Unit 1: Data Analysis

Lesson: Data Collection

DEFINITIONS:

QUALITATIVE VARIABLE	SIMILAR	SAMPLE	LONGITUDINAL STUDY	CENSUS
ENTIRE	POPULATION	ANALYZING	QUANTITATIVE VARIABLE	POPULATION
CROSS-SECTIONAL STUDY				

_____ All individual (people, places, things, etc...) belonging to the group of being studied.

_____ A section of individual taken from a population.

_____ Survey the ENTIRE population.

_____ _____ Variables that can be measured numerically.

_____ _____ Variables that cannot be measured numerically and are categorical nature.

_____ A study that considers individuals from different groups at the same time.

_____ A study that considers individuals over a prolonged period of time.

What makes a good sample?

Ideally, a sample is a good representation of the _____ population. This means that the results you get from analyzing the sample should be _____ to the results you would have gotten from _____ the entire _____.

Practice for Data Collection:

For each of the following

i) identify the population
ii) Provide a possible Quantitative/Qualitative variable that can be used for the research
iii) determine if the study should be longitudinal or cross-sectional

1) Conducting a poll for the voters' choice on the next Ontario election

   Population: _____

   Type of Variable: _____

   Type of Study: _____

2) Preference on wearing ordinary glasses or contact lenses for women

   Population: _____

   Type of Variable: _____

   Type of Study: _____

3) Comparison on depreciation rate for domestic and foreign cars in North America for the first year

   Population: _____

   Type of Variable: _____

   Type of Study: _____

4) The effect of breakfast on high school students' academic performance

   Population: _____

   Type of Variable: _____

   Type of Study: _____

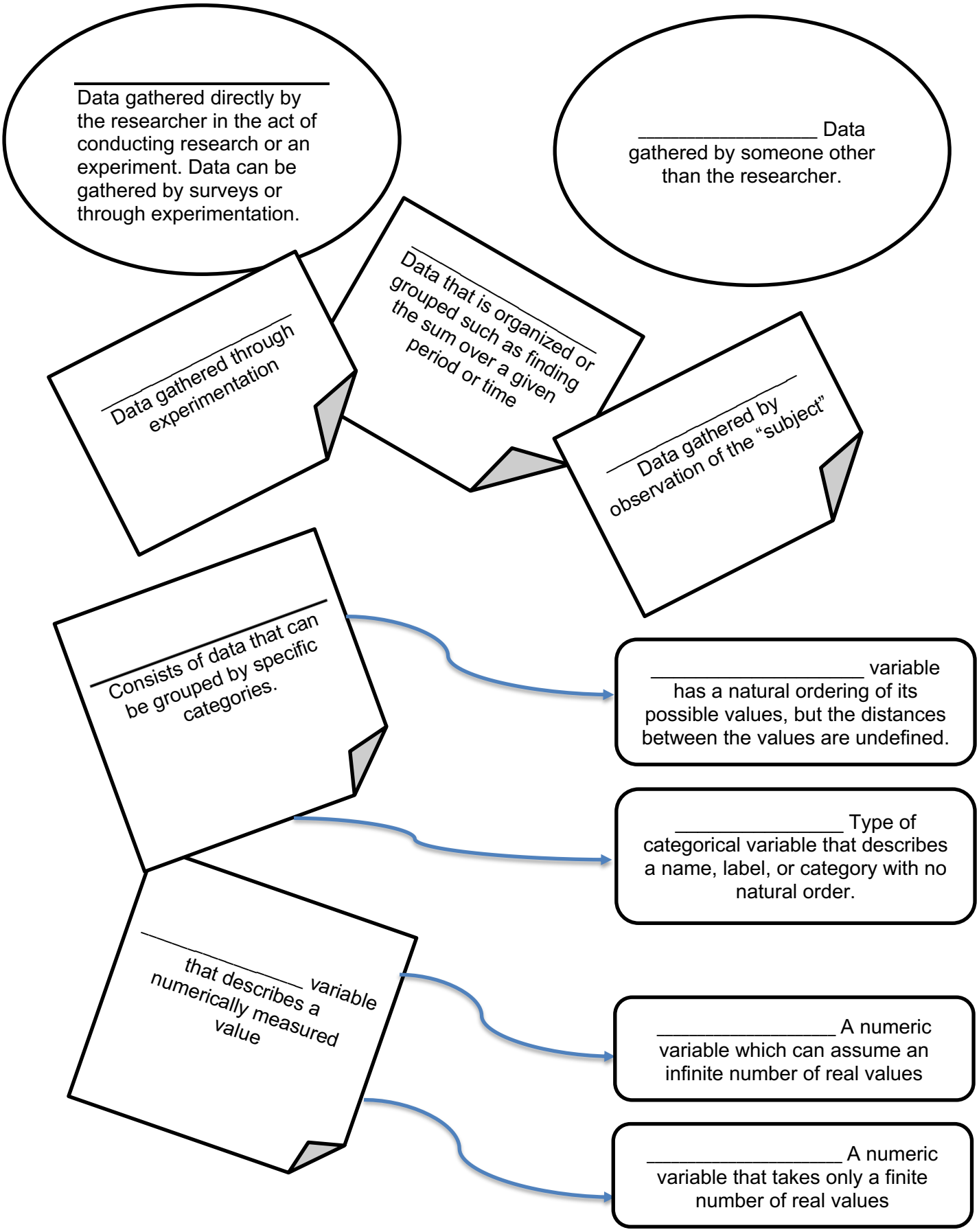
5) Which gender would more likely be getting speeding tickets in the age between 17-21?

   Population: _____

   Type of Variable: _____

   Type of Study: _____

Lesson: Types of Data



Practice: Types of Data

1. Use the words given to complete the sentences below.

   Discrete, Continuous, Categorical, Secondary

   32 pupils in a classroom fill out surveys….

   a) Their favourite type of fruit is _____ data.
   b) The span of the pupils hands is _____data.
   c) The number of siblings each pupil has is _____data.

2. Data is collected on the following areas. For each, identify all variables in the sentence as: continuous, discrete, or qualitative. The data may fit more than one of these categories.
   a) Favourite colours of 30 students _____
   b) Time taken to run 100 m (to the nearest second) _____
   c) Duration of journey to school _____
   d) Which subjects were chosen as an option _____
   e) Number of students in the school each year _____
   f) Cost of the journey to school _____

3. Sarah is reading an energy saving article on the internet. The article covers a broad range of topics. She learns that the average UK household uses 349.44254 litres of water per day.
   a) Is household usage of water discrete or continuous data? _____
   b) Is Sarah working with primary or secondary data? _____

4. Give one advantage and one disadvantage of using secondary data. _____

5. Which of the following is a significant problem specifically associated with secondary data?
   a) It is not be accurate.
   b) It is biased.
   c) It may not address the aims.
   d) It is expensive.

Answers:
1.a)categorical b)continuous c)discrete 2.qualitative for colours, discrete for # of students b)discrete for rounded time recorded c) continuous for duration and continuous for distance to school d)discrete for # of subjects and qualitative for type of subjects e)discrete for number of students f)discrete for cost and continuous for distance to school 3.a)continuous b)secondary 4.advantage: take less time than collecting it yourself; disadvantage: unsure whether the sample is biased/representative/fair 5.it may not address the aims

_____

Every member of the population has an equal chance of being selected
EX: Lottery, draw names from an envelope/box

_____

There is an open invitation to any member of the population to participate in a survey or experiment

_____

Surveying members of the population that are easily accessible

_____

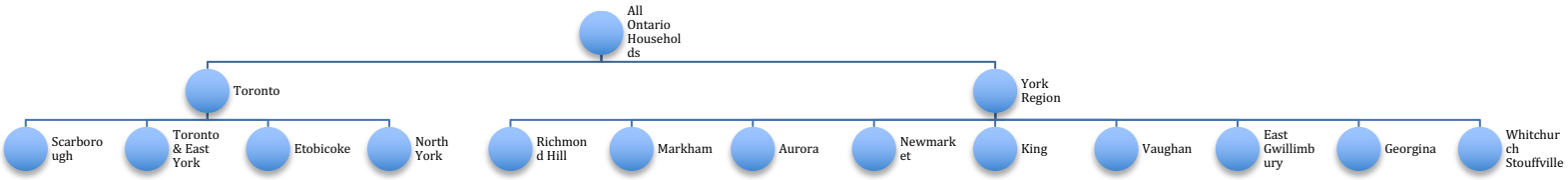
Selecting a sub-group as being representative of the entire population and every member of the sub-group participates

_____

requires that the samples that are taken are destroyed in the process of testing
Ex. Cars used in crash tests

_____

Selecting members of the population randomly in stages



_____

Selecting members that are in sequence at regular intervals

$$interval = \frac{population\ size}{sample\ size}$$

_____

Selecting members randomly from a population that may be divided into groups (strata) that share a common characteristic (such as gender, grade, education) such that the proportional representation of a strata in the sample is the same as the proportional representation in the population

Ex: survey 25% of the drivers in nearby 4 cities

Cities	Number of drivers	Sample size
Markham	4088	
Richmond Hill	5409	
North York	2459	
Scarborough	3236	
Total	**15 192**	

Practice for Sampling Techniques:

Identify the population and sampling techniques for each of the following scenarios.

Example1. For each of the following scenarios, identify the population and the type of sampling used.

a. A grade 12 class wants to determine the average height of the students in the class. The names of all students in the class are placed in a hat and then names are drawn.

b. A Canadian publisher wants to know which type of book high school students like best. The publisher chooses ten high schools at random and then surveys every student in each school.

c. You want to collect data about which TV shows teenagers like best. You ask your friends and teenage cousins which shows are their favourites.

d. The student council wants to know how much money students are willing to pay for a yearbook. They choose five boys and five girls from each period 1 class, from each grade, and survey them.

e. Police want to know the speed motorists drive along Bayview Avenue. They set up radar and measure the speed of every 10^{th} vehicle for a day.

Example 2. Before booking bands for the school dances, the Student Council at Bayview Secondary School wants to survey the music preferences of the entire student body. The following table shows the enrollment at the school.

Grade	Number of Students	Sample
9	455	
10	432	
11	409	
12	384	
Total		

a. Design a stratified sample for 20% of the student body

b. You have been given an alpha list of students and due to time-constraint, a sample size of about 50 students is deemed to be sufficient. Suggest a method for selecting a systematic sample.

Sampling Techniques Homework Do #1-4, 7-9

Practice A

1. Identify the population for each of the following questions
    a. Who should be the next president of the students' council?
    b. Who should be next year's grade-10 representative on the student council?
    c. What is your faviourite soft drink?
    d. What Beatles song was the best?
    e. How effective is a new headache remedy?
2. Choose the sampling method used in each of the following scenarios.
    a. A radio-show host invites listeners to call in with their views on banning smoking in restaurants.
    b. The Heritage Ministry selects a sample of recent immigrants such that the proportion from each country of origin is the same as for all immigrants last year.
    c. A reporter stops people on a downtown street to ask what they think of the city's lakefront.
    d. A school guidance counselor arranges interviews with every fifth student on the alphabetized attendance roster.
    e. A statistician conducting a survey randomly selects 20 cities from across Canada, then 5 neighbourhoods from each of the cities, and then 3 households from each of the neighbourhoods.
    f. The province randomly chooses 25 public schools to participate in a new fundraising initiative.
3. What type(s) of sample would be appropriate for
    a. A survey of engineers, technicians, and managers employed by a company?
    b. Determining the most popular pizza topping?
    c. Measuring customer satisfaction for a department store?

Apply, Solve, Communicate B

4. Natasha is organizing the annual family picnic and wants to arrange a menu that will appeal to children, teens, and adults. She estimates that she has enough time to survey about a dozen people. How should Natasha design a stratified sample if she expects 13 children, 8 teens, and 16 adults to attend the picnic?
5. **Communication** Find out, or estimate, how many students attend your school. Describe how you would design a systematic sample of these students. Assume that you can survey about 20 students.
6. The newly elected Chancellor of the Galactic Federation is interested in the opinions of all citizens regarding economic conditions in the galaxy. Unfortunately, she does not have the resources to visit every populated planet or to send delegates to them. Describe how the Chancellor might organize a multi-stage sample to carry out her survey.
7. **Communication** A community centre chooses 15 of its members at random and ask them to have each member of their families complete a short questionnarire.
    a. What type of sample is the community centre using?
    b. Are the 15 community-centre members a random sample of the community? Explain.
    c. To what extent are the family members randomly chosen?
8. **Application** A students' council is conducting a poll of students as they enter the cafeteria.
    a. What sampling method is the student council using?
    b. Discuss whether this method is appropriate for surveying students' opinions on
        i. The new mural in the cafeteria
        ii. The location for the graduation prom
    c. Would another sampling technique be better for either of the surveys in part b>?
9. **Application** The host of a call-in program invites listeners to comment on a recent trade by the Toronto Maple Leafs. One caller criticizes the host, stating that the sampling technique is not random. The host replies: "So what? It doesn't matter!"
    a. What sampling technique is the call-in show using?
    b. Is the caller's statement correct? Explain.
    c. Is the host's response mathematically correct? Why or why not?

Answers Section 2.3 pp. 117-118

Practice A

1. a) students in the particular school
   b) grade 10 students in the particular school
   c) you
   d) all those who have listened to the Beatles' music
   all those who have tried the new remedy
2. a) voluntary-response sample b)stratified sample
   c) convenience sample d) systematic sample
   e) multi-stage sample f) cluster sample
3. a) stratified sample
   b) simple random sample, convenience sample
   c) voluntary-response sample

Apply, Solve, Communicate B

4. sample $\frac{12}{37} \doteq 32\%$ of the members of each group; 4 children, 3 teens, 5 adults
5. Obtain a list of the students in the school. Calculate $\frac{n}{20}$ names. Then, select every $\frac{n}{20}th$ name.
6. Answers may vary.
7. a) cluster sample
   b) No, not every member of the community had an equal chance of being surveyed. To the same extent as the member of the community centre was randomly selected.
8. a) convenience sample
   b) i) yes ii) no
   c) In part ii), a simple random sample should be used.
9. a) voluntary-response sample
   b) yes
   c) not if the host claims that the callers are representative of the population
10. Answers will vary.
11. Answers will vary.
12. Answers will vary.

Lesson: Data Collection Bias

Collection Bias: Statistical bias is any factor that favors certain outcomes or responses and hence systematically influences (or skews) the results of a survey.

_____

Sampling bias occurs when the sample does not reflect the characteristics of the population. This can result from poor sampling technique or data collection methods. This should be reduced by planning the study utilizing the appropriate sampling technique.

_____Household bias occurs when one type of Household Bias respondent is over represented because different sized groupings are polled equally.

_____

This is a form of sampling bias caused when a particular group or groups are underrepresented in the survey because they have chosen not to participate.

_____

If the respondent **deliberately** gives false or misleading answers to the survey this would Response Bias be called response bias. The dishonesty in the answers may be caused by embarrassment or fear to answer the question honestly.

_____

Measurement bias is a consistent measurement error which skews the results of the survey. Often the data collection process affects the variable that is being measured. Someimes these effects (if known) can be accounted for in the analysis.

Two types of measurement bias questions:

(a) _____

is a question that suggests an answer which otherwise may not have been chosen by the respondent without prompting.

(b) _____

is a question contains wording or information which influences the respondents.

Practice for Data Collection Bias:

Choose the MOST obvious type of bias.

a) A survey asked students at a high school basketball game whether a fund for extracurricular activities should be used to buy new equipment for the basketball team or instruments for the school band.

Types of Bias	Explanation	Solution

b) A social science class asks every tenth student entering the cafeteria to answer a survey on environment issues. Less than half complete the questionnaire. The completed questionnaires show that a high proportion of the respondents are concerned about the environment and about environmental issues.

Types of Bias	Explanation	Solution

c) A civil engineer suggests that a cost-effective way to survey traffic speeds on a highway would be to have a polic officer patrol the highway and record the speed of the traffic every half hour.

Types of Bias	Explanation	Solution

d) An aid agency in a developing country wants to know what proportion of households have at least one personal computer. One of the agency's staff members conduct the survey by calling household randomly selected from the telephone directory.

Types of Bias	Explanation	Solution

e) The following questionnaire was developed by Sarah's friends"

Election Survey Sponsored by the friends for the election of Sarah committee.
Circle the appropriate response.
Gender: Male Female

Grade: 9 10 11 12 On election day, I intend to vote for:

Sarah Zahir Melanie

I want more:
dances hat days holidays fun

Types of Bias	Explanation	Solution

f) One student from each home room is surveyed about which new menu choices they would prefer in the cafeteria. Below is information about the number of people in each home room: 8 Home rooms with 1-12 students 42 Home rooms with 13-24 students
20 Home rooms with 25-36 students

Types of Bias	Explanation	Solution

g) A teacher finishes explaining a new concept to the class and wants to check that all the students have grasped the concept. The teacher asks those who do not understand to raise their hands.		
Types of Bias	Explanation	Solution
h) As part of a survey of the "Greatest Hits of All Time," a radio station asks its listeners: Which was the best Elvis songs? Blue Suede Shoes； Hound Dog； Heartbreak Hotel; All Shook Up; Suspicious Minds		
Types of Bias	Explanation	Solution
i) Members of a golf and country club are polled regarding the construction of a highway interchange on part of their golf course.		
Types of Bias	Explanation	Solution
j) A group of candidates running for the position of Prime Minister are asked whether they have ever taken part in an illegal protest.		
Types of Bias	Explanation	Solution
k) A random poll asks the following question: "The proposed nuclear power plant will produce a number of jobs and economic activity in and around your community. Are you in favour of this forward-thinking initiative?"		
Types of Bias	Explanation	Solution
l) A survey uses a cluster sample of Toronto residents to determine public opinion on whether the provincial government should increase funding on public transit.		
Types of Bias	Explanation	Solution
m) A researcher appears on a talk-show and conducts an on-air survey about re-instituting capital punishment in Canada.		
Types of Bias	Explanation	Solution
n) Parents of high school students were asked: "Do you think that students should be released from school an hour early on Friday, free to run around and get into trouble?"		
Types of Bias	Explanation	Solution
o) Audience members at an investment workshop are asked to raise their hands if they have been late with a bill payment within the last six months.		
Types of Bias	Explanation	Solution

Answers:

a) Sampling
b) Non-response
c) Response
d) Non-response
e) Measurement
f) Sampling
g) Response
h) Measurement
i) Sampling
j) Response
k) Measurement
l) Sampling
m) Non-response
n) Measurement
o) Response

Data Collecting Bias Homework Do #1-3,5,6,7

Practice A

1. Classify the bias in each of the following scenarios.
   a. Members of a golf and country club are polled regarding the construction of a highway interchange on part of their golf course.
   b. A group of city councilors are asked whether they have ever taken part in an illegal protest.
   c. A random poll asks the following question: "The proposed casino will produce a number of jobs and economic activity in and around your city, and it will also generate revenue for the provincial government. Are you in favour of this forward-thinking initiative?"
   d. A survey uses a cluster sample of Toronto residents to determine public opinion on whether the provincial government should increase funding for the public transit.

Apply, Solve, Communicate

2. For each scenario in question 1, suggest how the survey process could be changed to eliminate bias.
3. **Communication** Reword each of the following questions to eliminate the measurement bias.
   a. In light of the current government's weak policies, do you think that it is time for a refreshing change at the next federal election?
   b. Do you plan to support the current government at the next federal election, in order that they can continue to implement their effective policies?
   c. Is first-year calculus as brutal as they say?
   d. Which of the following is your favourite male movie star?
      i. Al Pacino
      ii. Keanu Reeves
      iii. Robert DeNiro
      iv. Jack Nicholson
      v. Antonio Banderas
      vi. Other:
4. **Communication**
   a. Write your own example of a leading question and a loaded question.
   b. Write an unbiased version for each of these two questions.
5. A school principal wants to survey data-management students to determine whether having computer Internet access at home improves their success in this course.
   a. What type of sample would you suggest? Why? Describe a technique for choosing the sample.
   b. The following questions were drafted for the survey questionnaire. Identify any bias in the questions and suggest a rewording to eliminate the bias.
      i. Can your family afford high-speed Internet access?
      ii. Answer the question that follows your mark in data management.
      Over 80%: How many hours per week do you spend on the Internet at home?
      60-80%: Would home Internet access improve your mark in data management?
      Below 60%: Would increased Internet access at school improve your mark in data management?
   c. Suppose the goal is to convince the school board that every data management student needs daily access to computers and the Internet in the classroom. How might you alter your sampling technique to help achieve the desired results in this survey? Would these results still be statistically valid?
6. **Application** A talk-show host conducts an on-air survey about re-instituting capital punishment in Canada. Six out of ten callers voice their support for capital punishment. The next day, the host claims that 60% of Canadians are in favour of capital punishment. Is this claim statistically valid? Explain your reasoning.
7. a) Locate an article from a newspaper, periodical, or Internet site that involves a study that contains bias.
   b) Briefly describe the study and its findings.
   c) Describe the nature of the bias inherent in the study.
   d) How has this bias affected the results of the study?
   e) Suggest how the study could have eliminated the bias.
8. **Inquiry/ Problem Solving** Do you think that the members of Parliament are a representative sample of the population? Why or why not?

Answers Section 2.4, pp. 123-124

Practise

1. a) sampling bias b) response bias
   c) measurement biase – loaded question
   d) sampling bias-the cluster may not be representative
2. a) A random sample of area residents should be taken.
   b) The responses could be returned anonymously.
   c) Remove the preamble to the question and the phrase "forward-thinking."
   d) A random sample of area residents should be taken.
3. Answers may vary.
   a) Which party will you vote for in the next federal election?
   b) same as a)
   c) Do you think first-year calculus is an easy or difficult course?
   d) Who is your favourite male movie star?
   e) Do you think that fighting should be eliminated from professional hockey?
4. Answers may vary.
6. Not statistically valid; the sample may contain response bias and also sampling bias. The radio station may have a conservative audience.
7. Answers may vary. 8. Answer may vary.

Lesson: Creating a Survey

Format of a Survey/Questionnaire

(a) _____

(b) _____

(c) Survey **Design:**

- **K**eep **I**t **S**hort and **S**imple
- Types of questions (* definitions look up on your own)

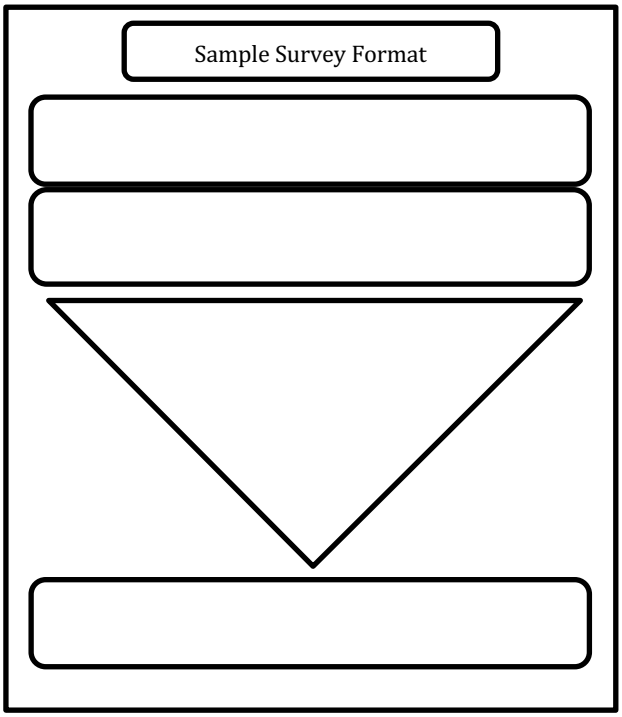
Closed-ended questions:

- Factual
- Rating scale
- Likert-type
- Semantic differential
- Multiple-choice/Dichotomous
- Rank order

Open-ended questions:

- Opinion and attitudes – questions related to opinions, values, and attitudes that are subjective about the thought processes and feelings of a respondent

(d) _____



Characteristics of a Good Survey

- Includes questions on all subjects essential to the research problem
- Does not include questions not related to the research problem
- Appears brief and easy to complete
- Respondents should feel like they are participating in an important project
- No questions that could bias the answers
- Designed to elicit clear and precise answers to all questions
- Phrasing, structure, and layout designed with the problem of tabulating the findings in mind

Characteristics of a Poor Survey

- Range of responses for questions is too narrow
- Misunderstood by respondents
- Too vague
- Ask information the respondent cannot remember or does not have
- Wrong level of intimacy
- Takes too much for granted
- Double negatives
- Overlong lists
- Overlapping alternatives

Survey question types and survey structure

The types of survey questions used in a survey will play a role in producing unbiased or relevant survey responses. As the survey designer, consider the types of questions to use and when it is appropriate to use them. Question types range from open-ended (comments to essays) to closed-ended (yes/no, multiple choice, rating scale, etc). In the end, it is the question types that determine what type of information is collected.

Question Types

1. Open-Ended Types

Open-ended questions are those that allow respondents to answer in their own words. Open-ended questions seek a free response and aim to determine what is at the tip of the respondent's mind. These are good to use when asking for attitude or feelings, likes and dislikes, memory recall, opinions, or additional comments. However, there can be some drawbacks to using open-ended questions:

- Sometimes respondents may find it difficult to express their feelings. This can result with respondents answering "I don"t know" or skipping it.
- They do take more time and effort to fill out and at times they can have a larger skip rate.
- In addition, analyzing open-ended comments can be time consuming and difficult. 2. Closed–Ended Types (Multiple Choice – One Answer or Multiple Answers)

2. Closed-ended questions are those with pre-designed answers with a small or large set of potential choices. One type of closed-ended question is a "dichotomous" question which allows respondents to choose one of two answer choices (e.g. Yes or No), while another type is the "multi-chotomous" question, which allows respondents to choose one of many answer choices.

3. Ranked or Ordinal Questions

Ranking questions are best to use when all the choices listed should be ranked according to a level of specification (e.g. level of importance). If you have a question in which you need the respondents to indicate what items are the "most important" to "least important" then you can set up a ranking question.

4. Rating Types

The rating type questions are used when surveying the frequency of something like behavior or attitude. It is best to present the rating scale in a logical or consistent order. Therefore, it makes sense to order the ranking or rating choices from low to high (e.g. Strongly Disagree to Strongly Agree going from left to right).

If you set up the rating scale in your survey in this format of "Strongly Disagree" to "Strongly Agree" make sure that the rest of the survey is consistent and all rating scales go from the low to the high frequency throughout (or vice versa). In addition, some surveys may only label the outliers or endpoints of the scale, but it is good practice to

A five-point rating scale typically gives sufficient discrimination and is easily understood by survey participants. This is usually recommended for most survey settings. However, there is no set limit on the number of categories to use. Using too few could give less-cultivated information, while using too many could make the question hard to read and answer. Related to this setup is the decision of incorporating a "middle category".

Question Sequence

A good survey design should help to stimulate recall (if necessary); it should motivate the respondent to reply; and the survey should flow in an orderly fashion. The sequence of questions will help to create a certain flow to the survey. This flow will also ease or arouse the respondent's interest and overcome his/her doubts about the survey's intent. As a general guideline, there are three areas regarding question sequence: opening questions, question flow, and location of sensitive questions.

1. Opening questions – The first few questions in the survey should be easy and interesting in order to calm any participants' suspicions about the survey's integrity. This allows the participants to build up confidence in the survey's objective. In return, this may stimulate their interest and overall participation.

2. Question flow – The question sequence in the survey body should take on a flow of ideas and be geared towards the respondents' abilities. After you have established the first general topic, all related questions should come up before a second topic is raised. It is a good idea to use "pages" in the online design to house each different section of the survey. Here you can raise one topic on one page and include the instructions/information for this section in the Page Description area. When you are then ready to introduce a new topic to the survey, you can create a new or second page to include that page's description and purpose. Conditional or Skip Logic questions are also a good way to control the respondent's flow or route through the survey. You can apply question or page skip logic to the survey when you want to guide respondents and exclude them from certain pages of questions that do not apply to them.

3. Location of sensitive questions – Some suggest that sensitive questions should not be included at the beginning of the survey. However, there are no set rules on this. If you do

include sensitive questions at the beginning of the survey, then you may run into respondents rejecting the survey and exiting early. They may not have built up confidence yet in the survey's integrity quite so early. Questions like demographics or personal information are usually best to introduce towards the end of the survey. This way, respondents are likely to have already developed confidence in the survey's objective.

Survey Layout

1. Basic guidelines

a. Introduction

When designing your survey structure, the overall format and layout is important from beginning to end. A poorly organized survey may cause respondents to skip questions or completely opt out of answering your survey. It is good practice to begin your survey with an introduction that explains the survey's purpose. Within the introduction, you may want to include the name of the organization conducting the survey, the confidentiality information, and how the data collected will be used. Many participants like some kind of assurance in regards to their responses; providing that kind of information before the survey starts can help ease those concerns. You may also want to provide an estimate of how long the survey might take or whether you are offering any kind of incentive or prize for taking the survey. Remember to deliver on your promised gift! If you provide this information up front it usually leads to honest responses and more completed surveys.

Providing general instructions on how to progress through the survey in the introduction or within each new section is important in letting your audience know how the survey works. From here respondents will not have to look back and forth in the survey to see what they are supposed to do.

b. Body of the Survey Design

The use of space throughout the survey is also important. Trying to fit too much information (e.g. too many questions) on a single page may cause respondents to struggle through the survey. If your survey has multiple sections or parts, then it is good to introduce each new section as suggested previously. Keep in mind to make the sections and questions flow in a sequential order that makes sense to the respondents.

Here are some tips to remember when designing the look of your survey:

1. Make the survey visually appealing and user-friendly.
2. Try not to use small fonts or fonts that are not easy to read. Some participants may have a difficult time reading small print.
3. To avoid clutter, use white space.
4. Ask only one question per line.
5. Group similar question together or in the same area of the survey.
6. Ask interesting questions in the beginning of the survey to grab the participants" attention. This helps to stimulate interest.
7. Place demographic and/or sensitive questions at the end of the survey. If they are in the beginning, participants may opt out early.
8. Finally, test the survey before going live. A small sample of test respondents can help verify if your survey is working properly. This enables you to revise and edit questions and the survey design.

c. End of Survey or Thank You Page

Once your respondent has reached the end of your survey, you can create a Thank You page. Here you can thank the respondent for their time. This may help to build rapport with the respondent; possibly increasing the likelihood that they will participate in your future survey invites.

2. Layout for coding and identification

As the designer of the survey, pay attention to the physical layout of the survey to reduce the likelihood of errors on the respondents" end or on your end regarding areas of coding or editing. Here are some principles to follow to make the survey logical for all people accessing the survey, as well as easy to identify, code, and store:

1. Identification – You can add a unique number or identifier to each questionnaire.

2. Numbering – Questions can be numbered sequentially throughout the survey, even if the survey is divided by pages or sections. You can choose to have our tool number the questions throughout the entire survey as a whole or have the questions numbered according to each individual page. This may help you in coding your survey.

3. Instructions – General instructions are important for the administration of the survey as well as for the collection of accurate data. The following two types of information ought to be distinguishable in the survey: questions to be read / answered and instructions to be followed. You may want to customize your survey to include different fonts for the instructions or page descriptions vs. the survey questions themselves. Place any special instructions on either the page description/section or directly above the question itself.

4. Fonts & Formats – If you want to emphasize important words, then underline or bold them in the survey question or page description. This makes it easier for your respondents to identify key points or items.

4 Common Sense Tips for Creating Surveys that Work

Author: Scott Smith, Ph.D. |December 31, 2012

Creating online surveys is as much an art as it is a science.

It involves attention to detail in the design and flow of your survey. Creating an effective survey that yields actionable insights can be difficult.

Effective survey design and flow gives power to your research. But *great questions* are the foundation for great research.

There are fundamental best practices for creating survey questions that all researchers must know.

Here are 4 tips for **creating surveys** that work.

(1) Keep It Simple

Do you remember taking the SAT or ACT? It's a long and boring process.

Your average survey respondent can start to feel that way about 15 minutes into a survey. Fifteen minutes is a good upper-limit for most surveys.

When a survey is too long, three bad things can happen:

1. **Respondents drop out:** They simply quit taking the survey. It costs money to find respondents, and a high drop-out rate can not only cost a lot, but can influence the quality of your results. Having a reward for completion can reduce drop-outs, but you can't stop it completely.
2. **People stop paying attention:** Remember your elementary-school classmate who just filled in random bubbles during a test? He grew up. If it takes too long to take your survey, he might do it again. We actually see this a lot, and encourage researchers to use attention filters.
3. **Clients get angry:** The irony of upsetting customers with an overly long satisfaction survey is not lost on your respondents. The best way to collect quality data is to keep your surveys short, simple, and well organized.

(2) Use Scales Whenever Possible

Scales are more than a little important.

Rather than asking respondents a basic yes or no question, use scales that measure both the direction and the intensity of opinions.

This is critical for research.

Someone who "Strongly Supports" a decision is very different from someone who only "Slightly Supports" it.

Scales extend the power of analysis from basic percentages to high-level analyses based on means and variance estimates (like t-test, ANOVA, regression, and the like).

Use scales whenever you can. You will get more information from each question.

(3) Keep Coded Values Consistent

Every survey response, option, question, or answer is coded as a numeric value that is reported as a percent of responses or as a mean, median, range, etc.

These values are the basis for analysis.

- **Mean:** Often referred to as an average, it is the sum of all the values divided by the number of values.
- **Median:** The middle point in a data set. To determine the median, lay out a distribution from lowest to highest and select the middle value.
- **Range:** The highest and lowest data points in a distribution form the range. VARIANCE: A dispersion measure of how far a set of numbers is spread out.
- **Example:** Assuming we have data points 1, 2, and 6: Mean: 3 = (9 / 3) Median: 2 Range: 1-6 Variance: 7

Values must be coded consistently. Generally, we assign the highest value to the best outcome (ie "Strongly Agree" that customer service is responsive) and then move down from there.

For simplicity, keep your scale direction consistent throughout your survey. This makes it easier for respondents to answer and for you as a researcher to conduct your analysis.

If scales have the same scale of points, you can quickly compare responses to different questions. For example, if a survey asks respondents to rate a series of

statements from Strongly Disagree to Strongly Agree, the responses are given these values:

Standard scaling helps managers to quickly understand customer service ratings by simply looking at averages.

For example, once managers understand that a 5-point agreement scale is being used, they could be given the mean results for the following customer evaluation (agreement) statements:

- I am completely satisfied with the customer service — 3.15
- The customer service is prompt — 4.12
- Customer service representatives are polite — 4.67
- Customer service representatives are knowledgeable — 2.08

Since all the statements are positive and the values are scaled consistently, a higher mean reflects better results in that area. A manager can look at these means and quickly identify the 2.08.

We see that customer service representatives are prompt and polite, but they don't seem to know what they're talking about. As a result, overall satisfaction with customer service is perhaps much lower than it could be.

You can reverse scales (or word questions negatively) to encourage respondents to read more carefully.

However, if you use reversed scales or negative wording for some items, be sure to recode the scales so that all scales point in the same direction. This will allow you to quickly compare multiple areas of customer service. (You can do recodes easily in Qualtrics.)

The simplest solution is just to keep all scales consistent throughout every survey.

(4) Explain Why

Respondents are more likely to help you if they see something of positive value for them.

Value offerings can range from a very general altruistic appeal for their help to a very specific offer of an economic incentive. For instance, with a customer feedback survey, you can explain that feedback will help improve customer service.

Here are some quick examples:

1. **Make it specific to them:** With employee evaluations, you can explain that feedback will be used to determine awards, promotions, and pay raises and will help management make organizational decisions that will affect them.
2. **Explain unexpected questions:** For instance, if it's important for you to ask toy store customers their preferred color of jeans, you might want to explain why that is relevant.
3. **Justify requests for sensitive information:** For instance, you can explain that purchasing habits will only be analyzed in aggregate for benchmarking purposes or that results will not be shared outside your organization.

So what do you think? What are some tips you'd give to fellow researchers? Leave a comment below and let's discuss.