A Level Statistics
Practice Test 2: 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: Data Types and Classification [25 marks]

1. **[8 marks]** Define and distinguish between different types of data:

   (a) Define primary data and secondary data with examples.

   (b) Explain the difference between ordinal and nominal data.

   (c) Define grouped and ungrouped data.

   (d) Give two examples each of ordinal and nominal qualitative data.

2. **[10 marks]** Classify the following variables and explain your reasoning:

   (a) Speed of cars on a motorway

   (b) Educational qualifications (GCSE, A Level, Degree, Masters)

   (c) Mobile phone brands

   (d) Daily rainfall measurements

   (e) Number of siblings

   (f) Customer satisfaction ratings (Very Poor, Poor, Average, Good, Excellent)

   (g) Types of music genres

   (h) Ages of people to the nearest year

   (i) Postcode areas

   (j) Weekly earnings of employees

3. **[7 marks]** A sports club is conducting research on member fitness levels. They collect:

   - Member's age (in years)
   - Fitness rating (Beginner, Intermediate, Advanced, Elite)
   - Weekly training hours
   - Sport preference

- Monthly membership fee paid

(a) Classify each variable as qualitative or quantitative.

(b) For quantitative variables, determine if they are discrete or continuous.

(c) Suggest the most appropriate data collection method for each variable.

Section B: Advanced Sampling Techniques [30 marks]

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

(a) Quota sampling: Describe the method and compare it with stratified sampling.

(b) Multi-stage sampling: Explain when this approach is necessary.

(c) Snowball sampling: Describe the technique and its applications.

(d) Purposive sampling: Explain when this non-probability method is appropriate.

5. [**18 marks**] A university has 15,000 students across 5 faculties: Science (4,500), Arts (3,200), Business (2,800), Engineering (2,500), Medicine (2,000). A study requires 300 students.

(a) Design a proportional stratified sample using faculties as strata. Calculate sample sizes.

(b) If systematic sampling is used with a starting point of 7, describe the selection process.

(c) For cluster sampling using tutorial groups of 25 students each, determine how many clusters to select.

(d) Design a quota sampling scheme based on year of study (assuming equal distribution across 4 years).

(e) Explain why multi-stage sampling might be necessary for this population.

(f) Compare the representativeness of each sampling method for this scenario.

Section C: Data Quality and Survey Design [35 marks]

6. [**15 marks**] Analyze sources of error and bias in data collection:

(a) Define measurement error and explain how it affects data quality.

(b) Describe interviewer bias and methods to reduce it.

(c) Explain leading questions and provide examples.

(d) Define recall bias and situations where it commonly occurs.

(e) Describe social desirability bias and its impact on survey responses.

7. [**20 marks**] Evaluate these data collection scenarios for bias and quality issues:

Scenario 1: A survey about exercise habits is conducted via smartphone app, targeting users aged 16-65.

Scenario 2: A market research study about luxury car preferences interviews people at an expensive shopping center.

Scenario 3: A health questionnaire asks: "How many cigarettes do you smoke per day?" with options 0, 1-5, 6-10, 11-20, 21+.

Scenario 4: An education survey is conducted by emailing university students during exam period with a 15

For each scenario:

(a) Identify potential sources of bias or error.

(b) Assess the impact on data reliability and validity.

(c) Suggest specific improvements to enhance data quality.

(d) Recommend alternative data collection approaches.

Answer Space

Use this space for your working and answers.

Formulae and Key Concepts

Sampling Methods:

Simple Random: Every member has equal chance of selection
Systematic: Select every kth member where k = $\frac{N}{n}$
Stratified: $n_i = \frac{N_i}{N} \times n$ (proportional allocation)
Cluster: Select entire groups randomly

Data Classification:

Qualitative: Nominal (categories), Ordinal (ranked categories)
Quantitative: Discrete (countable), Continuous (measurable)
Primary: Collected firsthand, Secondary: Previously collected

Sample Size Considerations:

Larger samples → Lower sampling error
Cost vs. precision trade-off
Diminishing returns beyond certain size

Common Biases:

Selection bias: Non-representative sampling
Response bias: Systematic error in responses
Non-response bias: Differences between responders/non-responders
Measurement bias: Systematic measurement errors
Social desirability bias: Responses to appear favorable

Survey Design Principles:

Clear, unambiguous questions
Avoid leading or loaded questions
Appropriate response scales
Logical question ordering
Pilot testing before full implementation

Data Quality Criteria:
Accuracy: Freedom from error
Precision: Level of detail
Reliability: Consistency of measurement
Validity: Measures what it claims to measure
Completeness: Minimal missing data

END OF TEST

Total marks: 90

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