L3 Business Statistics

Overview of content

1. Management Information: The External and Internal Business Environment
   Data collection
   Descriptive statistics

2. Business Planning Models
   Correlation and regression
   Time-based data

3. Risk Management and Business Decision Making
   Probability, including the normal distribution
   Estimation and confidence intervals
   Significance testing
   Chi squared test

4. Quality Assurance and Control
   Quality control

GROUP 1 – SELECT SIX (x4 marks each)

Qu1
Match the descriptions to the correct type of data.

[Primary data] [collected by an organisation itself for its own specific purposes]
[Secondary data] [collected by other organisations and used for a range of different purposes]
[Primary data] [exactly fits the need being current and reliable]
[Secondary data] [cheaper to obtain, and easier to collect, often coming from sources that are not generally accessible]

MATCH ITEMS

Qu2X
It is often difficult to identify all members of the appropriate population from which to obtain data. As result, a list of the relevant populations may be used.

What is this known as?

A	Sampling frame	✔
B	Sample	
C	Population group	
D	Census	

MCQ

Qu3X
Match the correct description to each term.

[Population] [the people or things that could provide the required data]
[Sample] [the people or things that actually provide the data]
[Census] [a means of asking questions to all members of a group]
[Survey] [a means of asking questions to a selected section of a group]

MATCH ITEMS

Qu4X
If a sample has been poorly chosen, results may be inaccurate and unreliable. This is caused by using an unrepresentative sample of the population, What is this known as?

A	Statistical bias	✔
B	Error	
C	Research noise	
D	Survey positioning	

MCQ

Qu5X
Which of the following are essential considerations for an effective questionnaire design?

A	Ask short and easily understood questions	✔
B	Avoid hypothetical questions	✔
C	Include general, open-ended questions	

	throughout	
D	Use conditional clauses where possible	

MRQ 2 marks each but 2 off for each incorrect answer

Qu6
Choose the appropriate responses to complete the table showing methods of data collection.

	Postal questionnaire	Telephone interview	Face-to-face interview
Cost	Low	Moderate	High
Response rate	Low	Moderate	High
Speed	[Slow]	[Fast]	Fast but travel time may be a factor
Quantity if information	[Limited]	Moderate	High
Quality of information	Depends of effectiveness of design	Good	[High]

Options:
[Slow] [Limited] [Fast] [Moderate] [Good] [High] [Poor]

PICK LIST 1 mark each

Qu7
By using suitable mid-points, calculate the standard deviation of the frequency data shown in the table. Give your answer to one decimal place.

Class interval	Frequency
1,100 – 1,200	5
1,200 – 1,300	9
1,300 – 1,400	14
1,400 – 1,500	15
1,500 – 1,600	7

[119.5]

INSERT VALUE

Qu8
Calculate the variance of the following data. Give your answer to two decimal places.

5, 8, 15, 29, 47, 47, 64, 71, 74

[723.25]

INSERT VALUE

Qu9
Calculate the standard deviation of the following data. Give your answer to two decimal places.

5, 8, 15, 29, 47, 47, 64, 71, 74

[26.89]

INSERT VALUE

Qu10
Fill in the blanks in the passage.

For effective statistical sampling, a sample must always be of at least a certain size, as the larger the sample, the more reliance can be placed on the results as truly [representative] of the population of the targeted [group]. Each item should have an [equal] chance of being chosen to provide a [realistic] cross-section of the population.

Options:
[representative] [realistic] [valuable] [group] [team] [equal] [low] [high] [possible]

PICK LIST

GROUP 2 – SELECT SIX (x4 marks each)

Qu11
Fill in the blanks in the passage.

The degree of [correlation] between two variables can be measured and using two main measures, denoted r and R.

The Product Moment Coefficient of Correlation (r) measures the strength of association between two variables; one the [dependent] variable and the other the [independent] variable.

The Rank [Correlation] Coefficient (R) measures the association between two sets of ranked or ordered data

FILL IN BLANKS

Qu12X
Within regression analysis, to find the line of best fit mathematically it is necessary determine the line that minimises which of the following?

A	The total of the squared deviations of the actual observations from the calculated line.]	✔
B	The total of the standard deviations of the actual observations from the calculated line	
C	The total of the deviations of the actual observations from the calculated line.	
D	The difference in the deviations of the actual observations from the calculated line.	

MCQ

Qu13
Which of the following statements about the equation of straight line are true?

A	The general form of the equation of a straight line of y = a + bx	✔
B	The variable a represents the slope of the line	
C	The variable b represents the point of intersection of the line and x-axis	
D	The variable a represents the point of intersection of the line and y-axis	✔

MRQ 2 marks each but 2 marks off for each incorrect answer.

Qu14X

A measure for determining the accuracy of the line of best fit is the coefficient of determination. This is calculated using which of the following?

A	Explained variation / total variation	✔
B	(Correlation coefficient)^2	✔
C	√(Correlation coefficient)	
D	Total variation / Expected variation	

MRQ

Qu15X

The ranking of performance of eight employees in two assessments are shown in the table. Calculate the Spearman rank correlation coefficient. Give your answer to two decimal places.

Employee	**Assessment**	**Assessment**

	1	2
A	2	3
B	7	6
C	6	4
D	1	2
E	4	5
F	3	1
G	5	8
H	8	7

[+0.74]

INSERT VALUE

Qu16X

A survey looking at customers' attitudes to anew food product has found the following results:

Person	Age range (years)	Response comment
A	Under 10	Very good
B	15 – 20	Fair
C	20 – 25	Fair
D	10 – 15	Excellent
E	Over 25	Disliked

Rank the ages and responses and then determine Spearman's rank correlation coefficient. Give your answer to three decimal places.

A	-0.825	✔
B	0.825	
C	-1.175	
D	1.175	

MCQ

Qu17X

The manager of Southern Engineering wants to find out if there is a relationship between production volume and production cost.

Using the data in the table, calculate the coefficient of determination. Give your answer to three decimal places.

Units produced	1,000	2,000	3,000	4,000	5,000	6,000
Production costs ($)	5,000	10,500	15,500	25,000	16,000	22,500

A	0.671	✔
B	0.819	
C	0.648	
D	0.774	

MCQ

Qu18X
Forecasting in business is considered in relation to different time frames. Match the descriptions to the correct time frame.

[Long-term forecasts] [dealing with major plant developments or company reorganisations]
[Medium-term forecasts] [looking at the development and launch of a new product or service]
[Short-term forecast] [dealing with needs of existing products or services]

MATCH ITEMS 1 mark each but 4 marks if all correct

Qu19X
Match the value of a correlation coefficient, r, to the correct interpretation.

[r = 1] [indicates a perfect positive relationship]
[r close to zero but positive] [indicates a weak positive relationship]
[r = 0] [indicates no linear relationship]
[r = -1] [indicates a perfect negative relationship]

MATCH ITEMS 1 mark each

Qu20
The sales figures for a small company are reproduced below for the first nine months of the year 2012.

Month	Sales	Moving average
1	112	
2	122	
3	132	122.0
4	124	[126.0]
5	142	[132.7]
6	170	
7	154	
8	162	[162.0]
9	178	[164.7]

Find the three-month moving averages for months 4, 5, 8 and 9.

INSERT VALUES 1 mark each

GROUP 3 – SELECT SIX (x4 marks each)

Qu21
Which of the following statements about a normal distribution are true?

A	The total area under the curve is directly related to the probability.	✔
B	The area under the curve between two points on the horizontal axis represents the probability that the value of the variable lies between these two points	✔
C	The curve is asymmetrical about the median.	
D	The vertical axis represents a continuous variable.	

MRQ 2 marks each but 2marks off for each incorrect answer

Qu22X
For a normal distribution, the position and shape of the curve varies in which of the following ways?

A	As the standard deviation becomes larger, the curve will get flatter.]	✔
B	As the standard deviation becomes smaller, the curve will get flatter.	
C	As the standard deviation becomes larger, the curve will get extend further either side of the mean.	✔
D	As the standard deviation becomes smaller, the curve will get extend further either side of the mean	

MRQ

Qu23
The standard normal distribution has which of the following characteristics?

A	Mean of zero	✔
B	Standard deviation of 1	
C	Mean of 1	
D	Standard deviation of zero	

MRQ

Qu24
Fill in the blanks in the passage.

The reliability of estimates of [probability] can be stated by indicating limits within which the true mean is expected to lie. These are known as [confidence] intervals. The most common level used is [95]%, which indicates that there is this level of expectation that the true mean lies [within] the calculated limits.

FILL IN BLANKS

Qu25X

A survey of 250 fans at a pop concert showed that 147 were female.

What are the upper and lower limits of the 95% confidence interval for the true percentage of female students?

A	Lower limit = 52.7%]	✔
B	Upper limit = 64.9%]	✔
C	Upper limit = 68.9%	
D	Lower limit = 58.8%	

MRQ

Qu26X

The number of standard errors for a given level of significance is different for one-tailed tests compared to two-tailed tests.

Place the values for the most common levels of 1% and 5% significance into the table shown.

	Number of standard errors	
	Two-tailed test	One-tailed test
5% level of significance	[1.96]	[1.65]
1% level of significance	[2.58]	[2.33]

DRAG AND DROP

Qu27

Which of the following statements about the possible results when we test the hypothesis, H_0, are true?

A	The true hypothesis is accepted – a correct decision.	✔
B	The false hypothesis is rejected – a correct decision.	✔
C	The true hypothesis is rejected – an incorrect decision. This is a Type II error.	

D	The false hypothesis is accepted – an incorrect decision. This is a TYPE I error.	

MRQ 2 marks each but 2marks off for each incorrect answer

Qu28
Fill in the blanks in the passage.

The t distribution can be used to test for differences in population means if the samples are [small]. In such cases, the two sampled populations should be near normally [distributed] and the two [standard deviations] should not be significantly [different.]

FILL IN BLANKS

Qu29X
A taxi driver claims that an equal number of passengers use his taxi cab each weekday. To test this hypothesis, he records the number of passengers for a week.

These are shown below:

Monday – 50 passengers
Tuesday – 60 passengers
Wednesday – 40 passengers
Thursday – 47 passengers
Friday – 53 passengers

Using the chi-squared test, calculate the p-value for his data. Give your answer to three decimal places.

[0.359]

INSERT VALUES

Qu30X
Using a chi-squared test, a beautician records four different categories of customers over a period of a week and suggests an equal number for each category. The recorded information shows the following:

Category a – 22 customers
Category B – 20 customers
Category C – 23 customers
Category D – 35 customers

What is the p-value for this data? Give your answer to three decimal places.

[0.137]

INSERT VALUE

GROUP 4 – SELECT SIX (x4 marks each)

Qu31X

Control charts are used for which of the following?

A	Depict whether sampled products meet their intended specification	✔
B	Analyse specific attributes of a product to show the degree to which they vary from specifications	✔
C	Indicate the profitability of r sampled product	
D	Randomly select products for testing	

MRQ

Qu32X

In an x-bar chart used to depict quality control, which of the following does the y-axis track?

A	Variance of the tested attribute in a univariate way	✔
B	Variance of the tested attribute in a multivariate way	✔
C	The degree to which all the products vary from the specification	
D	The variation of product quality	

	benchmarks	

MRQ

Qu33X
Match the type of quality control chart to the correct description.

[x-bar chart] [The chart tracks the degree to which the variance of the tested attribute is acceptable.]
[R chart] [The chart is used to monitor the variation of a process based on small samples taken at specific time.]
[U chart] [The chart is used when the number of defects per unit and the sample size vary.]
[Np chart] [The chart is used when the number of variations and the sample size are fixed.]

MATCH ITEMS 1 mark each

Qu34
Fill in the blanks in the passage.

Control charts contain control [limits], calculated from statistical methods, with sample values checked against them. The charts provide a [visual] means of distinguishing between variables due to [inherent] causes - process in control - and the variability due to [special] causes -process out of control).

FILL IN BLANKS 1 mark each

Qu35X
There are two levels of control limit. Which of the following are the correct limits?

A	Warning limit	✔
B	Action limit	✔
C	Activation limit	
D	Determined limit	

MRQ

Qu36
A process may have a satisfactory average but the variation about the mean may cause concern.

What should be monitored in these circumstances?

A	The range of values	✔
B	The standard deviation	
C	The median value	
D	The outlier values	

MCQ

Qu37
When using control charts. It is usually necessary to define which of the following?

A	The upper warning limit and upper action line	✔
B	Both the upper warning limit and lower warning limit	
C	Only the upper warning limit	
D	The upper action line and the lower action line	

MCQ

Qu38
Fill in the blanks in the following passage.

Control charts in [quality] control are powerful tools for providing valuable insight into process [performance] by tracking data over time, enabling businesses to identify [variations], prevent potential issues arising, detect [trends] and supporting data-driven [decisions].

FILL IN BLANKS

Qu39X
Using mean deviation as a measure of statistical dispersion is limited for which of the following reasons?

A	It does not differentiate between positive and negative differences	✔
B	It is only effective for large differences	
C	It requires linking to the median and mode values	
D	It does not take account of quartile values	

MCQ

Qu40X

Which of the following is not a component of time series analysis?

A	Magnitude of differences	✔
B	Trend	
C	Cyclical fluctuations	
D	Seasonal variations	

MCQ

MCQ GROUP – SELECT SIX (x2 marks each)

Q41X
Most data are collected from a representative group of people or things and used to estimate the characteristics of a wider range of all people and things.

Such a group is known as which of the following?

A	Sample	✔
B	Survey	

C	Segment	
D	Section	

MCQ

Qu42X
What type of graph would you use when dealing with unequal class intervals?

A	Histogram	✔
B	Pie chart	
C	Block graph	
D	Box plot	

MCQ

Qu43
Calculate the mean of the following data.

5, 8, 15, 29, 47, 47, 64, 71, 74

A	40	✔
B	47	
C	38	
D	35	

MCQ

Qu44
What is the term given to the changes in one factor or variable that arise from movements in the other variable?

A	Causal relationship	✔
B	Regressive relationship	
C	Direct proportional relationship	
D	Inverse proportional relationship	

MCQ

Qu45X
The equation for the regression line of y ($000) on, x ($000) passes through the following pairs of (x, y) coordinates (1, 3) and (2,4).

What is the regression equation?

A	y = x + 2	✔
B	y = x - 2	
C	y = 2x + 1	
D	y = 2x - 1	

MCQ

Qu46
What percentage of the total area under the standard normal distribution curve is contained within the region between +1 and -1 standard deviations? Give your answer to one decimal place.

A	68.3	✔
B	29.1	
C	58.2	
D	34.2	

MCQ

Qu47X
A greengrocer receives an order of melons. The weight of the melons is normally distributed, with a mean of 800g and a standard deviation of 10g.

What percentage of melons would be expected to weigh between 810g and 812g? Give answer to two decimal places.

A	4.36	✔
B	95.64	
C	6.68	
D	93.32	

MCQ

Qu48
The chi-squared test is often referred to as which of the following test?

A	Goodness of fit test	✔
B	Observed frequency test	
C	Yates correction	
D	Null Hypothesis test	

MCQ

Qu49
A greengrocer receives an order of melons. The weight of the melons is normally distributed, with a mean of 800g and a standard deviation of 10g.
What percentage of melons would be expected to weigh more than 815g? Give answer to two decimal places.

A	6.68%	✔
B	93.32%	
C	4.36%	
D	95.64%	

MCQ

Qu50X
Which statistical process is an extension of hypothesis testing and is used when the actual, observed distribution is compared with a hypothesis?

A	Chi-squared test	✔
B	t test	
C	Two-tailed test	
D	One-tailed test	

MCQ

TOTAL MARKS AVAILBLE = 108