PRIFYSGOL ABERYSTWYTH UNIVERSITY

SEMESTER TWO EXAMINATIONS 2026

Aberystwyth Business School

AB15220: Data Analytics

Time allowed : **2 HOUR(S)**

Answer all questions. All questions carry equal marks.

CALCULATORS ARE PERMITTED (Casio FX83 or FX85 models only)

MAE MODD I FYFYRWYR GYFLWYNO ATEBION I'R PAPUR HWN NAILL AI YN Y GYMRAEG NEU'R SAESNEG

STUDENTS MAY SUBMIT ANSWERS TO THIS PAPER IN EITHER WELSH OR ENGLISH

Once seated, students may complete the front cover of the answer book and the attendance slip ONLY.

DO NOT TURN THE EXAMINATION QUESTION PAPER ONTO THE NEXT PAGE UNTIL INSTRUCTED TO DO SO.

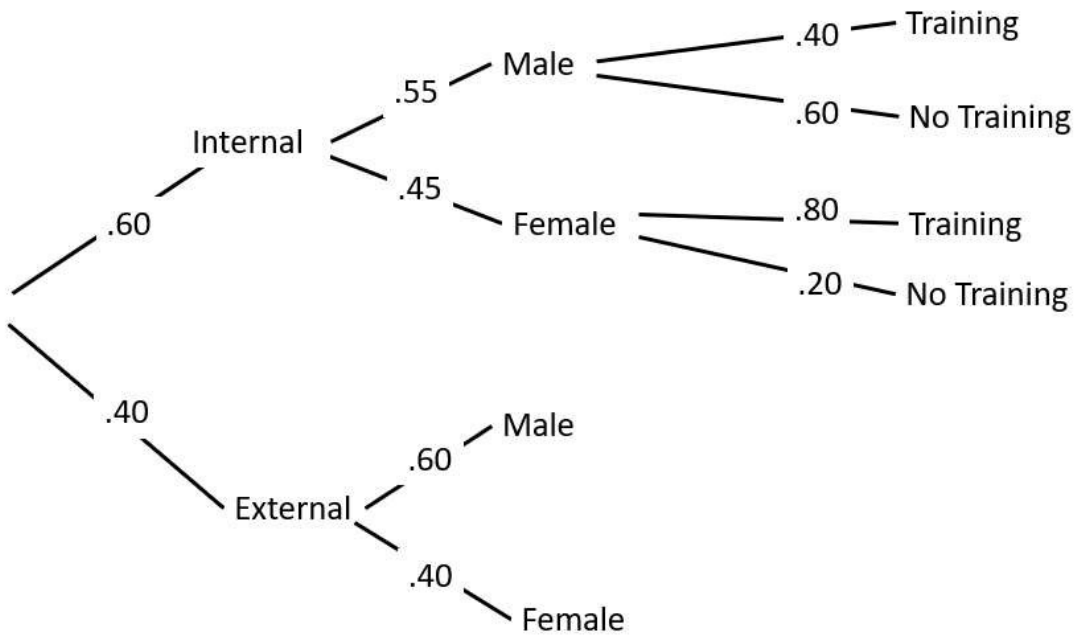
1. A firm advertises for applicants for a management post. The post is open to internal and external applicants. Internal applicants have the option of taking a management training course prior to applying.

   A total of 36 people applied for the post. 60% of the applicants were internal.

   Of the internal applicants, 55% were male and 45% were female. Of the males, 40% had taken the training course whereas 80% of the females had taken the course.

   Of the external applicants, 60% were male, the remainder female. The training course is not available to external applicants.

   The following tree diagram represents this situation:



   If an applicant is selected at random, calculate the probability that they:

   a) Are a male internal applicant.

   b) Are a trained internal female applicant.

   c) Are a female applicant.

   d) Are trained.

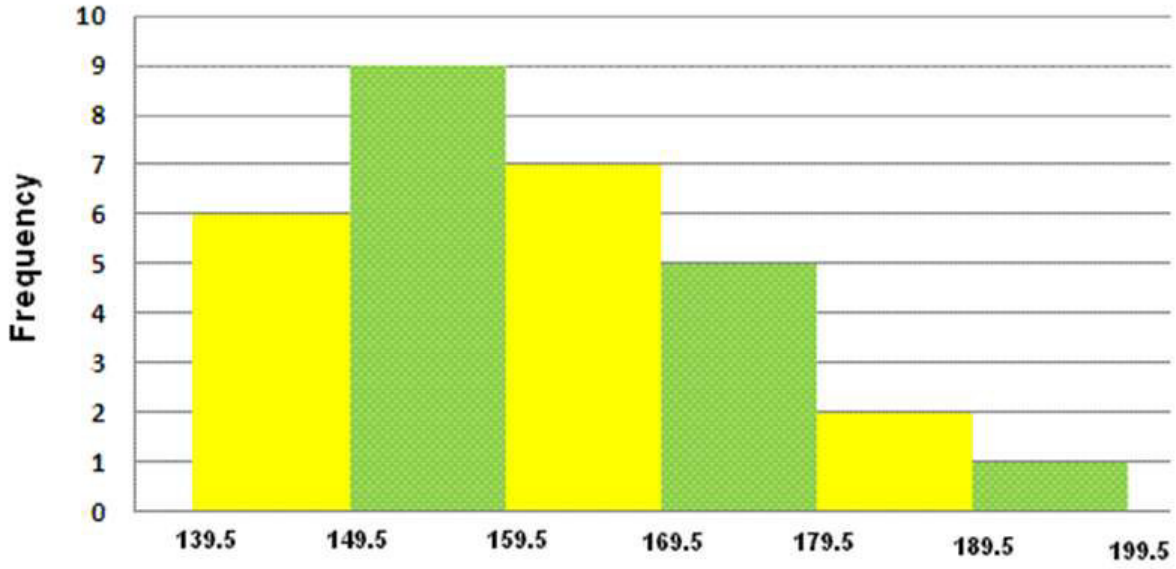
   e) Are not trained

   f) If they are internal, what is the probability that they are trained.

2.

   a) Bar charts and histograms are very similar in appearance. Explain the differences between them; use examples to illustrate your explanation.

   The following histogram shows the heights (in cm) of 30 people:



b) How many people have heights between 159.5 and 169.5 cm?

c) How many people have heights less than 179.5 cm?

d) How many people have heights greater than 169.5 cm?

e) What percentage of people have heights between 149.5 and 179.5 cm?

f) What is the best estimate you can make of the mean height of these 30 people? Comment on its accuracy.

3. Populations are often studied by use of sampling

a) Explain what the terms *population* and *sample* mean in the context of statistics.

b) Why is sampling used so frequently to examine characteristics of a population?

c) What are the main methods of selecting a sample?

d) What are the dangers of sampling?

e) You are thinking of booking a holiday on a French campsite. When you look at on-line reviews of the campsite you find that 22% of people who posted a review described the site "unsatisfactory".
What is your estimate of the probability that you would enjoy a holiday there? Explain the limitations of this estimate.

4.

a) Explain what is meant by the term *Expected Value* in the business context.

b) What are the dangers of using expected values?

c) Given the table of 5 possible outcomes of a business venture below, calculate the expected value; comment on its usefulness.

Profit (£m)	-500	-85	0	180	275
Probability	0.05	0.2	0.15	0.45	0.15

5.

a) Explain what is meant by the term *Correlation*.

b) Discuss the dangers and weaknesses of using this correlation to examine the relationship between two variables.

c) The following table presents data from records of a sample of 15 people. For each person their height (in metres) and their weight (in kilos) are recorded.

Person No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Height (m)	1.8	1.7	1.9	1.8	1.6	1.8	1.5	2.0	1.7	1.6	1.5	1.9	1.8	1.6	1.5
Weight (kg)	70	51	90	89	50	65	40	100	67	77	42	98	90	83	90

d) Calculate the mean, variance and standard deviation of the height of the 15 people.

e) Produce a scatter plot of height and weight.

f) Calculate Pearson's coefficient of correlation of height and weight.

g) Comment on the correlation coefficient.

Formulae

	Sample	Population
Mean	$\bar{x} = \frac{\sum x}{n}$	$\mu = \frac{\sum x}{n}$
Variance	$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$	$\sigma^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}$
Standard Deviation	$s = \sqrt{s^2}$	$\sigma = \sqrt{\sigma^2}$

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

Pearson's Coefficien t of Correlatio n :

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

END OF PAPER