A Level Statistics
Practice Test 4: Data Collection

Instructions:
Answer all questions. Show your working clearly.
Calculators may be used unless stated otherwise.
Draw diagrams where appropriate to illustrate your solutions.
Time allowed: 3 hours

Section A: Time Series and Trend Analysis [25 marks]

1. **[8 marks]** Define and explain time series concepts:

   (a) Define a time series and explain its components.

   (b) Distinguish between trend, seasonal variation, and random variation.

   (c) Explain what is meant by cyclical patterns in time series data.

   (d) Describe moving averages and their purpose in trend analysis.

2. **[10 marks]** A company's quarterly sales data (£000s) over 3 years is shown:

Year	Q1	Q2	Q3	Q4
2021	120	145	180	95
2022	135	160	195	110
2023	150	175	210	125

   (a) Plot the time series data on a graph.

   (b) Calculate 4-point moving averages for the data.

   (c) Identify any seasonal patterns in the data.

   (d) Estimate the underlying trend from your moving averages.

   (e) Calculate the seasonal effects for each quarter.

   (f) Predict the sales for Q1 and Q2 of 2024.

   (g) Comment on the reliability of your predictions.

3. **[7 marks]** A weather station records daily temperature data. The analyst notices: - Overall warming trend over decades - Regular seasonal patterns (winter/summer cycles) - Daily fluctuations due to weather systems - Occasional extreme values due to unusual weather events

   (a) Classify each observation as trend, seasonal, cyclical, or random variation.

   (b) Suggest appropriate methods to isolate each component.

   (c) Explain how you would handle the extreme values in analysis.

Section B: Correlation and Relationships [30 marks]

4. **[12 marks]** Define and explain correlation concepts:

(a) Define correlation and distinguish it from causation.

(b) Explain positive, negative, and zero correlation with examples.

(c) Describe the correlation coefficient and its range of values.

(d) Explain what is meant by "spurious correlation" with an example.

5. **[18 marks]** A study examines the relationship between advertising spend (£000s) and monthly sales (£000s) for 10 months:

Month	Advertising (x)	Sales (y)
1	12	85
2	15	92
3	8	76
4	20	105
5	18	98
6	10	82
7	25	118
8	14	89
9	22	112
10	16	95

Given: $\sum x = 160$, $\sum y = 952$, $\sum x^2 = 2740$, $\sum y^2 = 91956$, $\sum xy = 15764$

(a) Calculate the means $\bar{x}$ and $\bar{y}$.

(b) Draw a scatter diagram for this data.

(c) Calculate the product-moment correlation coefficient.

(d) Interpret the strength and direction of the correlation.

(e) Identify any potential outliers on your scatter diagram.

(f) Discuss whether the correlation implies causation in this context.

(g) Suggest other factors that might influence sales.

(h) Comment on the appropriateness of using this relationship for prediction.

Section C: Data Collection Planning and Evaluation [35 marks]

6. **[15 marks]** A local authority wants to investigate public transport usage patterns to improve services. Design a comprehensive data collection plan:

(a) Define the research objectives and target population.

(b) List the key variables you would collect data on.

(c) Choose appropriate sampling methods for different aspects of the study.

(d) Design data collection instruments (surveys, observation sheets, etc.).

(e) Identify potential sources of bias and propose solutions.

7. **[20 marks]** Evaluate the following research scenarios for methodology and data quality:

Scenario 1: A supermarket chain analyzes customer purchasing patterns using loyalty card data from 50,000 customers over 6 months.

Scenario 2: A university surveys student satisfaction by emailing questionnaires to all 15,000 students, receiving 2,000 responses (13.3

Scenario 3: A traffic planning study monitors vehicle counts at 20 locations using automated sensors for one week in July.

Scenario 4: A health study tracks fitness tracker data from 500 volunteers who downloaded a specific app over 3 months.

For each scenario, analyze:

(a) Strengths and limitations of the data source.

(b) Potential biases in the sample or methodology.

(c) Representativeness of the data collected.

(d) Reliability and validity concerns.

(e) Suggested improvements to enhance data quality.

Answer Space

Use this space for your working and answers.

Formulae and Key Concepts

Moving Averages:

4-point moving average $= \frac{\text{Sum of 4 consecutive values}}{4}$

Centered moving average for even number of points requires averaging pairs

Seasonal Adjustment:

Seasonal effect = Actual value - Trend value

Deseasonalized value = Actual value - Seasonal effect

Product-Moment Correlation Coefficient:

$$r = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}}$$

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

where $S_{xy} = \sum xy - n\bar{x}\bar{y}$, $S_{xx} = \sum x^2 - n\bar{x}^2$, $S_{yy} = \sum y^2 - n\bar{y}^2$

Correlation Interpretation:

$|r| = 1$: Perfect correlation

$0.8 \leq |r| < 1$: Strong correlation

$0.5 \leq |r| < 0.8$: Moderate correlation

$0.3 \leq |r| < 0.5$: Weak correlation
$|r| < 0.3$: Very weak/no correlation

Time Series Components:
Trend: Long-term direction
Seasonal: Regular patterns within year
Cyclical: Longer-term fluctuations
Random: Irregular, unpredictable variations

Research Design Principles:
Clear objectives and hypotheses
Appropriate target population
Representative sampling methods
Valid and reliable measurement instruments
Ethical considerations

Data Quality Criteria:
Accuracy: Freedom from systematic errors
Precision: Level of detail and consistency
Completeness: Minimal missing data
Timeliness: Data collected when relevant
Relevance: Fit for intended purpose

Common Research Biases:
Self-selection bias: Volunteers differ from population
Temporal bias: Time period affects results
Measurement bias: Systematic measurement errors
Survivorship bias: Only successful cases included
Confirmation bias: Seeking supporting evidence only

END OF TEST

Total marks: 90

For more resources and practice materials, visit:
stepupmaths.co.uk