

Data Types and Samples

1. A survey was carried out at a bank. Some numerical data was collected. State whether the following variables are continuous or discrete.
 - (a) The number of customers who entered the bank between 10:00 am and 11:00 am.
 - (b) The time taken to serve each customer.
 - (c) The total amount of money withdrawn on that day.
 - (d) The number of employees working on the day that the survey was carried out.
2. Complete the table by naming the type of data formed by each of the measurements.

| Measurement | Type of Data |
|--|-------------------------|
| Number of births each month during 2009 | Discrete numerical data |
| Concentration of volcanic ash particles in the atmosphere | |
| Weights of all beef slaughtered in Ireland during 2008 | |
| number of aeroplanes flying out of Shannon every day | |
| Number of pages in the books kept in the school library | |
| The grades of the students in your class in Junior Certificate | |

3. A sample of 500 households in Dublin was selected and several questions were asked of the householders. Which of the following is not correct?
 - (a) The total household income is ordinal categorical data.
 - (b) The number of persons in the household is discrete data.
 - (c) Socioeconomic status (coded as 1 = low income, 2 = middle income and 3 = high income) is nominal categorical data.

Now correct the incorrect statements.

4. Classify each variable as qualitative or quantitative.
 - (a) Colours of the automobiles in the faculty parking lot.
 - (b) Number of desks in classrooms.
 - (c) Classification of children in a day care centre (infant, toddler, pre-school).
 - (d) Weights of fish caught in Lake Emilie.
 - (e) Number of pages in statistics textbook.
 - (f) Capacity (in gallons) of water in selected dams.
 - (g) Number of off-road vehicles sold in the United States

5. Classify each variable as discrete or continuous.
- (a) Number of loaves of bread baked each day at a local bakery.
 - (b) Water temperature of the saunas at the given health spa.
 - (c) Incomes of single parents who attend a community college.
 - (d) Lifetimes of batteries in a tape recorder.
 - (e) Weights of newborn infants at a given hospital.
 - (f) Capacity (in gallons) of water in swimming pools.
 - (g) Number of pizzas sold last year in the United States.
6. Classify each as nominal, ordinal, interval or ratio level.
- (a) Number of exams given in a statistics course.
 - (b) Ratings of word-processing programs as user friendly.
 - (c) Temperatures of a sample of automobile tyres tested at miles per hour for six minutes.
 - (d) Weights of suitcases on a selected commercial airline flight.
 - (e) Classification of students according to major field.
7. Classify each variable as discrete or continuous.
- (a) The time it takes to drive to work.
 - (b) The number of credit cards a person has.
 - (c) The number of employees working in a large department store.
 - (d) the amount of a drug injected into a rat. the amount of sodium contained in a bag of potatoe chips.
 - (e) the number of cars stolen each week in a large city.
8. Read the article below and answer the questions that follow.
- The recent study reached the following conclusion: "A heart patient overly dependent on a spouse may have a harder time making necessary lifestyle changes in diet and exercise."
- (a) Do you agree or disagree with this statement? State your reasons.
 - (b) Comment on how this study's conclusion might have been reached?
 - (c) What are the variables used in the study and how might they have been defined?
 - (d) How might the researcher measure the variables?
 - (e) What would the population be for this study?
 - (f) What factors other than dependence on a spouse might have influenced the results of the study?
 - (g) Do you think the gender of the patients would make a difference in the results? Why or why not?

HEART TO HEART

There is a flurry of new research findings about how heart heals the heart.

A heart patient overly dependent on a spouse may have a harder time making necessary lifestyle changes in diet and exercise.

For women heart attack victims, spousal support is critical - but hard to come by. "The family sometimes feels abandoned," explains Martin Sullivan, "and they don't want the woman to take time out of her duties as a wife and mother to make important lifestyle changes. Women are more willing to change for men."

For men, a heart may shatter the sole definition of self (as family provider). The introduction of larger concepts of the self is therapeutic.

Patients who feel a sense of self efficacy and control over their disease do better than those who don't.

Depression and anxiety affect pain perception and the capacity to function in the face of medical symptoms.

In a study at Stanford University, behavioural counselling after heart attack, especially for hard-driving type A individuals, lowered the rate of recurrent heart attacks by 45 percent - the same as the most powerful prescription drugs.

Such findings have led Martin Sullivan to introduce innovative techniques at the Duke Centre. These include a program known as PAIRS (Practical Application of Intimacy Relationship Skills), which teaches couples healthy interactive skills, and a meditation program that teaches patients to freeze-frame a moment in time and look at the emotional content of what they are experiencing. Says Sullivan of the Duke Centre's work, "We're trying to take the best of everything."

9. An electronics company manufactures four different types of components. The number of each type manufactured per day is given in the table below. The company has a policy of randomly selecting 70 components a day for quality checks.

| Component | A | B | C | D |
|-----------|-----|-----|----|-----|
| Number | 200 | 350 | 50 | 100 |

- (a) How many of each type of component should be selected
- (b) What type of sampling is being done here?
10. What type of sampling is being used in these cases?
- (a) The Gardai set up a checkpoint to make sure drivers have car insurance. They stop every fifth driver.
- (b) On the day of a general election, RTE organise an exit poll. They randomly select 100 polling stations and all voters are surveyed as they leave the station.
- (c) A researcher in the Department of Social Welfare has partitioned all Dublin adults into the following categories: unemployed, employed full-time and employed part-time. She is surveying 50 people from the first category, 200 people from the second category and 2 from the third category.

11. What type of sampling is being used in these cases?
 - (a) In an MRBI poll of 1,000 adults, people were selected by using a computer to randomly generate telephone numbers that were then called.
 - (b) An engineer tests his new design on family members.
 - (c) The school principal decides to interview all students in the school to get their opinion on some school matter.
12. Explain why a large sample may not necessarily be a good sample.
13. After the World Trade Center in New York was destroyed, an internet poll asked the question, 'Should the World Trade Center be rebuilt?'

Of the 1.3 million respondents, 769,000 said 'Yes', 287,000 said 'No', and 249,000 were undecided. Given that this sample is very large, should it be considered representative of the views of the population of the USA? Explain.
14. Suppose we want to estimate the number of students who watch *Strictly Come Dancing* in school with a population of 1,000 students. We randomly select 50 students. Our sample happens to contain 40 girls and 10 boys.
 - (a) Is it possible that this sample could be biased? Explain.
 - (b) How could potential bias be eliminated in this particular case?
15. Classify each sample as random, systematic, stratified or cluster.
 - (a) In a large school district, all teachers from two buildings are interviewed to determine whether they believe the students have less homework to do now than in previous years.
 - (b) Every seventh customer entering a shopping mall is asked to select his or her favourite store.
 - (c) Nursing supervisors are selected using random numbers in order to determine annual salaries.
 - (d) Every hundredth hamburger manufactured is checked to determine its fat content.
 - (e) Mail carriers of a large city are divided into four groups according to gender (male or female) and according to whether they walk or ride on the routes. Then 10 are selected from each group and interviewed to determine whether they have been bitten by a dog in the last year.
16. A researcher divided subjects into two groups according to gender and then selected members from each group for his sample. What sampling method was the researcher using?
 - (a) Cluster
 - (b) Random
 - (c) Systematic
 - (d) Stratified