## Section 4.1a - Statistics

This booklet belongs to: $\qquad$ Block: $\qquad$

- Essentially influence our opinions, thoughts, and decisions.

How often have you heard the following starts to sentences?
$\checkmark \quad 1$ in 5 people....
$\checkmark$ The majority of Canadians....
Need to be looked at with a more skeptical eye
$\checkmark$ A new study shows that....

How about watching sports?
$\checkmark \quad$ There are statistics for everything

- Pitches thrown
- Batting average
- Pass accuracy
- Shots on net
- Save percentage

It goes on and on and on.

- All these statistics that you see, all the information that is presented to you, can be biased, skewed, or deceptive; you need to always look deeper.

Questions to ask yourself:
$\checkmark$ Who did the research?
$\checkmark$ What is the message and what are their references (sources)?
$\checkmark$ Who funded the research?
$\checkmark$ Who do the results benefit/harm?
$\checkmark$ What could cause bias in the results?

- This list is not the be all and end all, there are many more questions to ask to ensure that the information you are receiving is vetted with careful skepticism.

We need to be skeptical...
$>$ If I read a study that said smoking is good for me, I should probably look deeper before going and buying a pack.

- Every study that involves the collection of data (and you shouldn't believe any without data) needs what is called a SAMPLE.


## Sample vs Population

- A population is the totality of all things under consideration
- Example: All the students in school district 61
- A sample is a selection of participants from the population
- Example: Only students at Mount Doug
- In order for a study to be free of bias, or skewed results, you need a true RANDOM SAMPLE.


## Example:

If I were looking to do a survey of people in Victoria who think we need to improve the BC transit system, I need to make sure my sample is random.

- If I asked people at bus stops, will that have a bias on my results?

Absolutely! Since you are asking people at the bus stop, you can assume they ride the bus regularly and therefore have a skewed response regarding BC transit.

How do you get a Random Sample?
There are many methods:

- Open the phonebook to a random page
- Put all the people's names in a hat and draw them out
- Use a random selection computer generation program


## Sampling Methods

Convenient Sampling - Just ask people that are around
Example: TV journalist asking people on the street
Bias: Opinion limited to individuals present

Voluntary Response Sampling - Individuals choose to be involved.
Example: Answering a poll on a radio call in show
Bias: $\quad$ Not considered valid or scientific
$\underline{\text { Television/On-line Survey - Limited to people who watch or use specific shows/websites }}$
Example: $\quad$ Filling out a questionnaire on-line

Bias: Very opinionated, and usually does not reflect the general public

Random Samples - Made of randomly selected individuals. Every person in the population has the same chance of being in the sample. The larger the sample, the closer the random sample is to representing the population.

Unbiased: Sampling randomly gets rid of bias

## Influences on Sampling Methods

Bias - A sample that takes items from a wrong population
Example: Online and phone-in polls. Only highly motivated people will respond, typically those with strong opinions

Use of Language - How a question is phrased could affect the answer that is given
Example: "Were you bullied in school?" vs "Did you have any bad experiences in high school?"

Ethics - Using data for a purpose other than its intended use

Example: Obtaining e-mail addresses for a survey and then sending unsolicited messages
Privacy - Data information gathered from participants must only be used as intended

Example: Social Networking websites allow you to choose who gets to see your personal information

Cost - The cost of collecting data should be worth the effort, and the provider of the money has to have no ties to the study

Example: A cigarette company offers to fund your research on smoking because the data and sample size is too large

Cultural Sensitivity - Being aware of cultural differences and similarities, they may affect the values and behaviour of individuals answering questions

Example: Living in a multicultural society we have to be aware of cultural differences

## Central Tendency

- Is the Central of Typical value with respect to the distribution of the statistics
- We will look at and discuss these three:
- The MEAN - The average of the values you have collected
- The MEDIAN - The middle term of the list of data, written smallest to largest
- The MODE - The term that occurs the most in a given set of data

The Median is often used in describing data compared to the Mean in statistics because it is not skewed by outliers that are excessively high and/or low, and may give a better idea of the typical value.

Example 1: Calculate the Mean, Median, and Mode of the following data set.

$$
12,13,20,10,15,17,21,37,22,13,10,17,21,21,37
$$

Solution 1: Rewrite the info in ascending order first to find the MEDIAN AND MODE

$$
10,10,12,13,13,15,17,17,20,21,21,21,22,37,37
$$

So now we can see that the MODE (MOST FREQUENT) is: 21 (it appears three times)

So now we can see that the MEDIAN (MIDDLE) is: 17 (7 items on either side of it)
*With an even number of data pieces, you get the median by averaging the two middle values*

So now we can calculate the MEAN (AVERAGE) is: 19.07 (Add all the terms up and divide by the total number)

## Interpolating and Extrapolating Graphs

- Whenever you have collected data, that information can be used to create a graph
- Some different types of graph are:
- Line Graphs
- Bar Graphs
- Pie Graphs
- Graphs can be a handy way to represent your information visually
- Interpolating a graph is using the graph to take information at various points
- Extrapolating is using the graph to identify a trend a predict future behaviour

Below is a graph of the $\mathrm{CO}_{2}$ Emissions in parts per million, the data they have gathered goes back to 1880.


- We can EXTRAPOLATE by identifying a trend and predicting future behaviour
- Example: What do you think the temperature increase will be in 2020?

Solution: Draw a trend line and identify the results ------ Approx. 0.6 degree rise

- Both interpolating and extrapolating come with significant caveats. We need to look at trends to predict behaviour, but we also have to consider if/when we make behaviour changes to change a trend line.
- Interpolating can include asking questions like: "Why did the data change here, what was happening?"
- Extrapolating can include questions like: "If we drastically reduce emissions, how will this effect the trend we are seeing in the graph? How can it change our prediction?"


## Section 4.1a - Practice Problems

1. In your own words, explain the difference between and population and a sample. Give examples of how they would vary.
2. Name 3 different sampling methods. Which methods is the best to use and why?
3. What does it mean to have 'bias' or influence on statistics and what two examples you can think of and why?
4. Determine the Mean, Median and Mode salaries of the staff listed below.

| Staff | Salary |
| :--- | :---: |
| One Owner | $\$ 80000$ |
| One Manager | $\$ 60000$ |
| Two Sales <br> People | $\$ 48000$ Each |
| Six Technicians | $\$ 44000$ Each |

5. The incomes of 6 people are listed below, what is the Mean, Median, and Mode of their salaries?
\$41 500, \$44 900, \$39 700, \$62 300, \$58 500, and \$53 100
6. Given the following data:

## Daily cases, hospitalizations, and deaths Jan 1 to Dec 9, 2021


7. What information can you interpolate from the graph? What is the graph attempting to tell us, provide at least 3 pieces of information?
8. Using the same graph, what can you extrapolate with respect to trends. What factors may affect the possible trend?

