Solution *Section* 1.3 – Sampling Methods

Exercise

A student of the author collected measurements of arm lengths from her family members. Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Convenience, since the sample is those who happened to be in the student's family.

Exercise

On the day of the last presidential election, ABC News organized an exit poll in which specific polling stations were randomly selected and all voters were surveyed as they left the premises. Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Cluster, since all the voters at randomly selected polling stations were surveyed.

Exercise

The author was an observer at a town of Poughkeepsic Police sobriety checkpoint at which every fifth driver was stopped and interviewed. (He witnessed the arrest of a former student.) Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Systematic, since every 5^{th} driver was stopped

Exercise

You observed professional wine taster working at the Consumer's Union testing facility in NY. Assume that a taste test involves three different wines randomly selected from each of five different wineries. Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Stratified, since the population of wines was subdivided into 5 different subgroups (wineries), and then samples were drawn from each subgroup.

Exercise

The U.S. Department of Corrections collects data about returning prisoners by randomly selecting five federal prisons and surveying all of the prisoners in each of the prisons. Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Cluster, since all the prisoners at five randomly selected prisons were surveyed.

Exercise

You instructor surveyed all of his students to obtain sample consisting of the number of credit cards students possess. Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Convenience, since the sample is those who happen to be students of the author.

Exercise

In a study of college programs, 820 students are randomly selected from those majoring in communications, 1463 students are randomly selected from those majoring in business, and 760 students are randomly selected from those majoring in history. Identify what type is used: random, systematic, convenience, stratified, or cluster.

Solution

Stratified, since the population was subdivided into 3 different subgroups, and then samples were drawn from each subgroup. The population of interest appears to be only communications, business and history majors – most likely because the programs being studied involved only those majors.

Exercise

Pharmacists typically fill prescriptions by scooping a sample of pills from a larger batch that is in stock. A pharmacist thoroughly mixes a large batch of Lipitor pills, then selects 30 of them. Does this sampling plan result in a random sample? Simple random sample? Explain.

Solution

Yes, it is a random sample because each pill has an equal chance of being selected.
Yes, it is a simple random sample because each sample of size 30 has an equal chance of being selected.

Exercise

A quality control engineer selects every $10{,}000^{th}$ M&M plain candy that is produced. Does this sampling plan result in a random sample? Simple random sample? Explain.

Solution

Whether the sample is a random depends on how the first selection is made. If the engineer chooses the first one at random from 1 to 10,000 and every $10{,}000^{th}$ one thereafter, then every M&M has an equal chance of being selected (namely 1 in 10,000) and the sample is a random sample. If the engineer determines to start with #1 and choose every $10{,}000^{th}$ one thereafter, then some M&M's have no chance of being selected (e.g., #2) and the sample is not a random sample.

No, no matter how the first selection is made the sample will not be a simple random sample of size n. All possible grouping of size n are not possible – any grouping containing #1 and #1, for example, could not occur.

Exercise

NBC News polled reactions to the last presidential election by surveying adults who were approached by a reporter at a location in N.Y. City. Does this sampling plan result in a random sample? Simple random sample? Explain.

Solution

No, this is not a random sample of all New Yorkers because persons who did not visit the location had not chance of being selected.

No it is not a simple random because it is not a random sample.

Exercise

A classroom consists of 36 students seated in six different rows, with six students in each row. The instructor rolls a die to determine a row, then rolls the die again to select a particular student in the row. This process is repeated until a sample of 6 students is obtained. Does this sampling plan result in a random sample? Simple random sample? Explain.

Solution

Yes, the results in a random sample because each student has an equal chance of being selected.

Yes, this results in a simple random sample of size 6 because each possible grouping of size 6 has an equal chance to occur.

Exercise

A computer company employs 100 software engineers and 100 hardware engineers. The personnel manager randomly selects 20 of the software engineers and 20 of the hardware engineers and questions them about career opportunities within the company. Does the sampling plan result in a random sample? Simple random sample? Explain.

Solution

Yes, the sample is random because all employees have the same chance of being selected.

No, it is not a simple random sample because some samples are possible, such as a sample consisting of 30 software engineers and 10 hardware engineers.

Exercise

A polling company obtains an alphabetical list of names of voters in a precinct. They select every 20^{th} person from the list until a sample of 100 is obtained. They then call these 100 people. Does the sampling plan result in a random sample? Simple random sample? Explain.

Solution

No. The sample is not random because not all voters have the same chance of being selected.
No, the second person on the list has no chance of being selected. It is not a simple random sample because some samples are not possible, such as a sample containing the second person on the list.

Exercise

What is an inherent zero? Describe three examples of data sets that have inherent zeroes and three that do not.

Solution

An inherent zero is a zero that implies none.

1. Average age of college students in years
2. Maximum wing speed during a hurricane
3. Average monthly precipitation in inches

Exercise

What is the different between a random sample and a simple random sample?

Solution

With a random sample, each individual has the same chance of being selected. With a simple random sample, all samples of the same size have the same chance of being selected.

Exercise

Determine whether the statement is true or false. If false, rewrite it as a true statement

a) In a randomized block design, subjects with similar characteristics are divided into blocks, and then, within each block, randomly assigned to treatment groups.
b) Using a systematic sample guarantees that members of each group within a population will be sampled.
c) The method for selected a stratified sample is to order a population in some way and then select members of the population at regular intervals.

Solution

a) True
b) False. Using a stratified sample guarantees that members of each group within a population will be sampled.
c) False. The method for selecting systematic sample is to order a population in some way and then select members of the population at regular intervals.

Exercise

Which method of data collection should be used to collect data for the following study

a) A study of the health of 148 kidney transplant patients at a hospital.
b) A study of the effect on the taste of a snack food made with a sugar substitute
c) A study of how fast a virus would spread in a herd of cattle.

Solution

a) Census
b) Experiment
c) Simulation

Exercise

A pharmaceutical company wants to test the effectiveness of a new allergy drug. The company identifies 250 females 30-35 years old who suffer from severe allergies. The subjects are randomly assigned into two groups. One group is given the new allergy drug and the other is given a placebo that looks exactly like the new allergy drug. After six months, the subjects' symptoms are studied and compared

a) Identify the experimental units and treatment used in this experiment.
b) Identify a potential problem with the experiment design being used and suggest a way to improve it.
c) How could this experiment be designed to be a double-blind?

Solution

a) The experiment units are the 30- to 35- year old females being given the treatment is the new allergy drug.
b) There may be a bias on the part of the researcher knows which patients were given the real drug.
c) The study would be a double-blind study if both the researcher and the patient did not know which patient received the real drug or the placebo

Exercise

What type of sampling is used: random, stratified, convenience, cluster, systematic, in the following?

a) To estimate the percentage of defects in a recent manufacturing batch, a quality-control manager at Intel selects every 8^{th} chip that comes off the assembly line starting with the 3^{rd} until she obtains a sample of 140 chips.
b) To determine the prevalence of human growth hormone (HGH) use among high school varsity baseball players, the State Athletic Commission randomly selects 50 high schools. All members of the selected high schools' varsity baseball teams are tested for HGH.
c) To determine customer opinion of its boarding policy. Southwest Airlines randomly selects 60 lights during a certain week and surveys all passengers on the flights.
d) A member of Congress wishes to determine her constituency's opinion regarding estate taxes. She divides her constituency into three income classes: low-income households, middle-income households, and upper-income households. She then takes a simple random sample of households from each income class.

e) In an effort to identify whether an advertising campaign has been effective, a marketing firm conducts a nationwide poll by randomly selecting individuals from a list of known users of the product.

f) A radio station asks it listeners to call in their opinion regarding the use of U.S forces in peacekeeping missions.

g) A farmer divides his orchard into 50 subsections, randomly selects 4, and samples all the trees within the 4 subsections to approximate the yield of this orchard.

h) A college official divides the student population into five classes: freshman, sophomore, junior, and graduate student. The official takes a simple random sample from each class and asks the members' opinions regarding student services.

i) Toyota wants to administer a satisfaction survey to its current customers. Using their customer database, the company randomly selects 80 customers and asks them about their level of satisfaction with the company

j) To determine her power usage, Dan divides up his day into three parts: morning, afternoon, and evening. He then measures his power usage at 3 randomly selected times during each part of the day.

k) A newspaper asks its readers to call in their opinion regarding the number of books they have read this month.

l) Toshiba wants to administer a satisfaction survey to its current customers. Using their customer database, the company randomly selects 80 customers and asks them about their level of satisfaction with the company.

m) An education researcher randomly selects 48 middle schools and interviews all the teachers at each school.

n) A market researcher selects 500 drivers under 30 years of age and 500 drivers over 30 years of age.

o) To avoid working late, a quality control analyst simply inspects the first 100 items produced in a day.

Solution

a) ***Systematic sampling***. The quality control manager is sampling every 8^{th} chip.

b) ***Cluster sampling***. The commission tests all members of the selected teams (clusters).

c) ***Cluster sampling***. The airline surveys all passengers on selected flights (clusters).

d) ***Stratified sampling***. The congresswoman samples some individuals from each of three different income brackets (strata).

e) ***Simple random sampling***. Each known user of the product has the same chance of being included in the sample.

f) ***Convenience sampling***. The radio station is relying on voluntary response to obtain the sample data.

g) ***Cluster sampling***. The farmer samples all trees within the selected subsection (clusters).

h) ***Stratified sampling***. The school official takes a sample of students from each of the four grades (strata).

i) ***Simple random***

j) ***Stratified sampling***

k) ***Convenience sampling***

l) ***Simple random***

m) ***Cluster sampling***

n) Stratified sampling

o) Convenience sampling

Exercise

Determine whether you would take a census or use a sampling to collect data for the study described:

a) The average credit card debt of the 65 employees of a company

b) The most popular grocery store among the 40,000 employees of a company

Solution

a) Census

b) Sampling

Exercise

Determine if the survey question is biased. If the question is biased, suggest a better wording

a) Why drinking fruit juice good for you?

b) Why is eating ice cream bad for you?

Solution

a) Yes. How do you think drinking fruit juice affects your health?

b) Yes. How do you think eating ice cream affects your health?

Exercise

A company has been rating television programs for more than 60 years. It uses several sampling procedures, but its main one is to track the viewing patterns of 20,000 households. These contain more than 45,000 people and are chosen to form a cross section of the overall population. The households represent various locations, ethnic groups, and income brackets. The data gathered from the sample of 20,000 households are used to draw inferences about the population of all households in the U.S.

a) What strata are used in the sample?

b) Why is it important to have a stratified sample for these ratings?

c) Observation studies are sometimes referred to as natural experiments. Explain what this means

Solution

a) The various locations, ethnic groups, and income brackets that are represented.

b) Stratified sampling ensures that each segment of the population is represented.

c) In an observation study, a researcher measures characteristics of interest of a part of a population but does not change existing conditions.

Exercise

Some polling agencies ask people to call a telephone number and give their response to a question

a) What is an advantage of this type of survey?
b) What is disadvantage of this type of survey?
c) Identify the sampling technique used.

Solution

a) This usually results in a savings in the survey cost.
b) There tends to be a lower response rate and this may introduce a bias into the sample. Only a certain segment of the population might respond.
c) Convenience sampling.

Exercise

A computer company employs 100 software engineers and hardware engineers. The personnel manager randomly selects 20 of the software engineers and 20 of the hardware and questions them about career opportunities within the company. Does this sampling plan result in a random sample? Simple random sample? Explain.

Solution

Yes it is random sample and it is not a simple random sample.
The sample is random because all employees have the same chance of being selected. It is not a simple random sample because some examples are not possible, such as a sample consisting of 30 software engineers and 10 hardware engineers

Exercise

Suppose you are the president of the student government. You wish to conduct a survey to determine that student body's opinion regarding student services. The administration provides you with a list of the names and phone numbers of the 19,935 registered students.

a) Discuss the procedure you would follow to obtain a simple random sample of 25 students.
b) Obtain this sample

Solution

a) The list provided by the administration serves as the frame. Number each student in the list of registered students, from 1 to 19,935. Generate 25 random numbers, without repetition, between 1 and 19,935.
b) Answers will vary

Exercise

True or False

a) When taking a systematic random sample of size n, every group of size n from the population has the same chance of being selected

b) A simple random sample is always preferred because it obtains the same information as other sampling plans but requires a smaller sample size.

c) When conducting a cluster sample, it is better to have fewer cluster with more individuals when the clusters are heterogeneous.

d) Inferences based on voluntary response samples are generally not reliable.

e) When obtaining a stratified sample, the number of individuals included within each stratum must be equal.

Solution

a) ***False***. In a systematic random sample, every kth individual is selected from the population.

b) ***False***. In many cases, other sampling techniques may provide equivalent or more information about the population with less "cost" than simple random sampling.

c) ***True***. When clusters are heterogeneous, the heterogeneity of each cluster likely resembles the heterogeneity of the population. In such cases, fewer clusters with more individuals from each cluster are preferred.

d) ***True***. Because the individuals in a convenience sample are not selected is not representative of the population.

e) ***False***. With stratified samples, the number of individuals sampled from each strata should be proportional to the size of the strata in the population.

Exercise

The human resource department at a certain company wants to conduct a survey regarding worker morale. The department has an alphabetical list of all 4502 employees at the company and wants to conduct a systematic sample.

a) Determine k if the sample size is 50

b) Determine the individuals who will be administered the survey. More than one answer is possible.

Solution

a) $\frac{N}{n} = \frac{4502}{50} = 90.04 \to 90$; thus $\underline{k = 90}$

b) Randomly select a number between 1 and 90. Suppose that we select 15. Then the individuals to be surveyed will be the 15^{th}, 105^{th}, 19^{th}, 285^{th}, and so on up uo the 4425^{th} employee on the company list.

Exercise

To predict the outcome of a county election, a newspaper obtains a list of all 945,035 registered voters in the county and wants to conduct a systematic sample.

a) Determine k if the sample size is 130

b) Determine the individuals who will be administered the survey. More than one answer is possible.

Solution

a) $\frac{N}{n} = \frac{945,035}{130} = 7269.5 \rightarrow 7269$; thus $\boxed{k = 7269}$

b) Randomly select a number between 1 and 7269. Suppose that we select 2000. Then we will survey the individuals numbered 2000, 9269, 16,538, 23,807, and so on up to the individual numbered 939,701.