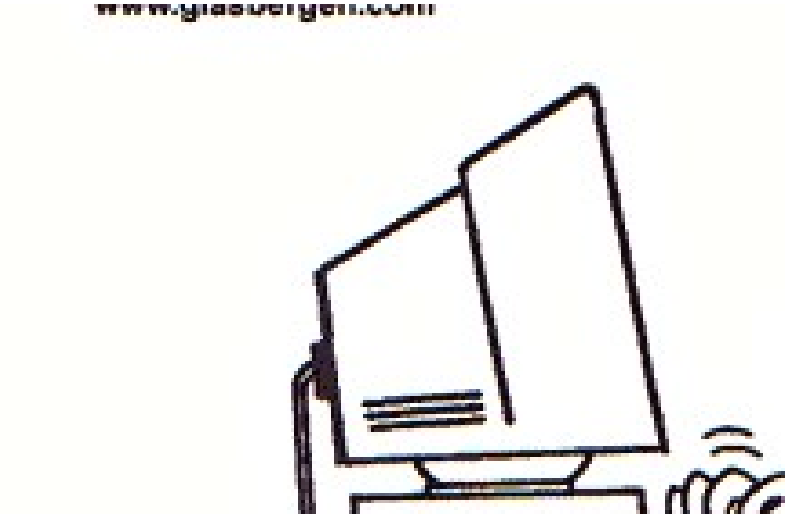
MBF 3C1 Name: _____

UNIT 7

Statistics & Probability



MBF 3C1 Name: _____

Sampling Techniques

Statistics involves _____, _____, and _____ data. Statistics are often used by businesses, advertisers, governments, and the media, both to inform us and persuade us.

In the context of mathematics, a _____ includes all members (people or objects) of a group that possess common characteristics from which information is being collected. A _____ is a part of the population chosen for participation in a study. Samples are often used because they are _____, _____ and _____. The accuracy of any statistical study depends on how the sample is chosen. The following are some of the different sampling methods available to researchers:

- cluster sample
- convenience sample
- simple random sample
- stratified sample
- systematic sample
- voluntary sample

SAMPLING METHOD	DEFINITION	EXAMPLE
	Every item in the population has an equal chance of being selected.	Drawing five names to survey from a hat containing 30 names.
	The population is divided into subgroups (by age, gender, nationality, etc.) and a random sample is selected from each subgroup in proportion to its size in the population.	A school is divided into 4 groups by grade. There are 300 grade nines, 350 grade tens, 270 grade elevens and 320 grade twelves. 10% of each group chosen to be a part of the sample.
	The population is divided into clusters and a certain number of clusters are chosen. Every member of these clusters is part of the sample.	A VP enters the cafeteria and randomly selects two tables. All students at those two tables are surveyed.
	The sample contains those members of the population from which data are most easily collected.	To survey woodworkers in Ontario, we ask people at several lumber yards and home improvement stores scattered about the province.
	The sample contains those members of the population who have chosen to respond to the survey. Often a reward is offered to those who participate in the survey.	The psychology students at University of Toronto are given an extra 2% at the end of the year if they volunteer for any two upper-year psychology surveys.
	Every n^{th} member of the population is selected.	Ronald McDonald hands out coupons for free Cheeseburger Happy Meals to every 10^{th} kid who enters the restaurant.

MBF 3C1 Name: _____

Example 1

Alicia wants to know which band Ontario high school students think is the best. Alicia's friend Jason goes to a different school, so they each survey students at their own school. Alicia uses the completed surveys from both schools to draw conclusions.

a. Identify the population and the sample.

b. Is the sample representative of the population?

Example 2

Determine the best sampling technique for each survey. Provide a reason for your answer

a. The school newspaper wants to determine which presidential candidate in the upcoming student council elections is supported by the majority of students.

b. A light bulb manufacturer wants to determine the lifespan of a certain type of light bulb, in hours.

c. The Parent – Teacher Association wants to determine the average number of hours per week that students spend on homework.

d. The producers of "Canadian Idol" want to determine which of the two remaining candidates should be the next Canadian Idol.

Example 3

There are 570 students taking mathematics this semester. The table below shows the number of math students in each grade. A total of 90 math students are to be surveyed using a stratified random sample. How many students from each grade level should be surveyed?

Grade	# of Math students	% of Math students	# of students to be surveyed
9	115		
10	125		
11	150		
12	180		
total	570	100	90

MBF 3C1 Name: _____

To make an accurate prediction about a population, it is necessary to use a **representative sample**. The characteristics of a representative sample are:

- _____ - every member of the population has an equal chance at being selected
- _____ - fair and impartial.

Bias is the intentional or unintentional prejudice of data collected in a survey. There are several types of bias:

TYPE OF BIAS	DEFINITION	EXAMPLE
Sampling Bias		A survey asks students at a high school football game whether a fund for extra curricular activities should be used to buy new equipment for the football team or instruments for the school band.
Response Bias		A group of professional football players are asked if they have ever taken banned performance enhancing substances.
Measurement Bias		A highway engineer suggests that an economical way to survey traffic speeds on an expressway would be to have police officers who patrol the highway record the speed of the traffic around them every 30 minutes.
Non-Response Bias		You hand out surveys to your classmates to be returned to you next week.

Example 4

Favourite Subject Survey

☺ Sponsored by the "Friends of Mathematics" Society ☺

Circle the most appropriate response.

Gender: Male Female

Grade: 9 10 11 12

My favourite subject is: MATH English Science

Please return the survey by the end of the day.

a. Identify the types of bias that *might* result from the survey.

b. Re-write the survey in a way that does not present the bias identified in part (a).

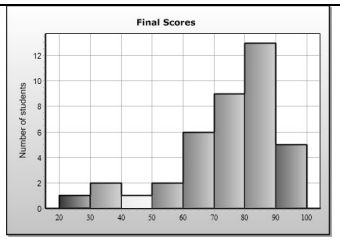
MBF 3C1 Name: _____

Collecting & Organizing One-Variable Data

One-variable data sets give measures of one attribute. They can be represented with:

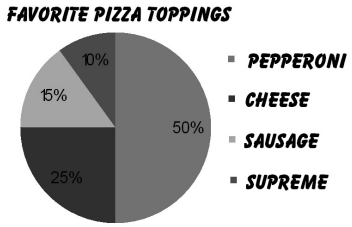
- **Favorite Color**

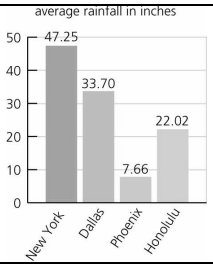
Color	Tally
red	𝍸
orange	III
yellow	II
green	III
blue	𝍸 I
indigo	II
violet	I

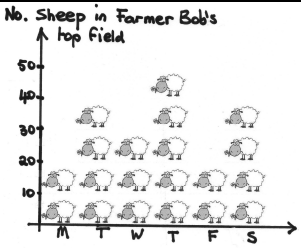
- 

-

Mark	Frequency
0 -10	2
10 -20	4
20 -30	7
30 -40	14
40 -50	23
50 -60	13
60 -70	8
70 -80	6
80 -90	2
90 -100	1
Total	80

- 

- 

- 

TYPE OF DATA	DEFINITION	EXAMPLE
Categorical Data	Data that is usually recorded as a _____ and not a _____. When recorded as a number, it is important to know what the number *represents* not its numerical value.	
Continuous Data	This is numerical (or quantitative) data where values can exist_____ recorded values, so _____ are allowed	Class interval / Tal 0 - 39 / I 40 - 79 / 𝍸 80 - 119 / 𝍸 𝍸
Discrete Data	This is numerical data, where values _____ exist between recorded values, so decimals are _____. There is a fixed number of possible values.	# of times / Tall 0 / |||| 1 / |||| | 2 / |||| || 3 / |||| || 4 / ||||

MBF 3C1 Name: _____

Example 1

For each state the data type.

a. Number of mugs of coffee drank in a day. _____

b. Type of pet at home (eg. dog, cat, bird, reptile, etc.) _____

c. Number of pets at home. _____

d. Amount of coffee in mL drank in a day. _____

For *categorical* and *discrete data*, **classes** are used to sort the data in a frequency table.

Example 2

Organize the following data about favourite types of movies into a frequency table.

sci fi	romance	comedy	action	romance	drama
romance	sci fi	action	romance	comedy	sci fi
romance	action	sci fi	comedy	romance	action
action	comedy				

CLASS	TALLY	FREQUENCY
Total		

MBF 3C1 Name: _____

For *continuous data*, **class intervals** are used to sort the data in a frequency table. When making frequency tables with class intervals,

- make sure intervals don't overlap by using decimals if necessary
- have a reasonable number of intervals, not too few nor too many

1. find the range (highest value – lowest value)
2. find interval length (divide range by 5 and 20 (the min and max interval length))

Example 3

Use a frequency table to organize the pulses of 30 people.

66	79	53	81	84	76	76	67	64	83
92	56	67	77	91	61	71	86	73	87
71	67	71	81	86	62	77	91	72	68

range = 92 - 53
= 39

39 ÷ 5 = 7.8 and 39 ÷ 20 = 1.9, so let the interval be 5

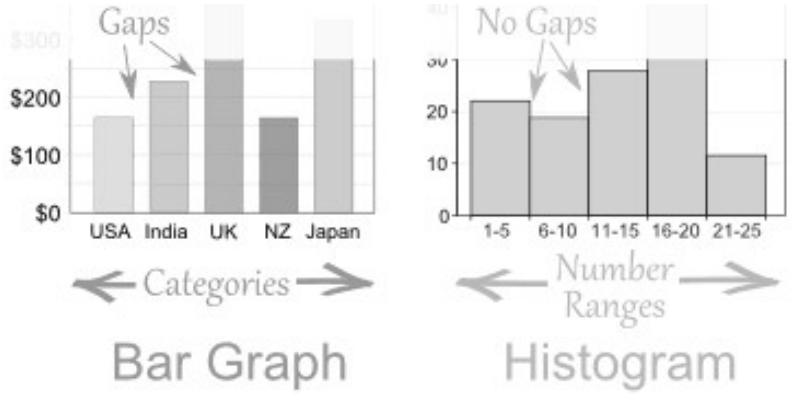
INTERVAL	TALLY	FREQUENCY
Total		

MBF 3C1 Name: _____

Displaying One-Variable Data

A **bar graph** uses bars to display discrete data which has been organized into classes. The bars have spaces between them since the data is _____.

A **histogram** uses bars to display continuous data which has been organized into intervals. The bars in a histogram touch since the data is _____.



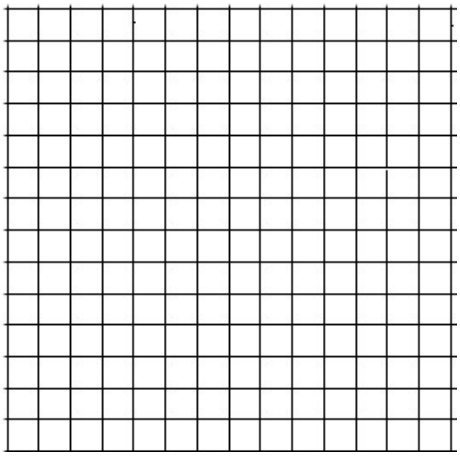
Bar Graph Histogram

Example 1

Use the frequency table from Example 2 (about movies) from DAY 2 lesson to construct a bar graph.

Process:

1. organize data into frequency table with classes if necessary
2. draw and label axes
   * labels go under bars
3. plot bars
   * bars must have spaces between them
4. add a title

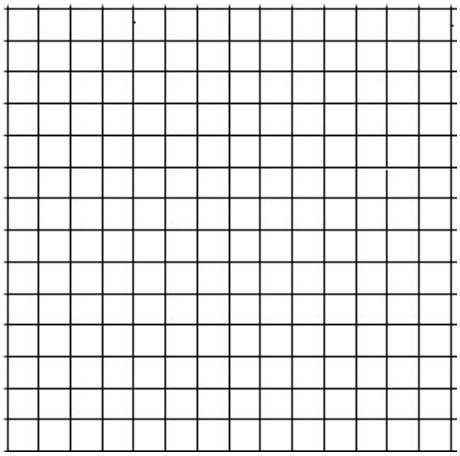


Example 2

Use the frequency table from Example 3 (about pulses) from DAY 2 lesson to construct a histogram.

Process:

1. organize data into frequency table with class intervals if necessary
2. draw and label axes
   * labels go between bars
3. plot bars
   * bars cannot have spaces between them
4. add a title



MBF 3C1 Name: _____

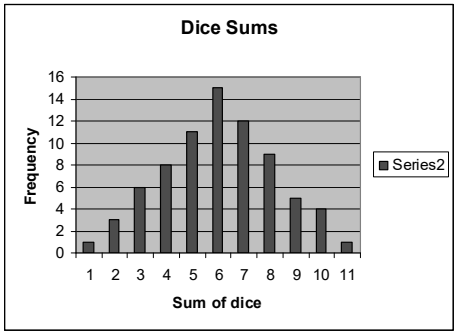
Bar graphs and histograms can take on any of several common shapes. Among these distributions are both ***SYMMETRICAL*** and ***SKEWED*** graphs.

SYMMETRICAL DISTRIBUTIONS can be:

1. **_____ Distributions**
   - commonly referred to as _____ or mound-shaped distributions
   - the middle interval(s) will have the greatest frequency (i.e. the tallest bar)
   - all other intervals will have decreasing frequencies as you move away from the centre of the graph (i.e. the bars get smaller as you move out to the edges)

   Example: A pair of dice were rolled 75 times. After each roll, their sum was recorded and graphed.

Sum on dice	Frequency
2	1
3	3
4	6
5	8
6	11
7	15
8	12
9	9
10	5
11	4
12	1

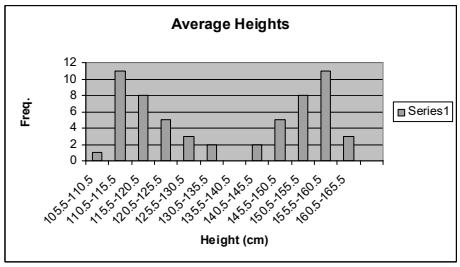


Note: Even though it isn't perfectly symmetrical, it still fits the definition of a normal distribution.

2. **_____ Distributions**
   - these look like inverted normal distributions
   - the intervals with the highest frequencies (i.e. tallest bars) are at either end of the graph and the interval with the lowest frequency is in the centre
   - frequencies increase as you move away from the centre of the graph.

   Example: Grade 6 and grade 1 students each measured, recorded and graphed their heights.

Height (cm)	Freq.
105.5-110.5	1
110.5-115.5	11
115.5-120.5	8
120.5-125.5	5
125.5-130.5	3
130.5-135.5	2
135.5-140.5	0
140.5-145.5	2
145.5-150.5	5
150.5-155.5	8
155.5-160.5	11



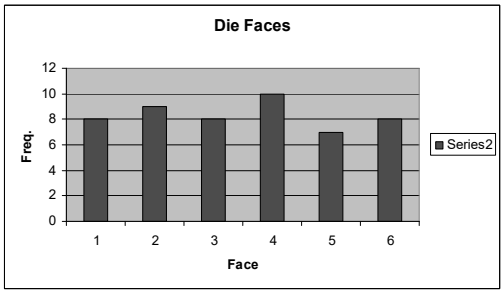
MBF 3C1 Name: _____

3. _____ **Distributions**

- the frequencies of each interval are approximately equal

Example: A die is rolled 50 times. The face is recorded and graphed.

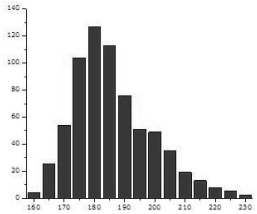
Die Face	Frequency
1	8
2	9
3	8
4	10
5	7
6	8



SKEWED DISTRIBUTIONS can be:

1. _____ **Graphs**

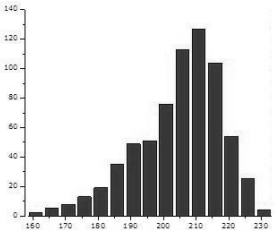
- the bars with the highest frequencies are on the left side and the frequencies decrease as you move right



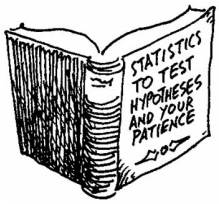
Note: Even though there is a low-frequency bar on the left side, the trend is still right-skewed.

2. _____ **Graphs**

- the bars with the highest frequencies are on the right side and the frequencies decrease as you move left



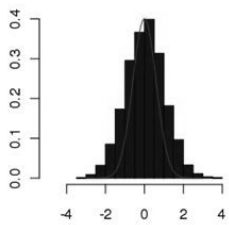
Note: Even though there is a low-frequency bar on the right side, the trend is still left-skewed.



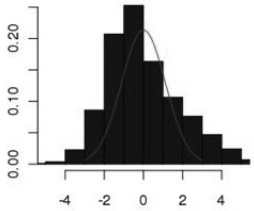
Example 3

Describe the distribution for each graph.

a.

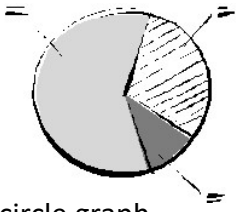


b.



MBF 3C1 Name: _____

A **circle graph** uses sectors of a circle to show how discrete data is divided.

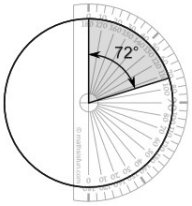


Example 4

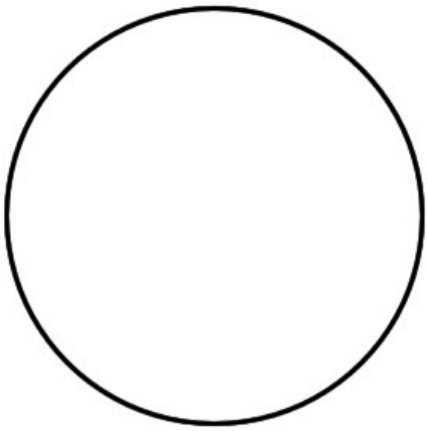
Use the frequency table from Example 2 (about movies) from DAY 2 to construct a circle graph.

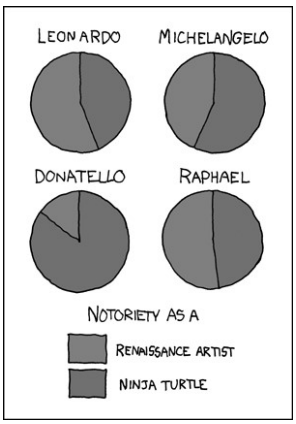
Process:

1. organize data into frequency table with classes if necessary
2. calculate the percent of each class
   * leave as decimal for future calculations
3. use the percentages to determine how many degrees each class consists of
4. draw a circle and divide it according to the calculations using a protractor
5. either label each section or provide a legend
6. add a title



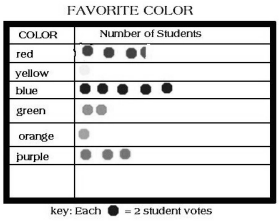
CLASS	FREQUENCY	PERCENT	DEGREES
action	5	$\frac{5}{20} = 0.25$	$0.25 \times 360° = 90°$
comedy	4		
drama	1		
romance	6		
sci fi	4		
TOTAL	**20**	**1.0**	**360°**

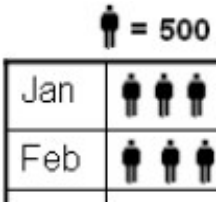




MBF 3C1 Name: _____

A **pictograph** uses images to display discrete data. Often 1 image represents more than 1 data piece. This type of graph is less accurate than other graphing styles.





Example 5

Use the frequency table from Example 2 (about movies) from DAY 2 to construct a pictograph.

Process:

1. organize data into frequency table with classes if necessary
2. draw and label axes
   * labels go under columns or beside rows of pictures
3. plot pictures
   * each picture represents a number of data pieces - this value must be stated
4. add a title

MBF 3C1 Name: _____

Measures of Central Tendency & Spread

The **MEASURES OF CENTRAL TENDENCY** are numbers that refer to the middle value in the data. Measures of central tendency are most useful when comparing sets of data that are similar.

The _____ is the average (the total of all values divided by the number of values). It is the most common way of finding the central number.

The _____ is the middle value in a list arranged in numerical order. It is used when two or more numbers are very different from the rest of the group. It emphasizes the position of the number rather than the value. If there is an even number of data, there are two middle numbers. In this case, average the two middle numbers to find the median.

The _____ is the value that occurs most often. It is used when it is important to know how often a number occurs rather than its value. When more than one number appears most often, there is more than one mode. If no number appears most often, there is no mode.

Example 1

David bowls for a youth league. Here are his scores for the past season: 135, 148, 120, 158, 162, 192, 162, 178, 162, 150, 138.

a. Calculate David's **mean** bowling score.

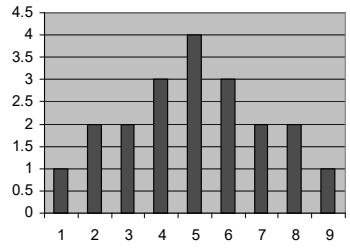
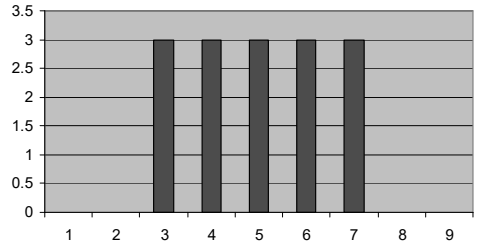
b. Find the **median** for David's bowling scores.

c. Identify the **mode** for David's bowling scores.

"Add the numbers, divide by how many numbers you've added and there you have it-the average amount of minutes you sleep in class each day."

MBF 3C1 Name: _____

Sometimes measures of central tendency don't explain the differences in data sufficiently. For example, the following sets of data both have a mean of 8 and a median of 8.

Set A	**Set B**
0, 2, 2, 4, 4, 6, 6, 6, 8, 8, 8, 8, 10, 10, 10, 12, 12, 14, 14, 16	4, 4, 4, 6, 6, 6, 8, 8, 8, 10, 10, 10, 12, 12, 12
	

When considering the calculations of mean and mode, the two data sets appear very similar. When looking at the graphs, they are quite different; Set A has a _____ distribution while Set B is _____.

The **MEASURES OF SPREAD** show how far apart the data are. Measures of spread are used to help compare data sets that vary from each other.

The _____ is the difference between the highest and lowest values in a set of data. It can be misleading if there are extremely high or low values that don't follow the rest of the data.

The range can be useful to compare data sets which appear similar. Consider sets A and B above.

Range of Set A	**Range of Set B**

Set ___ is more consistent since it has a _____ range.

The _____ is the measure of dispersion of the data. It is the average of the squared differences from the mean.

The _____ _____ is the best measure of the spread of data and is represented with the symbol σ (sigma). It is calculated by finding the square root of the variance. The smaller the standard deviation, the closer it is to the mean of the data and better it represents the data.

MBF 3C1 Name: _____

Example 2

Toby and Polly both work at a local pizza shop. Their manager has decided to give a raise to her best employee. She looks at the number of pizzas each made during their shifts.

Toby	54	152	180	12	72	126	104	132
Polly	132	104	102	120	86	12	180	96

The manager calculates the mean and range to help her make her decision.

	Toby	**Polly**
MEAN	$\frac{832}{8} = 104$	$\frac{832}{8} = 104$
RANGE	180 – 12 = 168	180 – 12 = 168

Unfortunately, these statistics leave both employees equal. The manager notices that Polly's data looks more consistent, but she needs proof to support her claim. Help the manager decide who is more deserving by calculating the standard deviation for each.

CALCULATING STANDARD DEVIATION BY HAND	**CALCULATING STANDARD DEVIATION USING THE TI-83+**
1. Find the difference between each value and the mean of the data. 2. Square each difference. 3. Add up all of your answers from Step 2. 4. Find the variance → the average of the differences squared. 5. Find the standard deviation → Find the square root of your answer in Step 4.	1. Press **STAT** and then **1**. 2. Enter the data into L_1 by pressing **ENTER** after each entry. 3. Press **STAT** and cursor right once for **CALC**. 4. Press **1** for 1-Var Stats. 5. Type L_1 by pressing **2nd 1 ENTER**. 6. **Sx =** the sample standard deviation (used when results are to be applied to an entire population, not just the data entered) **σx** = the population standard deviation (used when only data entered should be considered)

MBF 3C1 Name: _____

Solving Problems Using Statistics

Compare sets of data by analysing and interpreting measures of central tendency and measures of spread.

Example 1

Consider the following list of data collected from MLS:

AVERAGE HOME PRICES BY PROVINCE (IN $)

PROVINCE	JUNE 2006	JUNE 2005
Yukon	177191	159668
Northwest Territories	243745	250222
British Columbia	399829	330333
Alberta	294282	215964
Saskatchewan	134161	121984
Manitoba	155531	139195
Ontario	280263	268074
Quebec	284747	252745
New Brunswick	127406	123732
Nova Scotia	170547	157524
Prince Edward Island	134115	114223
Newfoundland/ Labrador	132571	140958

Compare the two years by finding all measures of central tendency and measures of spread.

	2005 DATA	2006 DATA
Mean		
Median		
Mode		
Conclusions		

	2005 DATA	2006 DATA
Range		
Standard Deviation		
Conclusions		

MBF 3C1 Name: _____

Theoretical Probability

Probability is used to predict the outcomes of various events. It is the chance of something happening.

Probability is used to make predictions about _____, _____, _____.

Probabilities can be written in 3 ways:

			Example 1	Example 2		
i. as a fraction	$\frac{\#\text{ of successful attempts}}{\text{total }\#\text{ of attempts}}$		$\frac{1}{2}$			$\frac{\#\text{ of successful attempts}}{\text{total }\#\text{ of attempts}}$
ii. as a decimal	divide the fraction to convert to percent					use place value to convert to fraction
iii. as a percent	multiply by 100 to convert to percent			60%		divide by 100 to convert to decimal

When two or more things have the same probability of happening, they are considered to have **equally likely outcomes**.

Theoretical Probability is the chance of something happening in a perfect world. It can be calculated using the formula:

$$P = \frac{\#\text{ of favourable outcomes}}{\#\text{ of possible outcomes}}$$

Example 1

A number from 1 to 50 inclusive is chosen at random. What is the probability that the number . . .

a. is even?		b. ends in a 3?	
c. is odd?		d. does not end in a 3?	
e. is greater than 13?		f. is prime?	
g. is divisible by 5?		h. has 2 digits the same?	

MBF 3C1 Name: _____

Example 2

The Toronto Maple Leafs are playing the Detroit Redwings in the 2003 Stanley Cup Finals. The probability of the Maple Leafs winning are $^{4}/_{11}$.

a. What is the probability of each team winning a single game? Write your answers as percentages.

b. If these teams were to play 6 games, how many is each expected to win?

Example 3

Suppose you toss a coin three times.

a. Which event do you think is more likely: you get 3 heads, or you get 1 head and 2 tails? Explain your thinking.

b. Draw a tree diagram to show the possible outcomes when a coin is tossed three times. Are the outcomes equally likely?

c. Use your diagram to determine the probability of each event.
   i. 3 heads ii. no heads iii. 1 head and 2 tails

MBF 3C1 Name: _____

Experimental Probability

Experimental Probability is the chance of something happening based on experimental results. It can be calculated using the formula:

$$P = \frac{\#\text{ of favourable outcomes observed}}{\text{total }\#\text{ of observations}}$$

Example 1

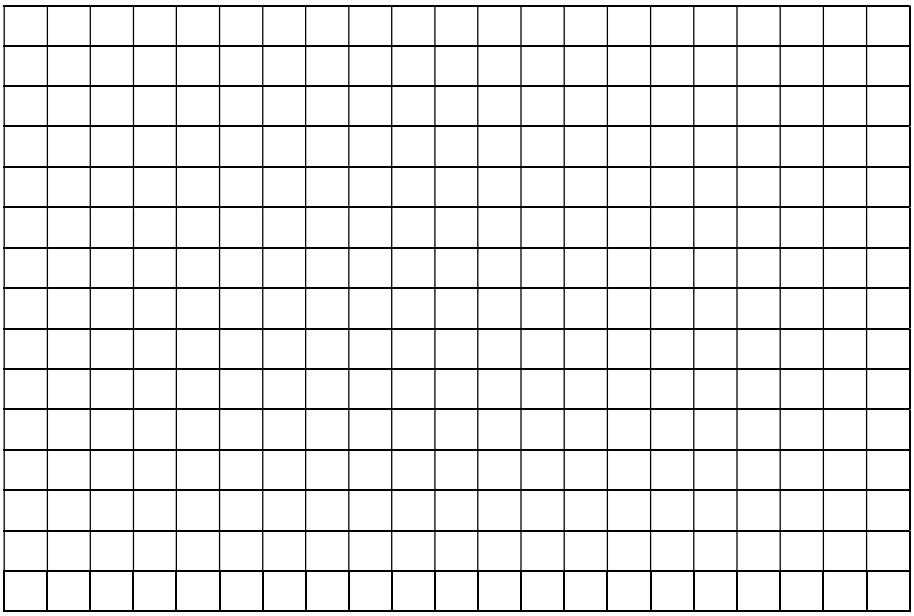
A new cereal is giving away a prize with each box of cereal. There are 6 different prizes and you want to win all of them. How many boxes of cereal do you have to eat to win them all?

Complete an experiment using 1 die to determine if the prizes in the cereal box are worth it. Each roll signifies one cereal box and you must get each prize, numbered 1 through 6. Keep track of how many rolls it takes you to get all of the prizes 1 through 6. Repeat this simulation 10 times.

Trial	Tally of Prizes Received						Total # of Cereal Boxes (frequency)
	1	2	3	4	5	6	
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							

MBF 3C1 Name: _____

1. Graph the data using a bar graph. Put the trial number on the *x*-axis and the frequency on the *y*-axis.



2. What does this experiment tell you about the number of boxes you would have to eat to get all the prizes?

3. On average, how many boxes of cereal must you eat to get all six prizes?

4. How do your results compare with the rest of the class?

MBF 3C1 Name: _____

Comparing Theoretical and Experimental Probability

	Theoretical Probability	**Experimental Probability**
Calculation Method?		
Time Frame?		
When Used?		
Product?		

As the number of trials in an experiment increases, experimental probability often approaches theoretical probability.

Example 1

You roll a die 20 times and observe an even result 13 times.

a. What is the experimental probability of rolling an even number?

b. What is the theoretical probability of rolling an even number?

c. What could be done to make the experimental value more accurate and closer to the theoretical value?

MBF 3C1 Name: _____

Example 2

An ice cream shop offers multiple options. The ice cream can be served on a chocolate dipped cone or a plain cone, with chocolate, raspberry or vanilla ice cream. A customer can also request sprinkles or no sprinkles. All ice cream cone options have an equal chance of being chosen.

a. Create a tree diagram to list all possible ice cream cones.

b. How many ice cream cone combinations can be made?

c. What is the probability that a customer orders a raspberry ice cream cone?

d. What is the probability that a customer orders a chocolate ice cream cone with sprinkles?

MBF 3C1 Name: _____

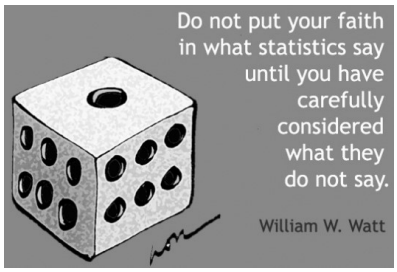
Statistics & Probability in the Media

What is the connection between statistics and probability?

Where are statistics and probability are used in everyday life?

- _____
- _____
- _____

Statistics and probability are used to provide information and often this information is meant to influence you. Graphs can be drawn and statements can be made to create false impressions. For this reason, it is important to make sure you have a good understanding of the graph or data so that you are not misled.



Things to consider when statistics and probability are used:

Misuse of language	• The words *average* or *typical* are sometimes used without identifying whether the number used is the mean, median, or mode. • Survey questions can be biased. • Relevant information can be left out if it does not get across the desired information.
Distorted visuals	• In some pictographs or 3-D graphs, the sizes of parts of the graph can make numbers appear greater or less than they are. • When axes do not start at 0, it is easy to conclude that differences between numbers are greater than they are. • Statistics can be fabricated or exaggerated and may not come from mathematical analysis.
Questionable Sources	• Do the data come from a random, unbiased sample: What is the population? How was the sample chosen? When, where and how was the survey conducted? What were the questions in the survey - where they biased? • The word *expert* implies that a person has a great deal of knowledge. Ask what may make a person an expert and whether they are an expert in the appropriate field? • It is important to distinguish whether the presented data shows facts or opinions. Just because people believe something doesn't make it true.

MBF 3C1 Name: _____

Example 1

Examine the following statistics and determine where they try to mislead.

a. 4 out of 5 doctors recommend *Sparkle Toothpaste* for your family.

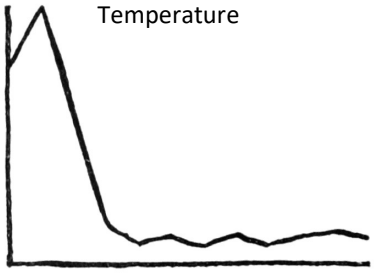
b. *Taylor Towels* are 30% stronger.

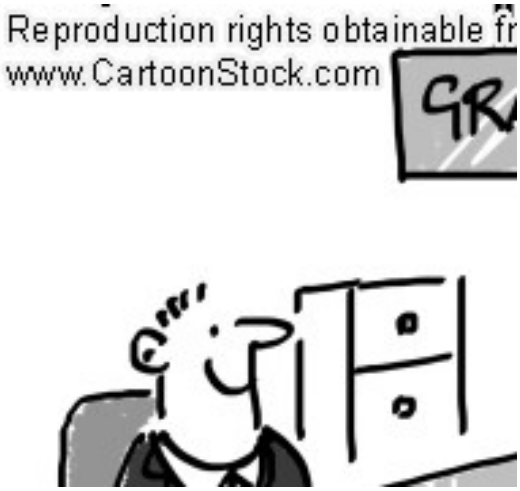
c. Lee earned $1000 a week selling our product door-to-door. You can too!

d. CDs on sale, from $8.99.

e. Last year there were 55 motorcycle accidents involving people ages 16 to 30 and only 25 motorcycle accidents involving people over 30. People under 30 are careless motorcycle drivers.

f.





MBF 3C1 Name: _____

Example 2

GM's OnStar service is a wireless communication system between a central call centre and an OnStar enabled car, allowing drivers to call for help or ask for other services. The volume of monthly OnStar activity reported follows:

- 1,000 airbag deployment calls
- 300 crash notifications
- 11,000 emergency service calls
- 5,400 'Good Samaritan' calls; Orange Alerts; someone in trouble etc.
- 325 stolen vehicle location assists
- 41,000 remote unlock calls
- 24,000 requests for roadside assistance
- 329,000 requests for route assistance
- 412,025 total calls

Determine the probability of each type of call occurring.

TYPE OF CALL	NUMBER OF CALLS	TOTAL CALLS	PROBABILITY OF CALL
airbag deployment		412,025	
crash notifications			
emergency service			
'Good Samaritan'; Orange Alerts; someone in trouble			
stolen vehicle location assists			
remote unlock			
roadside assistance			
route assistance			

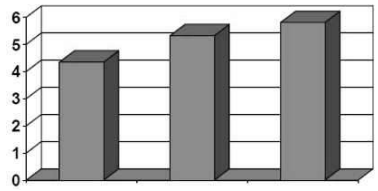
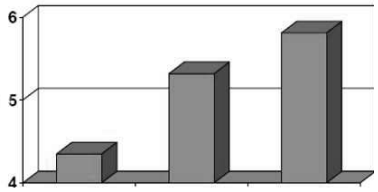
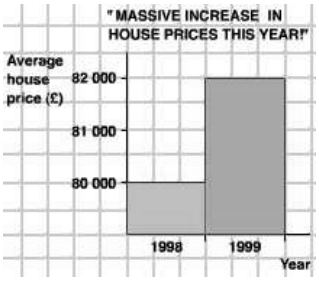
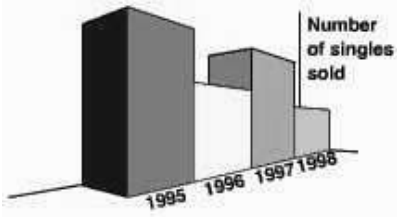
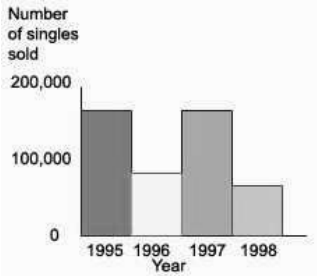
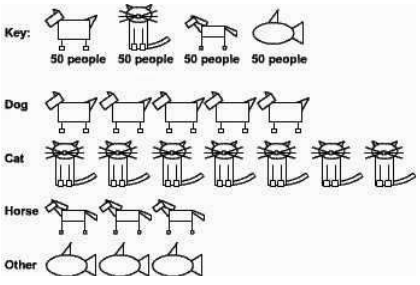
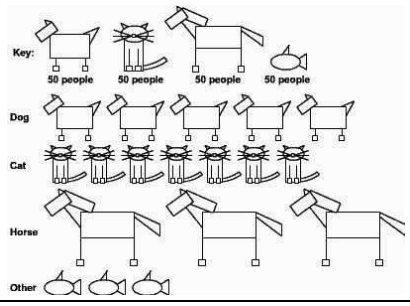
a. Based on what you've seen in the media what type of headline might you see in a newspaper based on the statistics given above?

b. How accurate might those headlines be based on what you know about statistics?

MBF 3C1 Name: _____

Example 3

Compare each set of graphs. Which is each misleading? Why?

	GRAPH 1	GRAPH 2	MISLEADING BECAUSE . . .
a.			
b.			
c.			
d.			

MBF 3C1 Name: _____

EXTRA: Driving To Work Probability Experiment

You and your brother drive to school together. Each week you fight over who will pay the $25 for gas each week. You decide to toss two coins. If the coins are both heads, or both tails, your brother pays for gas. If the coins are different, you pay.

1. What are the possible outcomes? Is the coin toss for gas a fair game? Explain.

2. How much should you expect to pay in a four week period? Explain.

3. Each week there are two outcomes: you pay or your brother pays. Draw a tree diagram to illustrate the outcomes of who pays over a four week period.

MBF 3C1 Name: _____

4. What is the theoretical probability that you will pay _____ in a four week period?

a) zero times b) once c) twice d) three times e) four times

5. Complete the chart to simulate who pays each week for a year (52 weeks).

Week #	1	2	3	4	5	6	7	8	9	10	11	12
Toss Result												
Who Pays												

Week #	13	14	15	16	17	18	19	20	21	22	23	24
Toss Result												
Who Pays												

Week #	25	26	27	28	29	30	31	32	33	34	35	36
Toss Result												
Who Pays												

Week #	37	38	39	40	41	42	43	44	45	46	47	48
Toss Result												
Who Pays												

Week #	49	50	51	52
Toss Result				
Who Pays				

MBF 3C1 Name: _____

6. From the chart, what is the experimental probability that you will pay _____ in a four week period?

a) zero times b) once c) twice d) three times e) four times

7. Compare your answers from question 6 to your answer from question 4. Explain the results.

a) zero times	b) once
c) twice	d) three times
e) four times	

MBF 3C1 Name: _____

EXTRA: Collect, Organize, Display & Analyse Data

Collect data for each question.

1. How old are you?

CLASS	TALLY	FREQUENCY
14		
15		
16		
17		
18		
TOTAL		

2. How tall are you?

INTERVAL	TALLY	FREQUENCY
Under 5'		
5' - 5'5"		
5'6" - 5'9"		
6' - 6'5"		
6'6" - 6'9"		
7' & Over		
TOTAL		

3. What colour are your eyes?

CLASS	TALLY	FREQUENCY
brown		
blue		
green		
hazel		
other		
TOTAL		

4. How many siblings do you have?

CLASS	TALLY	FREQUENCY
0		
1		
2		
3		
4		
5		
TOTAL		

5. What size sweater do you wear?

CLASS	TALLY	FREQUENCY
XS		
S		
M		
L		
XL		
TOTAL		

6. What is your overall average in math?

INTERVAL	TALLY	FREQUENCY
Under 50%		
50% - 59%		
60% - 69%		
70% - 79%		
80 - 89%		
90 - 99%		
TOTAL		

7. How many times are you absent or late for math class per week?

CLASS	TALLY	FREQUENCY
0		
1		
2		
3		
4		
5		
TOTAL		

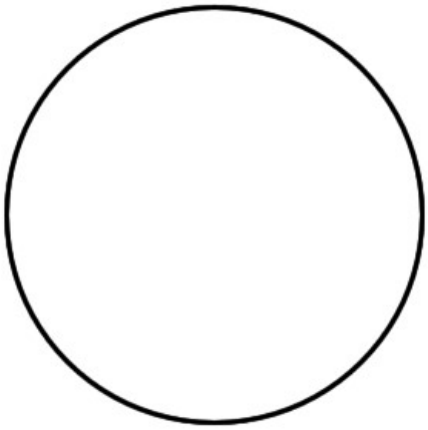
8. How many hours do you spend on math homework per day?

INTERVAL	TALLY	FREQUENCY
Less Than 1		
1		
2		
3		
More than 3		
TOTAL		

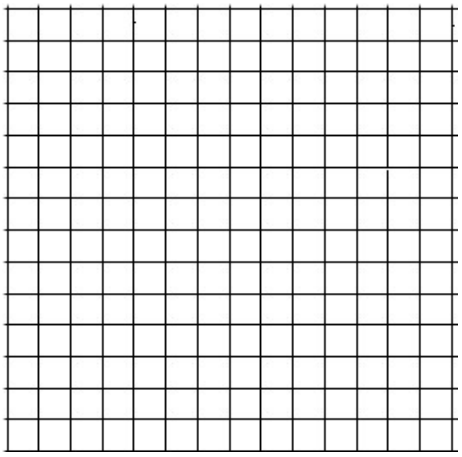
MBF 3C1 Name: _____

1. Choose one data set to represent with a pictograph.

2. Choose one data set to represent with a circle graph.



3. Choose one data set to represent with a bar graph.
    a. Create the graph.



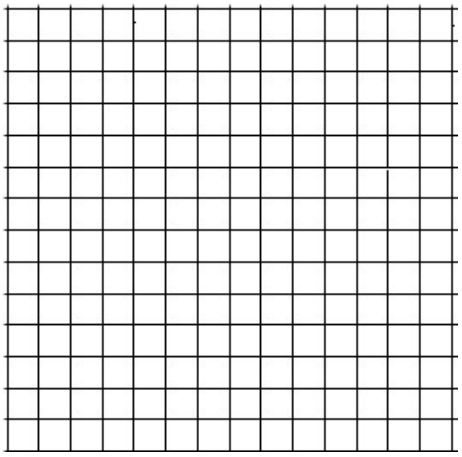
    b. Describe the distribution of the data.

    c. Calculate the measures of central tendency and spread for the data.

    d. What do these measures tell you about the class?

MBF 3C1 Name: _____

4. Choose one data set to represent with a histogram.
    a. Create the graph.



    b. Describe the distribution of the data.

    c. Calculate the measures of central tendency and spread for the data.

    d. What do these measures tell you about the class?

MBF 3C1 Name: _____

EXTRA: Conducting A Survey

Suppose you need data about the opinions and habits of people in your school. You will need to collect your own data.

1. **Designing the Questionnaire**
   - Choose a topic. If possible, choose a topic where the information you collect might help to achieve some goal or to make a decision.
   - Write the questions. Keep these guidelines in mind:
     - → Make the questions short and easy to understand.
     - → Respect the respondents' privacy: avoid asking questions unrelated to the survey.
     - → Avoid questions that may bias the results of a survey or provide vague responses.
     - → Include a few questions that collect demographic data such as age, gender, and grade.
     - → Questions can often be asked in an open way or with choices that help organize the responses. For example,

Open	With choices
How much did you spend on entertainment last week? _____	How much did you spend on entertainment last week? Less than $16 ____ $16–$30 ____ over $30 _____
What is your favourite subject? _____	Which is your favourite subject? (choose one) Math ____ Science _____ English _____ Tech ____ Other _____ Don't have one ____
Should school cafeterias be banned from selling certain foods? If so, which foods? _____ _____	Foods such as pop, chips, and French fries should be banned in school cafeterias. Strongly disagree _____ Disagree _____ Agree ____ Strongly agree ____ No opinion _____

     - → Organize your questions in a logical way.
     - → Write a brief introduction that explains the survey and how you will use the results. Your introduction should grab the attention of potential respondents and encourage their participation.
     - → Make it clear that the responses will be anonymous.

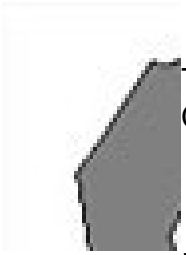
2. **Collecting the Data**
   - What is the population for your survey?
   - What type of sample will you choose?
   - Write a brief explanation of the strategy you will use to conduct the survey.
   - State the pros and cons of your strategy.
   - Carry out your survey.

3. **Organizing and Displaying the Data**
   - Organize the data you have collected in a frequency table.
   - Choose the most appropriate type of graph to represent the data you have collected.
   - Create a visual display to represent the data.

4. **Analyse the Data**
   - Comment on the type and distribution of the data.
   - Calculate the measures of central tendency for the data and explain what it tells you about the data.
   - Calculate the measures of spread for the data and explain what it tells you about the data.
   - State the results of your survey. What is the answer to the question you asked?
   - Were the results what you expected or hoped for? Explain.
   - How could you improve your questionnaire?

MBF 3C1 Name: _____

EXTRA: Bullying—studying it to curb it....



The bullying phenomenon is on the rise in schools. Studies show that 10% of Quebec students are victims of acts of bullying at least once per week.

Is the bullying situation just as disturbing in the rest of the country? Does it affect boys as much as girls? Does it affect younger students as much as older students? Is your class affected by this phenomenon?

In 2004-05, many Canadian students indicated the number of times that they had been victims of bullying at school over the course of the last year.

1. To begin your analysis, obtain a **sample** of 200 students in Canada. Go to the Census at School site (www.censusatschool.ca) and under "Data and Results", click on "Random data selector". Follow the directions to obtain your sample about Canadian schools from 2004-5.

   Then fill in a **frequency table** like the one below:

Group	Number of bullying incidents reported over the course of the last year				
	0	1 to 3	4 to 9	10 or more	Total
Frequency					
Relative Frequency (Freq÷total)					

Relative frequency is like **probability**. Does the probability help to illustrate your data? Why?

You can also use a **graph** to communicate more clearly what appears important to you. Which of two graphs types, the **bar** or **circle** graph, would you use in this case? Why?

MBF 3C1 Name: _____

To better understand the phenomenon and identify the **variables** that influence it, you can analyse the distribution of responses in the **sample** of Canadian students based on some of their characteristics.

2. Start by noting the different responses by *sex* (boy or girl) to assess the effect of this **variable**.

Sex	Number of bullying incidents reported over the course of the last year				
	0	1 to 3	4 to 9	10 or more	Total
Boy					
Girl					
Total					

Based on this table, try to answer the following questions. Make a relative frequency table or a graph to help you answer them:

- **Proportionally**, is the number of bullying incidents that girls and boys are subjected to the same?

- Does the **probability** of being a victim of bullying vary based on whether you are a boy or girl? Explain your reasoning.

MBF 3C1 Name: _____

3. Then, see whether certain *age groups* are more affected than others:

Age	Number of bullying incidents reported over the course of the last year				
	0	1 to 3	4 to 9	10 or more	Total
9-10					
11-12					
13-14					
15-16					
17-18					
19+					
Total					

Use appropriate **graphs** to more easily compare the number of acts of bullying suffered by students of each age group.

- Which age group do you believe is at greater risk?

- Does the phenomenon increase or decrease with age?

MBF 3C1 Name: _____

Could being bullied also be related to the height of students?

Since height **depends** on both sex and age, we must try to keep these other two variables constant to be able to **isolate** the effect of the height variable.

4. The ideal method would be to create a new, sufficiently large sample of students of the same age and then analyse responses from girls and boys separately.

   Since height is a **continuous variable**, it is a good idea to define different classes. To keep it simple, you can define three classes based on the data: short, medium and tall. Your frequency table could look like the one below.

Height	Number of bullying incidents reported for boys 13 years of age				
	0	1 to 3	4 to 9	10 or more	Total
Short					
Medium					
Tall					
Total					

   Once again, use graphs to help you compare the data.

What variable(s) seem most related to the bullying phenomenon?

5. Compare your various tables and graphs to identify which **variable(s) are most related** to the bullying phenomenon. Identifying these relationships allows you to consider a **model**, that is, a simplified representation of the phenomenon and of certain key variables that can describe, explain or even predict it.

MBF 3C1 Name: _____

6. Conduct an anonymous survey of the class.

1. Are you male or female? _____

2. What is your age? _____

3. How tall are you? _____ cm.

4. In the *last month*, how many times have you been bullied at school?

(Being bullied means when someone does or says something to make you feel uncomfortable or afraid.)

Choose one: • 0 • 1 to 3 • 4 to 9 • 10 or more

MBF 3C1 Name: _____

7. To **validate** the results of question 5, now examine your class data. Please note that the students in your class reported the number of times they were bullied at school in the ***last month***.

   a**)** Do the sex, age and height variables appear to play the same role here?

   b) Does your class appear to be particularly affected by bullying?

   c) Remember that the data from the Canadian samples indicated the number of times that each student reported being bullied in the ***last year***. Do you think the fact of reporting the number of bullying incidents in the last year instead of the last month could affect the **accuracy** or the **reliability** of the data? Explain.

   d) What other reasons could explain the differences between your class data and the sample data?

To conclude:

8. Answer the following questions:

   - What did you learn from this analysis?

   - Did this exercise help you identify ways of stopping bullying at your school?