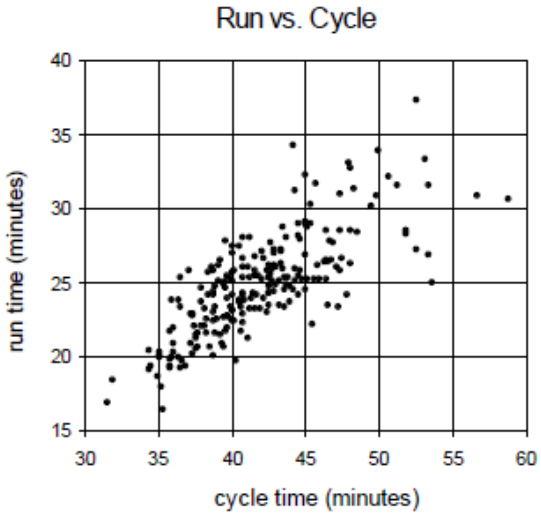
(ii) In all three histograms, the times are grouped into intervals of _____ minutes.

(iii) The time of the fastest person in the swim was between _____ and _____ minutes.

(iv) The median time for the run is approximately _____ minutes.

(v) The event in which the times are most spread out is the _____.

(b) Máire is interested in the relationship between the athletes' performance in the run and in the cycle. She produces the following scatter diagram.



(i) The correlation coefficient between the times for these two events is one of the numbers below. Write the letter corresponding to the correct answer in the box.

A. 0·95
B. 0·77
C. 0·13
D. –0·13
E. –0·77
F. –0·95

(ii) Frank was the slowest person in the run. How many people took longer to complete the cycle than Frank did?

Answer: _____

(iii) Brian did not enter this race. Suppose that he had, and suppose that he completed the cycle in 52 minutes and the run in 18 minutes. Explain why this performance would have been very unusual.

page | running

(c) Máire knows already that the male athletes tend to be slightly faster than the female athletes. She also knows that athletes can get slower as they get older. She thinks that male athletes in their forties might be about the same as female athletes in their thirties. She decides to draw a back-to-back stem-and-leaf diagram of the times of these two groups for the swim. There were 28 females in their thirties, and 32 males in their forties. Here is the diagram:

Female, 30 – 39 years		**Male, 40 – 49 years**
4	13	
	14	9
1 0	15	1 3 4 5 6
9 8 8 7 3 2 2	16	3 4 6 7 7 8
6 4 3 2	17	6 7 7
1	18	0 1 3 8 9
9 6 3 1 0 0	19	0 0 1 2 3 4
	20	3 9 9
3 3 2	21	2 2
4	22	
	23	0
8	24	
	25	
5	26	
	27	
	28	
7	29	

Key: | 14 | 9 means 14·9 minutes.

(i) Describe what differences, if any, there are between the two distributions above.

(ii) Máire drew the diagram because she thought that these two groups would be about the same. Do you think that the diagram would cause Máire to confirm her belief or change it? Give reasons for your answer.