Data Collection
7.1 and 7.2

➢ Data = a collection of facts such as numbers, measurements, words, descriptions of things, or observations.

➢ There are two types of data:

   o Quantitative Data = numerical data = numerical information

      Examples: number of students in math 9 course, number of teenagers who live in the Yukon, average time needed to commute to school, average height of a PCSS student, average   score in English 12.

   o Qualitative Data = categorical data = descriptive information

      Examples: breeds of dogs, colours, makes of cars, types of sports, types of movies, favourite fashion brand, …
      OR range of descriptors: "strongly agree – agree – somewhat agree – not sure – somewhat disagree – disagree – strongly disagree".

➢ There are two types of quantitative = numerical data

   o Discrete = comes from counting. No decimals and no fractions. Graphed as distinct points.

Examples: number of salmon swimming through the fish ladder a day, number of cars in a parking lot, number of passengers on a bus, number of languages   a person can speak.

   o Continuous = comes from measuring. Decimals or fractions are allowed. Graphed as connected lines.

Examples: length of each fish caught in Tagish Lake during August, Marsh Lake water temperature, average height of an 8th grader,

➢ There are two types of qualitative = categorical data

  o Nominal = data cannot be ordered.

Examples: Languages a person can speak: English, French, Southern Tutchone, Spanish, Tagalog, Mandarin, Cantonese, Korean, Japanese, German, …
        Sneakers people wear most often: VANS, Nike, Adidas, Sketchers, Fila, …

  o Ordinal= data can be ordered.

Examples: How satisfied are you with Whitehorse movie theatres? Select one:
        Very satisfied – satisfied – somewhat satisfied – somewhat dissatisfied – dissatisfied – extremely dissatisfied.

        How difficult is Math 9 course? Select one:
        Very difficult– difficult – somewhat difficult – neither easy nor difficult – somewhat easy – easy – very easy.

➢ Data can be collected in many different ways:
  o Observation
  o Experiment = a controlled test or investigation
  o Questionnaire/Survey
  o **<u>Census</u>** = a survey during which everyone in a population answers the questions.
  o Database

➢ Population = everyone or everything with certain characteristics

Examples: All people who live in Canada, All students who attend PCSS, All teenagers who live in Whitehorse, All Indigenous Students in Yukon schools, All senior citizens who are retired, All young adults (18-20 years old) who work more than 20 hours a week and live in the Yukon, …

➢ Sample = a part (subgroup) of a population. A sample is carefully selected from the population to fairly represent the population. Sampling saves time, human resources, $, and natural resources/energy.

Examples: Every 1000th person who lives in Canada, 5 random students from each grade in PCSS, 300 teenagers that are randomly interviewed at CGC, Better Bodies, Tim Horton's, and Starbucks, Every third indigenous student selected from an alphabetized list of all indigenous students in the Yukon, Every fifth person who uses senior discount at Shoppers Drug Mart on Thursday morning, CRA contacts every tenth person who is between 18-20 years old and their T4 shows earnings over $16 000.00.

Types of Samples

- ➢ **<u>Convenience Sample</u>**
- o To sample students at PCSS I stand by the main door at 8:35 and select every 15th person who enters the main door.
- o To sample people who use public transit, I select people who, like me, wait for the bus in front of the city hall at 7:35 AM.
- o To determine people's favourite pizza place in Whitehorse, I ask people from my work, my friends, family, and my neighbours.

- ➢ **<u>Volunteer Sample</u>**
- o I distribute a questionnaire to all students in PCSS and collect answers only from those who want to share their opinion.
- o I e-mail all my contacts asking them to share their opinion on the Black Panther movie. Only those who have time and either really loved or disliked the movie will answer.

- ➢ **<u>Random Sample</u>** = every individual in the population has the same chance of being selected to represent the population.

  - ❖ The best type of a sample as it is most likely to represent the population.
    - o Selecting every 10th person from an alphabetical list of all students in PCSS.
    - o Writing the name of every student in Math 9 on separate piece of paper. Folding the paper and putting all pieces in a box. Selecting three names from the box.

Questionnaires

➢ Questions have to be fair = unbiased. Unbiased questions do not influence a person's answer. A question that influences a person's answer is **<u>biased</u>** question.

Biased question:
Q: What do you think about the amazing fall collection by Gucci?
Q: Do you think that the Soccer World Cup should not have taken place in Qatar?
Q: Should our school vending machine offer junk food?

Unbiased question:
Q: What do you think about Gucci's fall collection?
Q: What do you think about Qatar hosting the Soccer World Cup?
Q: What types of snacks should be available in our school vending machine?

➢ Each person should find an answer he or she would choose.

➢ All people should understand the question the same way.

➢ Privacy or each respondent should be respected and protected.