CHAPTER 9

Culminating Investigations

In every province and territory except Nunavut, it is illegal to use a hand-held wireless communication device while driving. Why are these laws put in place? How are data collected to justify the laws? How do insurance companies apply probability theory to calculate insurance rates?

In this chapter, you will learn how to design and complete culminating projects for probability and statistics. Sections 9.1 and 9.2 will show you how to apply probability concepts to design a culminating probability project. Sections 9.3 and 9.4 will demonstrate how to use the research process to complete a culminating statistics project. Both of these projects will require you to design a framework for planning, completing, and presenting your work. Your teacher will provide you with a timeline for your project. Section 9.5 will review how your project will be evaluated and how you will evaluate your peers' projects.

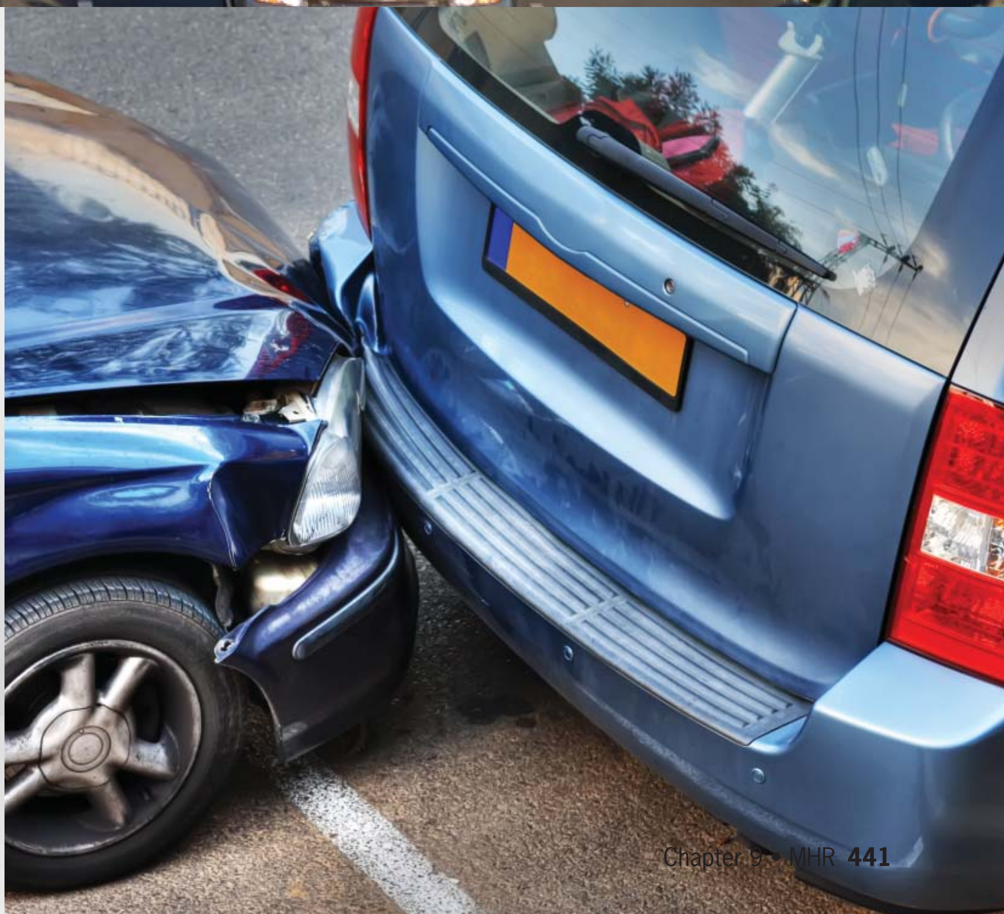
Key Terms

research question
hypothesis

Career Link

Statistician

People who effectively collect, analyse, summarize, and represent data can support law reform, social change, health education, business growth, and more. Samuel Perreault is a statistician who collects data on the incidence of impaired driving rates across the country. The information he collects helps police forces and educators target impaired driving. Statisticians study statistics in a mathematics or social science program at university. How might data collection and analysis be helpful in other sectors, such as marketing, environmental science, health and wellness, travel and tourism, technology, manufacturing, or transportation?



9.1

Planning and Carrying Out a Culminating Probability Project

Learning Goals

I am learning to

- assess games for elements of chance or randomness
- critique games for fairness
- plan a game of chance for a culminating probability project

Board games, lotteries, sweepstakes, and other contests have varying degrees of elements of chance. How can you calculate the probability of winning those games to become an informed player?

You can also use probability to simulate real-life events. For example, you could use Sidney Crosby's historical shooting percentage and simulate him taking 300 shots per season for 10 seasons to see what type of variation there would be in his goal-scoring totals. What are some other ways you could use probability or simulations in a culminating project?

Investigate A Game of Chance

Materials

- 3 dice per group
- 10 counters per player
- 100 counters per dice roller
- playing board with numbers 3 to 18

Triple Your Chances, Double Your Counters is a dice game in which players guess the sum of a roll of three dice.

3	5	7	9	11	13	15	17	18
	4	6	8	10	12	14	16	

1. Read the instructions for the game. In a group, play a few sample rounds, and record your data as you play. One person will be the dice roller and the rest will be players.

Instructions:

- Each player has 10 counters.
- The dice roller has 100 counters.
- A player places **one** counter on a sum on the game board.
- The dice roller rolls the dice.
- If the sum of the dice matches the sum chosen by a player, the player wins. The table on the next page shows the number of counters a player wins for a given sum.

- If the sum of the dice does not equal the sum chosen by the player, the dice roller wins those counters. The game ends when the players have no counters left.

Outcome of the Dice Roll	Counter Payout to the Player
Sum of 9, 10, 11, or 12	1 (win back the counter, plus one more)
Sum of 7, 8, 13, or 14	2 (win back the counter, plus two more)
Sum of 5, 6, 15, or 16	3 (win back the counter, plus three more)
Sum of 3, 4, 17, or 18	9 (win back the counter, plus nine more)

2. In your group, discuss and rank the game for each criterion in the chart.

Criteria	Ranking		
Description of game	Poorly described	Somewhat clear	Makes sense
Uniqueness of game	Not unique	Somewhat unique	Very unique
Complexity	Not complex	Somewhat complex	Very complex
Clarity of rules	Not clear	Somewhat clear	Very clear
Fun factor	Not fun	Somewhat fun	Very fun
Fairness toward players	Hard to win	Some chance to win	Easy to win
Chance vs. skill	All skill	Some chance	All chance

3. For your probability project, you could conduct a number of trials of a game and analyse the experimental and theoretical probability of the player winning. What are the pros and cons of using this game for a probability project, based on your assessment so far?

Investigate Creating a Game of Chance

1. Work with a partner or in a small group. Brainstorm a list of board games, card games, or computer games you have played.
2. Research various kinds of lottery games (Lotto 649), sweepstakes (Heart & Stroke Lottery, Princess Margaret Home Lottery), television game shows (*Deal or No Deal*), or other board games (Monopoly, Risk) to add to your list.
3. Sort the games in your list based on how you win: purely on chance, purely on skill, or both chance and skill.

4. Consider the games in your list that are based purely on chance. Choose one game and answer the following questions to analyse it.
   a) How do you play? Are the rules clearly described? How do you win?
   b) How complex is the game?
   c) How fun is the game?
   d) How is the game marketed to consumers?
   e) Is the probability of winning reported on the game? If so, report the probability of winning.
   f) Does the game seem fair to the players? Why or why not?
   g) What kind of data might the game generate?
   h) What other interesting information do you notice about this game?
   i) What elements/materials from this game could be used to create a new game of chance?

5. a) Is it possible to represent all the outcomes for your game in a chart or diagram? Why or why not?

   Example: In the lottery game Lotto 649, players choose six numbers between 1 and 49. They can choose repeat digits, so there are 49^6 or 13 841 287 201 total six-number combinations possible. There are too many outcomes to show them all in a list, chart, or diagram.

   b) Brainstorm how to calculate the theoretical probability of winning the grand prize.

   Example: In Lotto 649, only one set of six numbers wins the grand prize. There is only one way to win the game out of 13 841 287 201 possible outcomes.

   So, the theoretical probability of winning is $\frac{1}{13\,841\,287\,201}$.

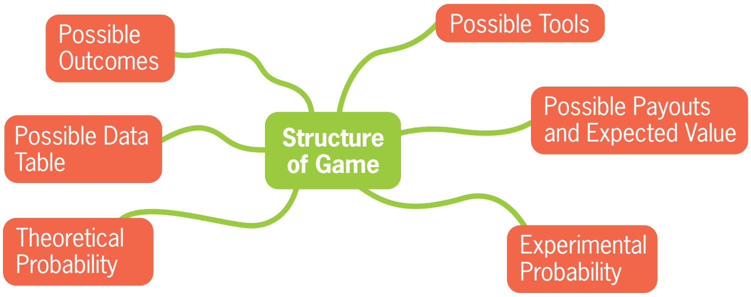
   c) If probability cannot be calculated based on the information given, what additional information is necessary?
   d) Could the game be changed to simplify the probability calculations?

6. **Reflect** What game ideas could you incorporate into your game of chance project? What tools, such as dice, could you use to generate random numbers?

7. **Extend Your Understanding** How could you change the game you analysed to increase the players' chance of winning? How about to decrease the players' chance of winning?

Step 1: Prepping for Game Construction

1. Create a mind map to brainstorm ideas about a game of chance you could create. Consider the games you investigated above, as well as any games discussed in class. Include details about how your data, calculations, and analysis might look.



2. Create a list of materials needed to construct your game.

3. Brainstorm the instructions for your game and how you will share them with players.

4. Decide on an appropriate payout structure for the winners.

5. What kind of data will your game generate? Design a chart you can use to record your data.

Round of Play	Desired Outcome	Actual Outcome	Number of Counters Won	Your Running Balance
1				
2				
3				

6. Simulate your game and record the data to test your instructions and data collection chart. Edit your instructions and chart if necessary.

7. Work with a partner to determine what probability calculations you can perform on the data generated. If you cannot calculate the theoretical probability of winning your game, you may need to change its structure. Does your game need to be
   - simplified? more complex?
   - easier to play? easier to track?
   - more fair to the players? more fair to you?

8. If your class is having a game day or fair to collect data, determine
   - how you will attract people to your game and how you will manage the data collection process
   - how you will educate people about games of chance in the real world and their probability of winning and fairness compared to your game

Project Prep

Refer to the analysis of a game of chance in section 9.2 for more ideas.

Step 2: Assessing Your Plan

Use the checklist to determine whether your game of chance plan is on track.

The Game	Yes	No
Is your game based solely on chance?		
Is your game • unique • fun • easy to understand • quick to play?		
Have you collected or created all the materials for your game?		
Have you prepared a chart to record data while playing your game?		
Have you calculated the theoretical probability of winning your game and planned an appropriate payout for the winners?		
If your class is having a game day or fair, have you developed a marketing strategy to attract players to your game?		
Are the rules clearly outlined and available to players? If your class is having a game day or fair, consider sharing the rules of the game as a poster, a short explanation, or a soft copy via LCD.		
If you are working with a partner, do you each have roles for game day?		
If your class is not having a game day or fair, have you determined how you will collect data for your game?		

Step 3: Collecting Your Data

1. Collect data for your game, either on a game day or by playing or simulating your game multiple times.

2. Be sure to keep track of all outcomes and payouts.

9.2

Analysing a Culminating Probability Project

Learning Goals

I am learning to

- critique a game of chance and the related analysis

Investigate Analysing a Culminating Probability Project

Kaelyn and Emma presented the following information about Triple Your Chances, Double Your Counters in a culminating project report. See section 9.1 for the game instructions.

Read Kaelyn and Emma's report. Then, complete the questions to critique and analyse it. Consider how you might analyse your game of chance and what elements you would include in your report. Record your ideas.

Triple Your Chances, Double Your Counters: Student Analysis of the Game

Theoretical Distribution of the Sum of Three Dice

There are 216 total possible sums when three dice are rolled at once.

	1						2						3						4						5						6					
	1	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
1	3	4	5	6	7	8	4	5	6	7	8	9	5	6	7	8	9	10	6	7	8	9	10	11	7	8	9	10	11	12	8	9	10	11	12	13
2	4	5	6	7	8	9	5	6	7	8	9	10	6	7	8	9	10	11	7	8	9	10	11	12	8	9	10	11	12	13	9	10	11	12	13	14
3	5	6	7	8	9	10	6	7	8	9	10	11	7	8	9	10	11	12	8	9	10	11	12	13	9	10	11	12	13	14	10	11	12	13	14	15
4	6	7	8	9	10	11	7	8	9	10	11	12	8	9	10	11	12	13	9	10	11	12	13	14	10	11	12	13	14	15	11	12	13	14	15	16
5	7	8	9	10	11	12	8	9	10	11	12	13	9	10	11	12	13	14	10	11	12	13	14	15	11	12	13	14	15	16	12	13	14	15	16	17
6	8	9	10	11	12	13	9	10	11	12	13	14	10	11	12	13	14	15	11	12	13	14	15	16	12	13	14	15	16	17	13	14	15	16	17	18

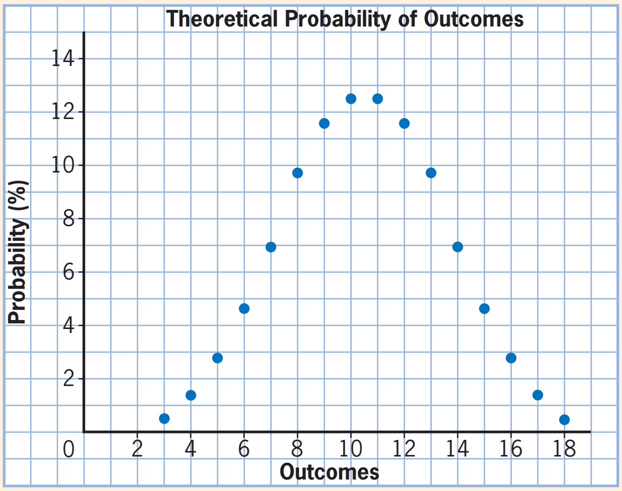
Possible Sums (x)	Probability $P(x)$	Expected Sum $x \cdot P(x)$	Counter Payout to Player (X)	Probability $P(X)$	Expected Payout $X \cdot P(X)$
3	$\frac{1}{216}$	$\frac{3}{216}$	9	$\frac{1}{216}$	$\frac{9}{216}$
4	$\frac{3}{216}$	$\frac{12}{216}$	9	$\frac{3}{216}$	$\frac{27}{216}$
5	$\frac{6}{216}$	$\frac{30}{216}$	3	$\frac{6}{216}$	$\frac{18}{216}$
6	$\frac{10}{216}$	$\frac{60}{216}$	3	$\frac{10}{216}$	$\frac{30}{216}$
7	$\frac{15}{216}$	$\frac{105}{216}$	2	$\frac{15}{216}$	$\frac{30}{216}$
8	$\frac{21}{216}$	$\frac{168}{216}$	2	$\frac{21}{216}$	$\frac{42}{216}$
9	$\frac{25}{216}$	$\frac{225}{216}$	1	$\frac{25}{216}$	$\frac{25}{216}$
10	$\frac{27}{216}$	$\frac{270}{216}$	1	$\frac{27}{216}$	$\frac{27}{216}$
11	$\frac{27}{216}$	$\frac{297}{216}$	1	$\frac{27}{216}$	$\frac{27}{216}$
12	$\frac{25}{216}$	$\frac{300}{216}$	1	$\frac{25}{216}$	$\frac{25}{216}$
13	$\frac{21}{216}$	$\frac{273}{216}$	2	$\frac{21}{216}$	$\frac{42}{216}$
14	$\frac{15}{216}$	$\frac{210}{216}$	2	$\frac{15}{216}$	$\frac{30}{216}$
15	$\frac{10}{216}$	$\frac{150}{216}$	3	$\frac{10}{216}$	$\frac{30}{216}$
16	$\frac{6}{216}$	$\frac{96}{216}$	3	$\frac{6}{216}$	$\frac{18}{216}$
17	$\frac{3}{216}$	$\frac{51}{216}$	9	$\frac{3}{216}$	$\frac{27}{216}$
18	$\frac{1}{216}$	$\frac{18}{216}$	9	$\frac{1}{216}$	$\frac{9}{216}$
		$\sum x \cdot P(x) = \frac{2268}{216} = 10.5$			$\sum X \cdot P(X) = \frac{416}{216} \approx 1.9$

The expected sum of three dice in this game on any given roll is 10.5, which means we expect the average sum of three dice will be 10.5.

The expected payout shown in the table is 1.9. So, we should expect to pay about 1.9 counters to players for each round of play. Assume one counter is placed on each sum in any given round. This means we would collect 16 counters from players and keep an average of 14.1 or 88% of the counters, suggesting this game is not very fair to the players.

Processes

Connecting
How would this calculation change if players could place 5 counters on the board? 10 counters? between 1 and 10 counters?



Notice the pattern. Sums of 3 and 18 have the least chance of occurring in the game because there is only one way to make each sum with three dice. Sums of 10 and 11 can each be made 27 ways with three dice, so they have the highest probability of occurring.

We used the probability of each possible outcome to help us decide how many counters to pay out when a player landed on a sum. Sums that are highly likely pay out fewer counters, compared to sums that are less likely.

Experimental Data: Sample of 10 Die Rolls Out of 120 Rounds of Play

Outcomes Chosen by Player	Actual Sum of Dice Rolled	Player Loss	Player Win	Number of Counters Won by Player	Balance of Counters for Die Roller (started with 100)
18	12	X		–1	101
17	10	X		–1	102
5	5		X	+3	99
16	10	X		–1	100
14	14		X	+2	98
8	8		X	+2	96
13	9	X		–1	97
6	10	X		–1	98
14	12	X		–1	99
4	4		X	+9	90

Kaelyn and Emma collected 120 rounds of experimental data. The table shows a sample of only the first 10 rounds to highlight how they recorded their data.

Calculating Experimental Probability for Sums

To collect data for our game, we played the game for 120 rounds and recorded our data.

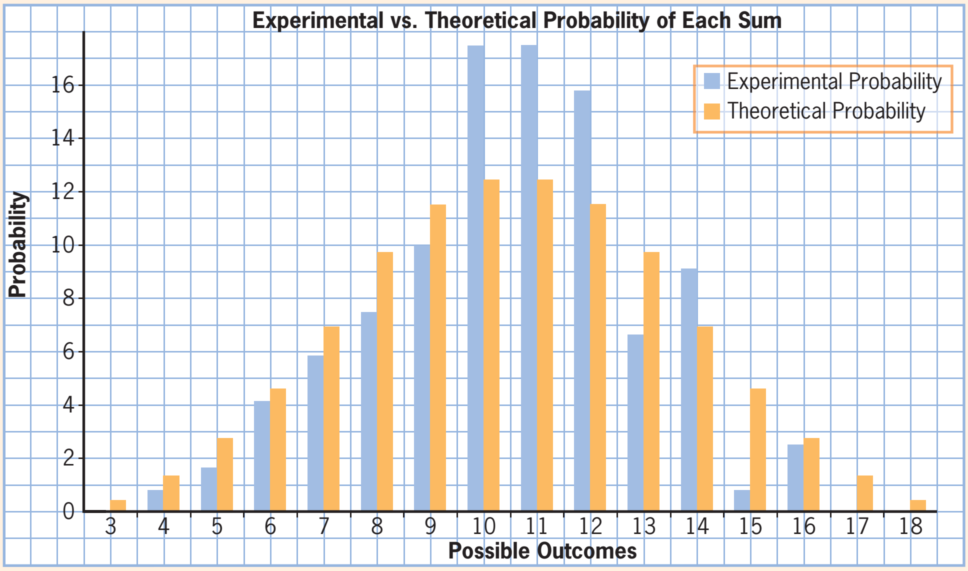
Sum	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Number of Times Sum Occurred	0	1	2	5	7	9	12	21	21	19	8	11	1	3	0	0

To calculate the experimental probability of each sum, we used the number of times a sum occurred out of the total number of rounds played.

$$P(9) = \frac{\text{\# of times occurred}}{\text{\# of rounds played}} \times 100$$
$$= \frac{12}{120} \times 100$$
$$= 10\%$$

$$P(12) = \frac{\text{\# of times occurred}}{\text{\# of rounds played}} \times 100$$
$$= \frac{19}{120} \times 100$$
$$\approx 16\%$$

The calculations shown are a sample of experimental probability calculations done for each of the possible sums.



Calculating Counter Payout Based on Experimental Probability From 120 Rounds

Sum	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Number of Times Sum Occurred	0	1	2	5	7	9	12	21	21	19	8	11	1	3	0	0
Number of Matches With Player	0	1	1	2	5	6	0	0	0	0	6	7	0	2	0	0
Payout Factor	9	9	3	3	2	2	1	1	1	1	2	2	3	3	9	9
Total Payout		9	3	6	10	12					12	14		6		

After playing 120 rounds, we had to pay out a total of 72 counters. This means we paid out 0.6 $\left(\frac{72}{120}\right)$ counters for every round played in our game.

Conclusions

The Experimental vs. Theoretical Probability of Each Sum graph shows which sums occurred more often than expected in our game. Sums of 9, 10, 11, and 12 occurred most often. Since these sums can be made in many ways with three dice, this is not a surprise. However, we would have expected sums 8, 13, and 15 to occur more often based on their theoretical probability.

In the experimental distribution, sums of 3, 17, and 18 show a probability of zero, which is also not a surprise because each of these sums also has the lowest theoretical probability. The theoretical probability suggests many of the outcomes should have occurred more often, even if only a little more often.

We would expect the theoretical and experimental probabilities to become more similar with more rounds of play.

In reality, the game was not very fair, as we had a significant advantage for winning counters. To make the game more fair we could alter the payout structure to provide greater payouts or allow players to put counters on more than one sum. For example, provide a grouping of {3, 4, 17, 18} with a 9:1 payout ratio. If a player placed a counter on the {3,4,17,18} group, the payout would be 9 counters if one of those sums were rolled.

1. What are the key elements of Kaelyn and Emma's analysis?

2. Could a tree diagram represent the possible outcomes of their game?

3. Does their analysis fully examine the probability of winning and any pros and cons of their game?

4. Does 120 rounds of play to calculate experimental probability seem large enough? What might happen if they played 500 or 1000 rounds?

5. a) Refer to page 15 for instructions on how to create a random number generator on a graphing calculator, or download an app. Simulate the outcomes of the game using 1000 and 5000 rounds of play.
   b) Compare the results to the experimental probability distribution for 120 plays. What do you notice?

6. Does the title of the game Triple Your Chances, Double Your Counters describe the probability of winning this game? Why might this title have been chosen?

7. Why do you think Kaelyn and Emma chose the payout structure they did? Do you agree with the counter payout offered to the player for each possible sum?

8. Do the charts and graphs provide a clear and accurate representation of their game and calculations?

9. What elements could be added or changed to more clearly explain the game, outcomes, and probability calculations?

Step 4: Planning Your Report

1. Are there any similarities between Kaelyn and Emma's game and your game of chance?
2. How might you represent the key elements of your analysis?
3. What graphs or charts could you use to represent data collected in your game?
4. What calculations could you include in your game of chance analysis? How will you show these calculations?
5. Would you make any adjustments to the payout structure in your game to ensure you win, or to make it more enticing for the player to continue playing?
6. What amount of data would you collect for your game of chance?
7. Could you use a random number generator to represent your data? If so, how many plays might you simulate? What results would you expect compared to the theoretical probability of winning your game, and to your experimental distribution if you simulated more rounds of play?
8. What would be an effective title for marketing your game of chance?
9. What elements of Kaelyn and Emma's game and analysis would you keep, change, or add to make a better game and report for your culminating probability project?

Step 5: Writing Your Report

Project Prep

Refer to section 9.5 for tools to self-assess your probability report.

A well-written report is organized, logical, and includes the necessary information with appropriate headings. Pull all the pieces together to tell the story of your game and its analysis in the form of a report. Your report should include

- a description of the game, rules, and required materials
- data from the game day or simulation
- calculations of the theoretical distribution, experimental distribution, and expected value
- comparison of the theoretical and experimental distributions
- summary of conclusions, including discrepancies in the results and insights and reflections about the game

Step 6: Finalizing Your Project

Look back at the theoretical probability for the game Triple Your Chances, Double Your Counters. Were the results very likely, given the number of trials Kaelyn and Emma ran?

While the game was being played, Logan was overhead saying, "This game is so easy! I could win five times in just five rounds! All I have to do is put my counters on 10!" Kaelyn and Emma did not agree and did the following calculations to disprove the statement.

Since a sum of 10 has the highest probability of occurring, we assumed that Logan will place his counter on 10 for all five plays. We can use a binomial distribution to represent the number of wins that occur out of five games because:

- Each roll of the dice is an independent event.
- There is a fixed number of dice being rolled.
- There are two possible outcomes: success or failure (either the sum on the dice is 10 or it is not).
- The chances of success remain constant, assuming Logan places his counter on 10 for all five rounds.

The table shows the chances of winning on each of the five rounds.

Number of Wins Out of Five Rounds	Probability of Winning Out of Five Rounds, $P(x) = {}_nC_x p^x q^{(n-x)}$
0	$P(x) = {}_5C_0(0.125)^0(0.875)^{(5-0)}$ $= 0.512\ 91$
1	0.366 36
2	0.104 68
3	0.014 95
4	0.001 068
5	0.000 030 5

Theoretical probability of rolling a sum of $10 = \frac{27}{216}$ $= 0.125$.

Our calculations show that Logan has a 51.3% chance of winning 0 out of 5 rounds played, assuming the same outcome is chosen for each of the five rounds.

Also, Logan's chance of winning 1, 2, 3, 4, or all 5 times out of five rounds gets progressively lower, suggesting that our game is actually not very easy to win.

1. Was a binomial distribution an appropriate choice to show the chances of winning from 0–5 times in five rounds?
2. Under what conditions does a binomial distribution model not work for this game?
3. Would another distribution model fit the data from this game better? Explain how you know.
4. What kind of distribution model might best fit the data from your game of chance? Include this information in your report and justify your calculations.

Step 7: Assessing Your Probability Project

Use the checklist to ensure your probability project is on track.

The Report	Yes	No
Does your report include: • Title page • Appropriate use of headings • Description of the game and relevant background information • Outline of the rules, including how to win • Explanation of the fee structure (cost to play) • Data table representing actual data from the game day • Actual winnings earned in the game based on your data table		
Have you calculated and/or graphed: • Theoretical distribution (possible outcomes vs. probability) • Experimental distribution (actual outcomes vs. probability) • Expected value of the game (profit based on your theoretical calculations)		
Does your analysis of the results: • Compare and contrast the theoretical and actual distributions • Discuss any discrepancies in the results • Provide insights and reflections about the game • Summarize your conclusions		
Review the following elements of your report: • Terminology: Are mathematical terms used correctly? • Neatness: Is your font clear and readable? Does your report look professional? • Writing Skills: Have you followed proper writing conventions? • Organization: Are key ideas presented logically, clearly, and concisely? Is your report focused and on topic? • Creativity: Have you included pictures or diagrams that make the report interesting? • Technology Skills: Are your visuals easily understood? Are your graphs clearly labelled?		

9.3

Planning and Carrying Out a Culminating Statistics Project

Learning Goals

I am learning to

- pose research questions and state hypotheses
- design a plan to investigate a research question
- gather, organize, and represent data

In 2009, legislation was passed in Ontario making it illegal to use hand-held devices while driving. What prompts the need for such a law? As cell phone use increases, police officers are reporting that more accidents are occurring while drivers are texting. Does texting while driving increase your chances of being in an accident? To answer this question, researchers are collecting data to look for trends.

Overview of the Research Process

Questions or ideas arise in nature, health, science, and various other fields.
Does texting while driving increase your chances of having an accident?
Does exercise reduce your risk of heart disease?

↓

Researchers explore the question further. They use the Internet, libraries, government publications, databases, and all other available resources.

↓

Based on the information found in their search, researchers revise the initial **research question** into a prediction statement called a **hypothesis**.
Texting while driving increases your risk of having an accident.
Regular exercise reduces your risk of heart disease.

↓

Researchers gather data to find out whether the prediction is accurate. They may use existing sources such as Statistics Canada or collect data via surveys, games, simulations, or experiments.

↓

Researchers use charts, graphs, pictures, or tables to organize the data. Then, they use statistical tools, calculations, or equations to see what story the data tell.

↓

Researchers summarize the story and share it with other researchers for critiquing. The validity of the research method and analysis is assessed.

↓

The results of research projects can affect policies and programs.

research question

- a question about a topic, problem, or area that can be investigated or solved

hypothesis

- a prediction about the relationship between variables or about the outcome of a research question
- used to guide an investigation or experiment

Step 1: Defining and Posing the Problem

Materials

- computer with Internet access

A good research question explores a relationship between two variables.

For example:
- What is the relationship between car insurance premiums and a driver's age?
- What is the relationship between education and income level?

1. Consider your interests and search for current issues or controversies in newspapers, magazines, or on the Internet.

2. Construct a graphic organizer showing your areas of interest, related issues, and possible research questions. You could use a concept map, a mind map, a graffiti sheet, or a chart.

3. Add a hypothesis to your graphic organizer for each research question in your list. Write your hypothesis as a prediction statement. The chart shows some examples. The dependent variable is underlined. The independent variable is italic. Identify the dependent and independent variable for each of your research questions.

Possible Research Questions	Hypothesis Written as Prediction Statement
What is the relationship between *homework completion* and <u>grades in math</u>?	Completing more homework questions will raise your math mark.
What is the relationship between the amount of *exercise* a person gets and their <u>resting heart rate</u>?	Regular cardiovascular exercise leads to a lower resting heart rate.
What is the relationship between *time spent playing video games* and <u>fitness level</u>?	As time spent playing video games increases, fitness level decreases.

4. **a)** Choose a question from the graphic organizer that you wish to use as your project question.

b) What kind of data would help you investigate your question?

c) Who is the population and sample for your research question?

Once you have chosen a research question and determined your hypothesis, you will need to collect some data to support your hypothesis. In some cases, data on your variables may already exist. If you cannot find existing data, you can alter your research question to match available data, or you might collect your own survey data.

Example 1

Searching for Data

Search for data on the following research question: What is the relationship between texting while driving and being in a collision?

Solution

An Internet search on texting, driving, and collisions suggests that texting while driving is becoming a problem and is a topic worthy of investigation.

"Instead of drinking, drivers are now distracting themselves to death."

Source: "How the War on Drunk Driving Distracts from the Real Danger," *Maclean's*, October 21, 2013.

"According to the Ontario Provincial Police, as of the beginning of September, 32 deaths in the province this year are attributable to impaired driving, while 47 were caused by distracted driving."

Source: "How the War on Drunk Driving Distracts from the Real Danger," *Maclean's*, October 21, 2013.

"Texting is the most deadly of these distractions. The Canadian Automobile Association claims texting drivers are 23 times more likely to be in an accident than a driver who's paying attention to the road."

Source: "How the War on Drunk Driving Distracts from the Real Danger," *Maclean's*, October 21, 2013.

"Experiments have also shown drivers who are actively texting have reaction times substantially slower than someone legally impaired by alcohol."

Source: "How the War on Drunk Driving Distracts from the Real Danger," *Maclean's*, October 21, 2013.

"Analysis of driving performance revealed that participants responded more slowly to the onset of braking lights while texting and driving. Moreover, text-messaging drivers were involved in more crashes than drivers not engaged in text messaging. Text messaging while driving has a negative impact on simulated driving performance. This negative impact appears to exceed the impact of conversing on a cell phone while driving."

Source: Drews, Frank A., et al., "Text Messaging During Simulated Driving," *Human Factors: The Journal of the Human Factors and Ergonomics*, December 16, 2009.

"Rigging a car with a red light to alert drivers when to brake, *Car and Driver* tested how long it takes to hit the brake when sober, when legally drunk at .08, when reading an e-mail, and when sending a text.

The results:
Unimpaired: .54 seconds to brake
Legally drunk: add 4 feet
Reading e-mail: add 36 feet
Sending a text: add 70 feet"

Source: LeBeau, Philip, "Texting and Driving Worse Than Drinking and Driving," Behind the Wheel, CNBC, June 2009.

After doing this initial research, you might hypothesize that texting while driving increases the chance that the driver will be in a collision. You could then collect data on how often people text while driving and on how often texting is listed as the cause of an accident. You might search websites such as Transport Canada, Statistics Canada, Traffic Injury Research Foundation (TIRF), or Canadian Automobile Association (CAA) to find this information.

Suppose that after extensive searching you cannot find the data you need to answer your research question, but instead find data on deaths and injuries from collisions and data on wireless subscribers for the same time period.

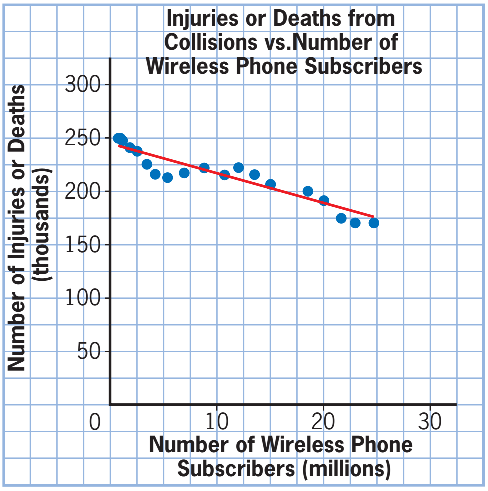
You could alter your research question to investigate the relationship between the number of cell phones in use and the number of accidents during the same time period. You might hypothesize that as cell phone purchases increase, accidents also increase because more cell phones likely means more opportunities for distracted driving and related accidents to occur.

Consider the table. You can use these data to calculate one-variable and two-variable statistics because they come from Canadians over the same time period.

Year	Total Deaths or Injuries From Collisions in Canada	Number of Wireless Phone Subscribers in Canada
1991	249 217	771 060
1992	249 823	1 023 810
1993	247 593	1 321 387
1994	241 899	1 868 882
1995	238 458	2 584 387
1996	227 283	3 414 711
1997	217 401	4 207 019
1998	213 319	5 317 247
1999	218 457	6 883 195
2000	222 869	8 731 220
2001	216 489	10 678 560
2002	222 707	11 934 565
2003	216 210	13 442 350
2004	206 229	14 984 396
2005	204 764	16 809 988
2006	199 970	18 425 194
2007	192 744	19 919 512
2008	176 455	21 513 862
2009	171 415	22 850 757
2010	170 629	24 567 947

Sources: Collision and Casualties 1991–2010, Transport Canada
Mobile Wireless Subscribers in Canada, Canadian Wireless Telecommunications Association

The graph shows injuries or deaths from collisions as a function of the number of wireless phone subscribers.

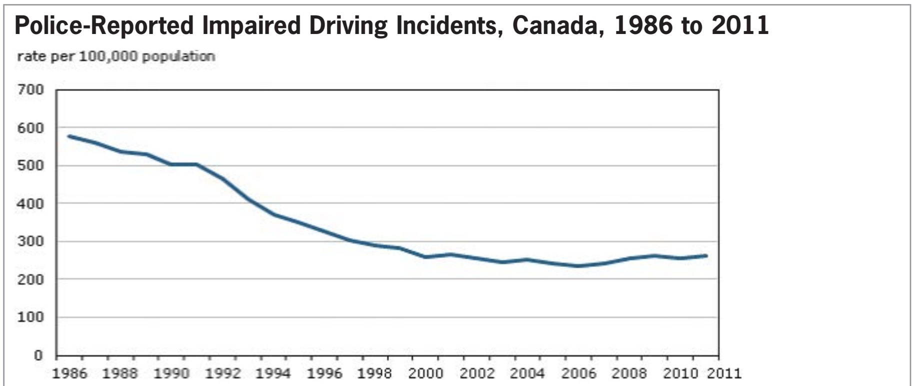


According to the graph, as the number of wireless phone subscribers increases, the accident rate appears to decrease.

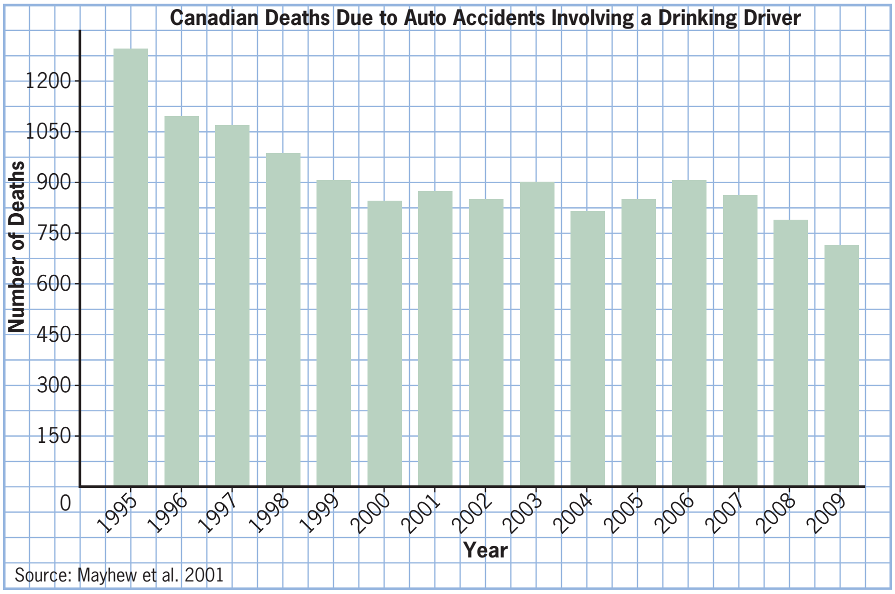
With a correlation coefficient of –0.94 and an r^2 value of 0.88, there appears to be a strong negative linear correlation. This result **contradicts** the hypothesis that as the number of cell phone subscribers increases, accidents will also increase.

When unexpected results arise, researchers must consider what other factors could account for the results.

Consider the following graphs. What trends do you notice?



Source: Chart 1, Police-reported impaired driving incidents, Canada, 1986 to 2011, Statistics Canada



Source: *The Road Safety Monitor* 2012 *Drinking and Driving in Canada: Trends*, Traffic Injury Research Foundation

The graphs show a decline in both impaired driving incidents and deaths resulting from a drinking driver. Education about the risks of impaired driving may have reduced the frequency of impaired driving, which could result in a fewer accidents.

Further, graduated licensing programs adopted starting in 1994 allow for more driving practice before full driving privileges are granted. Studies have shown that this contributes to the decrease in the overall accident rate.

Increased enforcement of speed limits and increased levels of fines and licence suspensions might be another factor contributing to the decreasing number of accidents.

Lastly, in recent years more drivers might be using Bluetooth technology leading to fewer accidents resulting from cell phone use.

To summarize, the negative correlation between the number of cell phones in use and accident rates could have occurred because accident rates are dropping for a variety of reasons.

Some important information is missing from the data that were collected:
- How many of the accidents were actually the result of cell phone use in the first place?
- Of these accidents, how many can be attributed to texting?
- Have either of these variables changed over time?

Without knowing the answers to these questions, the result of the regression analysis is not reliable. Even though the relationship has a correlation coefficient of –0.94 and an r^2 value of 0.88, **no conclusions** can be drawn about accident rates and cell phones in use, or about accidents resulting from texting while driving, because so many other variables could have been responsible for the decrease in accident rates.

Further investigation is required to support the hypothesis that increased cell phones in use leads to an increased accident rate.

If your research and analysis also lead to unexpected results, you don't need to start over. Instead, consider possible explanations for your results and discuss ideas for future study.

In terms of texting and driving research, teens participated in a plan to investigate distracted driving behaviours as follows:

"[H]igh school students in 10 Canadian cities [were asked] to conduct a national distracted-driving blitz. The students staked out major intersections and toted-up drivers doing things they shouldn't. Nearly 20 per cent of all drivers were spotted either talking or texting on their cell phones, something that's illegal in all provinces."

Source: "How the War on Drunk Driving Distracts from the Real Danger," *Maclean's*, October 21, 2013.

In the absence of ideal data, you may still be able to present a convincing argument if you apply your critical thinking skills.

Step 2: Finding and Collecting Secondary Data

1. Search the Internet for websites related to your question of interest.
2. Search the Statistics Canada website to determine whether data on your variables have already been collected.
3. Search newspapers, magazines, journals, or other literature for more background information or data. If you cannot find data, or if the amount of data is overwhelming, consider revising your research question.
4. Keep track of the websites you use and cite your online references properly.

APA Style

Format for referencing an article from a website:

Organization or Author's Last Name, Author's First Initial. (Year, Month Day Published). *Website Title*. Retrieved Month Day, Year, from URL.

Example:

Canadian Automobile Association. (2014). *CAA Bike Safety*. Retrieved Feb. 21, 2014 from URL.

Format for referencing a data set from a website:

Organization or Author's Last Name, First Initial. (Year). Table name [Data file], Retrieved from URL.

Example:

Statistics Canada. (2014). Labour Force Survey January 2014 [Data file]. Retrieved from URL.

MLA Style

Format for referencing an article from a website:

Editor or Compiler or Author's Last Name, Author's First Name (if available). *Website Title.* Name of institution/organization affiliated with the site (sponsor or publisher), date of creation (if available). Web. Date of access.

- If the publisher name or date is not available, include *n.p.* or *n.d.* to indicate.

Example:

Canadian Automobile Association. *CAA Bike Safety*. Canadian Automobile Association, 2014. Web. Feb. 21, 2014.

Format for referencing a data set from a website:

Editor or Compiler or Author's Last Name, Author's First Name (if available). "Title of page or table." *Website Title.* Name of institution/organization affiliated with the site (sponsor or publisher), date of creation (if available). Web. Date of access.

- If the publisher name or date is not available, include *n.p.* or *n.d.* to indicate.

Example:

Statistics Canada. "Labour Force Survey January 2014." Statistics Canada. (2014). Web. 21 Feb. 2014.

5. a) Are the data you found valid and related to your research question?

b) What is the source of the data? Is this a reliable source for data? How do you know?

c) How were the data collected?

d) What was the sample for this data set? Will it allow you to make predictions about the larger population?

e) Is there any evidence of bias in the data collection method?

f) Can the data be organized to facilitate retrieval and analysis?

g) How much and what kind of data do you need for your project?

h) Do you need to revise your research question?

Example 2

Surveying for Data

Secondary data are not always available, so researchers may need to collect their own data. Surveys are a common tool for collecting primary data. Cassie wants to investigate the relationship between time spent using technology (cell phones, texting, social media sites, video games, and so on) and students' academic performance.

a) Create a paper survey to help Cassie determine which questions to use. Consider what kind of data the survey questions could generate.

b) If Cassie wants to use an electronic survey, what criteria should she first consider?

Solution

a) Read each question and consider the response the question would generate.

Cassie's Survey

For #1 to #6, circle one answer per question.

1. What is your gender?	Male	Female			
2. What grade are you in?	9	10	11	12	12+
3. Do you own a cell phone?	Yes	No			
4. Do you use text messaging?	Yes	No			
5. Do you use social media websites?	Yes	No			
6. Do you play video games?	Yes	No			

Each question generates categorical data that would allow Cassie to compare results across a category. For example, she might compare the technology use of males and females, or of students by grade.

For #7 and #8, circle one answer per question.

7. How much time did you spend on your cell phone yesterday?

0-1 hrs	1-2 hrs	2-3 hrs	3-4 hrs	4-5 hrs	5-6 hrs	6-7 hrs	7-8 hrs	8-9 hrs	9-10 hrs	10-11 hrs	11-12 hrs

8. How many texts did you send yesterday?

0-20	20-40	60-80	80-100	100-120	120-140	140-160

For each question, the data generated will fall into one of the given ranges. Providing ranges on a survey can help respondents consider a wider variety of responses. However, the midpoint of each range would have to be used to calculate measures of central tendency, which could diminish the level of accuracy of the calculations and results.

For #9 and #10, think about a typical week and answer as accurately as possible.

9. How many hours did you spend on social media websites yesterday? ________

Why is this better than asking how many hours you spend on social media websites per day?

10. How many hours did you spend playing video games yesterday? ________

11. What is your current math mark? ________

12. What was your overall average mark on your last report card? ________

These questions allow respondents to enter their own data. The questions are open-ended with no restrictions, which can generate more accurate results as long as respondents answer truthfully.

Processes

Connecting
What data management tools can you use to investigate numerical data such as the answers for #9 to #12?

b) In chapter 5, you created a survey using Fathom™ and Google Docs. You can also create a survey using online survey tools. Before committing to an electronic survey, you need to be sure the technology fits your purpose.

Consider the following criteria to help you determine which type of electronic survey best meets your needs. Cassie's responses are shown.

Criterion	Cassie's Response
Is there a fee to use the service?	Fees may make it harder for me to get participants.
Is there a limit to how many survey questions can be asked?	10 to 20 questions should be enough.
Does the tool allow for open and closed questions?	Questions can be framed either way to get the data I need.
How is the survey sent out? Examples: email, telephone, posted for the public	It would be easiest to give the survey website to participants
Can anyone respond?	I will limit my sample to participants from school.
How is anonymity protected?	Accuracy may not be compromised since data I am collecting are not personal or sensitive.
Is there a time limit or expiration on the survey site?	As long as I get participants right away, this will not matter.
How are the survey data organized?	I can change the format if I need to.
When are the data available?	Right away is best.
How easy is the survey to use, for both the surveyor and respondents?	If it is too hard to complete, no one will participate.

Before doing a paper or electronic survey, test your survey on a small sample to check whether your questions make sense.

Step 3: Gathering Primary Data

If you cannot find secondary data to help answer your research question, you may need to create your own survey to collect primary data. Refer back to Cassie's survey in Example 2 on page 463. What kinds of questions did Cassie ask?

1. Brainstorm a list of questions for your survey.
2. How will you structure your questions to maximize clarity and minimize bias? Decide the best way to ask each of your survey questions. For example, open response, multiple choice, a scale, or a scale with intervals.
3. Construct your survey. Remember to protect the anonymity of survey respondents by avoiding questions that reveal their identity.
4. Do the survey to see how long it takes to answer. Check for errors.
5. Give a few copies of your survey to a small sample to ensure your survey questions make sense and provide appropriate, valid responses.
6. Edit your survey as required.
7. Determine what amount of data is required for your project. Decide who you will survey, and how you will distribute your survey to ensure valid results.

8. If your survey questions will allow you to graph your data, calculate one and two variable statistics, and tell a meaningful story, then it is time to collect your data.

Step 4: Graphing Your Data

Once you have collected data, it is helpful to see the data summarized in a graphical format, such as a table, chart, or graph.

1. What kind of graph would best represent your data?
2. Graph your data using more than one type of graph. This will allow you to determine which visual best represents your data. Be sure to consider the most appropriate scales and labels for your graph and variables.
3. What are the advantages or disadvantages of each type of graph?
4. Examine your graphs. Do you see any patterns?
5. Are there any outliers? How might these outliers affect your one-variable and two-variable statistics calculations?

Step 5: Assessing Your Project So Far

Use this checklist to ensure your statistics project is on track.

Defining the Problem	Yes	No
The chosen topic is relevant and appropriate.		
The hypothesis is clearly stated and justified.		
The dependent and independent variables are clearly defined.		
The population of interest is clearly identified.		
Background information is used to introduce the research question.		

Secondary Data Collection	Yes	No
The source of data (URL) is provided.		
The sampling technique is described and justified.		
Sources of sampling bias are identified.		
Sufficient data are collected.		
The raw data are effectively represented.		

Primary Data Collection	Yes	No
The sample survey is included in the report.		
The survey contains an appropriate introduction.		
The survey questions are well designed.		
The data are collected in a timely and responsible manner.		
Survey participants' anonymity is clearly protected.		
Sufficient survey data are collected.		
Raw survey data are effectively represented.		

Data Analysis	Yes	No
The data are represented using appropriate charts and graphs.		

9.4

Analysing a Culminating Statistics Project

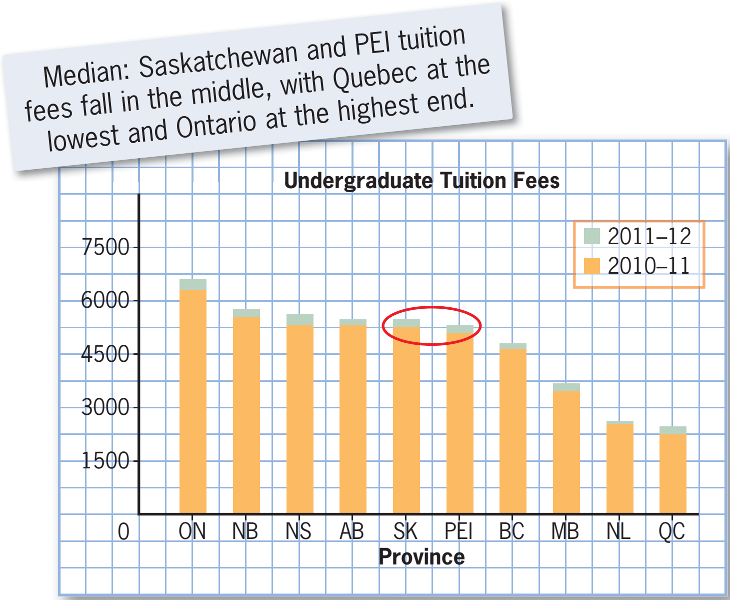
Learning Goals

I am learning to

- represent and interpret one- and two-variable statistics from a data set
- evaluate the impact of outliers on my calculations and justify my results
- organize, summarize, and interpret the key results of my data analysis

Mean, median, and mode are three common statistical calculations used to provide a snapshot or quick summary of a data set. The range, interquartile range, standard deviation, and variance also help describe the general distribution of a data set.

- How will you share your measures of central tendency and dispersion in your statistics project?
- What might be an effective and interesting way to summarize these calculations?



Source: CNW Group/Canadian Federation of Students

Step 6: Digging In to the Centre of Your Data

Telling the story of data begins with assessing the distribution of the data.

1. Consider the data you collected. Would it be useful to look at the mean, median, mode, standard deviation, or variance? Perform the relevant calculations and record your results in a table like the one below.

	Mean	Median	Mode	Standard Deviation	Variance
Independent Variable					
Dependent Variable					

2. What do these calculations tell you about your data?

3. Are there any data values for which you might wish to calculate a z-score? If so, calculate it now.

4. Do your data values appear to follow the pattern of a normal distribution? If not, what does this mean?

5. Test for outliers using a modified box and whisker plot or the 1.5 IQR test described in chapter 6. If your data contain outliers, describe their impact on the mean, median, mode, range, interquartile range, standard deviation, and variance.

Step 7: Searching for Patterns

Now it is time to determine whether your data follow a specific pattern or model, or whether a specific mathematical relationship exists among your variables.

1. Enter your data into appropriate lists using the technology of your choice.

2. Perform a regression analysis to determine whether a linear correlation represents your data. Record and interpret the r and r^2 values.

3. You may also wish to perform an additional regression analysis to see whether a quadratic, cubic, or exponential model fits your data better than a linear model. Record and interpret the r^2 values. Remember that the r or r^2 values on their own do not indicate which model is best, just if the model is a good fit. In addition, it does not imply cause and effect.

4. Represent your regression analysis on a scatter plot to show the line (or curve) of best fit.
5. If you had any outliers in your data set, describe their impact on the line (or curve) of best fit.
6. Show how the line (or curve) of best fit changes if you remove any outliers. Decide whether you should remove outliers from the data set. Justify your choice.
7. What might be an interesting way to share or represent these results in your project?
8. Are there any other statistical calculations you can include to further describe the story of your data?

Step 8: Assessing Your Data Analysis

Revisit the work you have done so far. Use the checklist to ensure you are on track.

Data Analysis	Yes	No
The data are represented using appropriate charts and graphs.		
The measures of central tendency are calculated and interpreted clearly.		
The measures of spread or dispersion are calculated and interpreted clearly.		
The appropriate tests have been done for outliers.		
Relevant *z*-score calculations are included and explained.		
The impact of any outliers on the results is thoroughly analysed.		
For the regression analyses: • all relevant models of best fit are explored, discussed, and reported. • the most appropriate model is chosen and justified. • the r and r^2 values are reported and explained. • a scatter plot shows a line (or curve) of best fit and the impact of potential outliers.		

Step 9: Assessing the Validity of Your Process and Results

Now that you have completed your statistical analysis, it is time to assess your research method and the validity of your results and sources.

1. Examine each of your statistical calculations. Do your results and conclusions support or refute your hypothesis?
2. How strong is the evidence?
3. Are there limitations to the evidence? Discuss bias, cause and effect, hidden variables, and how well your sample represents the population.
4. What questions arose that might require further investigation?
5. What might you do differently if you were to complete a similar project in the future?

Step 10: Assessing Your Results

Use this checklist to ensure your statistics project is on track.

Evaluating Your Results	Yes	No
Sources of bias (beyond sampling bias) are examined and discussed.		
Conclusions are reported and justified.		
Reflections on the methodology and improvements for future studies are thorough and insightful.		
References are appropriately sourced.		

A well-written report is organized, logical, and includes the necessary information with appropriate headings. Pull together all the pieces to tell the story of your data in the form of a report and/or presentation.

Project Prep

Refer to section 9.5 for tools to self-assess your culminating project report and presentation.

9.5

Creating a Report, Presenting, and Critiquing Other Projects

Learning Goals

I am learning to

- critique my own research report or presentation constructively
- critique the research reports or presentations of my peers based on predetermined criteria

Think of a time when you received constructive feedback. What was helpful? Did anything make you feel uncomfortable?

Assessing Your Culminating Probability Project

Materials

- criteria for effective feedback generated by class
- Probability Project Self-Assessment Checklist
- Probability Project Content Rubric
- Probability Project Presentation Rubric
- Probability Project Peer Critique Form
- Statistics Project Self-Assessment Checklist
- Statistics Project Content Rubric
- Statistics Project Presentation Rubric
- Statistics Project Peer Critique Form
- video recorder (optional)

1. Use the Probability Project Self-Assessment Checklist to ensure you have included all relevant elements in your project.

2. Use the Probability Project Content Rubric to assess the depth and quality of your report.

3. If you will be presenting your project to the class, consider recording your presentation on video and use the Probability Project Presentation Rubric to assess the quality of your presentation.

4. With a partner, brainstorm criteria for effective constructive feedback. Write a list of at least five tips to follow when providing feedback to a peer.

5. Trade projects with another peer or team. Use the Probability Project Self-Assessment Checklist and Probability Project Content Rubric to critique their work.

6. Use the Probability Project Peer Critique Form to provide written feedback. Keep in mind the criteria you brainstormed in #4 for effective feedback.

Reflect

a) Are there any aspects of your peers' projects you could incorporate into your project?

b) What feedback from your peers can you incorporate into your project/report?

c) What changes or improvements will you make before submitting your final project?

Assessing Your Culminating Statistics Project

1. Use the Statistics Project Self-Assessment Checklist to ensure you have included all relevant elements in your project.
2. Use the Statistics Project Content Rubric to assess the depth and quality of your report.
3. If you will be presenting your project to the class, consider recording your presentation on video and use the Statistics Project Presentation Rubric to assess the quality of your presentation.
4. With a partner, brainstorm criteria for effective constructive feedback. Write a list of at least five tips to follow when providing feedback to a peer.
5. Trade projects with another peer or team. Use the Statistics Project Self-Assessment Checklist and Statistics Project Content Rubric to critique their work.
6. Use the Statistics Project Peer Critique Form to provide written feedback. Keep in mind the criteria you brainstormed in #4 for effective feedback.

Reflect

a) Are there any aspects of your peers' projects you could incorporate into your project?

b) What feedback from your peers can you incorporate into your project/report?

c) What changes or improvements will you make before submitting your final project?

Probability Project Self-Assessment Checklist

The Game	Yes	No
Is your game based solely on chance?		
Is your game • unique • fun • easy to understand • quick to play?		
Have you collected or created all the necessary materials for your game?		
Have you prepared a table to record your data during the playing of your game?		
Have you calculated the theoretical probability of winning your game and planned an appropriate payout for the winners?		
If your class is having a game day or fair, have you developed a marketing strategy to attract players to your game?		
Are the rules clearly outlined and available to players? If your class is having a game day or fair, consider sharing the rules of the game as a poster, succinct explanation, or soft copy via LCD.		
If you are working with a partner, do you each have roles for game day?		
If your class is not having a game day or fair, have you determined how you will collect data for your game?		
The Report	**Yes**	**No**
Does your report include: • Title page • Appropriate use of headings • Description of the game and relevant background information • Outline of the rules, including how to win • Explanation of the fee structure (cost to play) • Data table representing actual data from the game day • Actual winnings earned in the game based on your data table		
Have you calculated and/or graphed: • Theoretical distribution (possible outcomes vs. probability) • Experimental distribution (actual outcomes vs. probability) • Expected value of the game (profit based on your theoretical calculations)		
Does your analysis of the results: • Compare and contrast theoretical and actual distributions • Discuss any discrepancies in the results • Provide insights and reflections about the game • Summarize your conclusions		
Review the following elements of your report: • Terminology: Are mathematical terms used correctly? • Neatness: Is your font clear and readable? Does your report look professional? • Writing Skills: Have you followed proper writing conventions? • Organization: Are key ideas presented logically, clearly, and concisely? Is your report focused and on topic? • Creativity: Have you included pictures or diagrams that make the report interesting? • Technology Skills: Are your visuals easily understood? Are your graphs clearly labelled?		

Probability Project Content Rubric

	Level 1 (50–59%)	Level 2 (60–69%)	Level 3 (70–79%)	Level 4 (80–100%)
Knowledge and Understanding				
Introduction • Description of the game and rules, relevant background information	• Limited information about the game	• Sufficient information about the game	• Detailed description of most elements of the game	• Clear and detailed description of all elements of the game
Data Collection • Data table represents actual data from the game day or simulation	• Insufficient data collected	• Sufficient data collected	• Sufficient data collected; data are appropriate to the task	• Extensive data collected; data are appropriate to the task and clearly represented
Mathematical Understanding • probability calculations and data analysis	• Reflects limited mathematical understanding	• Reflects some mathematical understanding	• Generally reflects mathematical understanding	• Reflects thorough mathematical understanding
Application				
Data Analysis • theoretical distribution • experimental distribution • expected value	• Limited use of appropriate mathematical procedures and strategies	• Mathematical strategies and procedures generally appropriate and sometimes correct	• Appropriate strategies used most of the time • Mathematical procedures generally correct/appropriate	• Mathematical strategies and procedures always correct/appropriate and are highly detailed
Thinking				
Summary and Conclusions • compare/contrast theoretical and actual distributions • discuss discrepancies in the results • provide insights and reflections about the game and clear conclusions	• Performs limited comparisons • Uses mathematical reasoning to justify few conclusions	• Performs some comparisons • Uses mathematical reasoning to justify some conclusions	• Performs detailed comparisons • Uses mathematical reasoning to justify all conclusions	• Performs and mathematically justifies conclusions based on insightful comparisons and reflections
Communication				
Clarity • explanations clear, accurate, and thorough	• Explanations are incomplete, inaccurate, and/or lack clarity	• Explanations are somewhat clear and accurate	• Explanations are mostly clear and accurate	• Explanations are clear, accurate, and thorough
Math Terminology and Writing • mathematical terms used correctly • proper writing conventions followed	• Limited use of correct terminology and notation • Proper writing conventions followed to a limited extent	• Some use of correct terminology and notation • Proper writing conventions followed sometimes	• Considerable use of correct terminology and notation • Proper writing conventions mostly followed	• Thorough and meticulous use of correct terminology and notation • Proper writing conventions extensively followed
Organization • key ideas presented logically • appropriate headings (title page, table of contents, etc.)	• Report not clear or logical; ideas rarely connect	• Report somewhat clear or logical; ideas connect sometimes	• Report most often clear, logical; ideas mostly connect	• Report clear, logical; ideas thoroughly connect

Probability Project Presentation Rubric

	Level 1 (50–59%)	Level 2 (60–69%)	Level 3 (70–79%)	Level 4 (80–100%)
Introduction				
Game Description • rules, objectives, and key information about the game	• Provides an unclear description of the game	• Provides a somewhat focused description that states the objective and some of the key points about the game	• Provides a focused description that clearly states the objective and the key points of the game	• Provides an effective description that captures the attention of the audience and clearly states the objective and the key points of the game
Visual Aids				
Visuals (graphs and slides) • clear/readable • neatly presented • easy to understand • graphs clearly labelled	• Visual aids do not support the conclusions or are almost non-existent	• Visual aids sometimes support the conclusions	• Visual aids generally support the conclusions	• Visual aids used consistently and thoroughly support the conclusions
Presentation Skills				
Communication • voice is clear, relaxed, audible • appropriate eye contact/body language • appropriate pace • clear, concise explanations • engaging and enthusiastic	• Presenter's voice rarely clear/audible • Presenter frequently reads from the screen or notes • Presentation is slow and choppy	• Presenter's voice somewhat clear • Presenter sometimes reads from the screen or notes • Presentation flows but moves slowly	• Presenter's voice usually clear • Presenter occasionally reads from the screen/notes • Presentation flows well; pace keeps the audience interested	• Presenter's voice always clear/audible • Presenter maintains eye contact with the audience • Presentation is highly engaging
Organization • key ideas presented logically • focused and on topic	• Not very clear or logical; ideas rarely connect • Few aspects are organized	• Somewhat clear or logical; ideas connect sometimes • Some aspects are organized	• Often clear, logical; ideas mostly connect • Many aspects are well-organized	• Clear, logical; ideas thoroughly connect • Presentation is well-organized
Ability to Engage the Audience				
Response to Audience Questions	• Responds to few audience questions, with limited justification or mathematical reasoning	• Responds to some audience questions, with somewhat detailed justification or mathematical reasoning	• Responds to most audience questions, with detailed justification or mathematical reasoning	• Responds to all audience questions, with very detailed justification or mathematical reasoning

Probability Project Peer Critique Form

Rate the quality of the presentation or report in each category using the following scale:

1 = poor 2 = fair 3 = good 4 = very good 5 = outstanding

	1	2	3	4	5
Description of game • outline of the rules, including how to win • explanation of the probability of the player winning					
Data table representing actual data from the game day or simulation					
Data analysis: • theoretical distribution (possible outcomes vs. probability) • experimental distribution (actual outcomes vs. probability) • expected value of the game (profit based on your theoretical calculations)					
Visuals (graphs and slides)					
Organization					
Communication					
Technology					
Conclusions					

Highlights:	Things to improve or add:

Statistics Project Self-Assessment Checklist

Defining the Problem	Yes	No
The chosen topic is relevant and appropriate.		
The hypothesis is clearly stated and justified.		
The dependent and independent variables are clearly defined.		
The population of interest is clearly identified.		
Background information is used to introduce the research question.		
Secondary Data Collection	**Yes**	**No**
The source of data (URL) is provided.		
The sampling technique is described and justified.		
Sources of sampling bias are identified.		
Sufficient data are collected.		
The raw data are effectively represented.		
Primary Data Collection	**Yes**	**No**
The sample survey is included in the report.		
The survey contains an appropriate introduction.		
The survey questions are well designed.		
The data are collected in a timely and responsible manner.		
Survey participants' anonymity is clearly protected.		
Sufficient survey data are collected.		
Raw survey data are effectively represented.		
Data Analysis	**Yes**	**No**
The data are represented using appropriate charts and graphs.		
The measures of central tendency are calculated and interpreted clearly.		
The measures of dispersion are calculated and interpreted clearly.		
The appropriate tests have been done for outliers.		
Relevant *z*-score calculations are included and explained.		
The impact of any outliers on the results is thoroughly analysed.		
For the regression analyses: • all relevant models of best fit are explored, discussed, and reported. • the most appropriate model is chosen and justified. • the r and r^2 values are reported and explained. • a scatter plot shows line (or curve) of best fit and the impact of potential outliers.		
Evaluating Your Results	**Yes**	**No**
Sources of bias (beyond sampling bias) are examined and discussed.		
Conclusions are reported and justified.		
Reflections on the methodology and improvements for future studies are thorough and insightful.		
References are appropriately sourced.		

Statistics Project Content Rubric

	Level 1 (50–59%)	Level 2 (60–69%)	Level 3 (70–79%)	Level 4 (80–100%)
Knowledge and Understanding				
Topic Question and Introduction	• Hypothesis stated, but not clear or justified • Background knowledge stated to a limited extent	• Hypothesis makes sense but could be more clear • Background knowledge somewhat evident	• Clear hypothesis • Background knowledge evident	• Clear, justified hypothesis • Background knowledge clearly evident and extensive
Data Collection and Sampling Method • understands methods for collecting and validating data	• Insufficient data collected • Data limited to one element of the course • Inadequate validation of the data source	• Some data collected • Data limited to few aspects of the course • Some validation of the data source	• Sufficient data collected • Data mostly appropriate to the task • Adequate validation of the data source	• Extensive data collected • Data appropriate to the task • In-depth validation of the data source
Mathematical Understanding • calculations and data analysis	• Reflects limited mathematical understanding	• Sometimes reflects mathematical understanding	• Generally reflects mathematical understanding	• Reflects thorough mathematical understanding
Application				
Data Analysis • central tendency • spread or dispersion • outliers • *z*-scores • regression • graphs	• Statistical tools incorporated to a limited extent • Limited use of appropriate mathematical procedures	• Statistical tools incorporated sometimes • Mathematical procedures sometimes appropriate and sometimes correct	• Statistical tools incorporated most of the time • Mathematical procedures generally correct/appropriate	• Statistical tools incorporated extensively • Mathematical procedures always correct/appropriate and highly detailed
Communication				
Math Terminology and Writing Conventions • uses proper math conventions for graphs, calculations, tables, etc. • uses proper writing conventions	• Limited use of correct math terminology and notation • Proper writing conventions followed to a limited extent	• Some use of correct math terminology and notation • Proper writing conventions followed sometimes	• Considerable use of correct math terminology and notation • Proper writing conventions mostly followed	• Thorough and meticulous use of correct math terminology and notation • Proper writing conventions followed extensively
Organization • key ideas presented logically • appropriate headings (title page, table of contents, etc.) • sources cited appropriately	• Report not clear or logical; ideas rarely connect • Few aspects are organized • Some references included • Sources cited incorrectly	• Report somewhat clear or logical; ideas connect sometimes • Some aspects are organized • Some references missing • Sources cited mostly following proper format (few errors)	• Report mostly clear, logical; ideas mostly connect • Many aspects are well-organized • All references included • Sources cited mostly following proper format (few errors)	• Report clear, logical; ideas thoroughly connect • Most aspects are well-organized • All references included • Sources cited following proper format
Thinking				
Conclusions, Assumptions, Limitations, and Reflection	• Identifies few assumptions, limitations, or ideas for future study • Conclusions supported minimally by the mathematical analysis and reasoning	• Identifies some assumptions, limitations, or ideas for future study • Conclusions somewhat justified by the mathematical analysis and reasoning	• Identifies many assumptions, limitations, and viable ideas for future study • Conclusions generally justified and supported by the mathematical analysis and reasoning	• Assumptions, limitations thoroughly discussed; detailed suggestions for future study included • Conclusions consistently justified by the mathematical analysis and reasoning

Statistics Project Presentation Rubric

	Level 1 (50–59%)	Level 2 (60–69%)	Level 3 (70–79%)	Level 4 (80–100%)
Hypothesis and Introduction				
Topic Question and Introduction	• Hypothesis stated, but not clear or justified • Background knowledge stated to a limited extent	• Hypothesis makes sense; could be more clear • Background knowledge somewhat evident	• Clear hypothesis • Background knowledge usually evident	• Clear, justified hypothesis • Background knowledge of topic clearly evident
Technology Skills (If Technology Is Used in the Presentation)				
Technology	• Uses technology with difficulty	• Uses technology with some proficiency, or overuses it	• Makes effective use of technology	• Technology enhances the quality of the presentation
Visual Aids				
Visuals (graphs and slides) • clear/readable • neatly presented • easy to understand • graphs clearly labelled	• Visual aids do not support the conclusions or are almost non-existent	• Visual aids sometimes support the conclusions	• Visual aids generally support the conclusions	• Visual aids are used consistently and thoroughly support the conclusions
Presentation Skills				
Communication • voice is clear, relaxed, audible • appropriate eye contact/body language • appropriate pace • clear, concise explanations • engaging and enthusiastic	• Presenter's voice rarely clear/audible • Presenter frequently reads from the screen or notes • Presentation slow and choppy	• Presenter's voice somewhat clear • Presenter sometimes reads from the screen or notes • Presentation flows, but moves slowly	• Presenter's voice usually clear • Presenter occasionally reads from the screen/notes • Presentation flows well; pace keeps the audience interested	• Presenter's voice always clear/audible • Presenter maintains eye contact with the audience, rarely reading from the screen/notes • Presentation is highly engaging
Organization • key ideas presented logically • focused and on topic	• Not very clear or logical; ideas rarely connect • Few aspects are organized	• Somewhat clear or logical; ideas connect sometimes • Some aspects are organized	• Often clear, logical; ideas mostly connect • Many aspects are well-organized	• Clear, logical; ideas thoroughly connect • Presentation is well-organized
Ability to Engage the Audience				
Response to Audience Questions	• Responds to few audience questions, with limited justification or mathematical reasoning	• Responds to some audience questions, with somewhat detailed justification or mathematical reasoning	• Responds to most audience questions, with detailed justification or mathematical reasoning	• Responds to all audience questions, with very detailed justification or mathematical reasoning

Statistics Project Peer Critique Form

Rate the quality of the presentation or report in each category using the following scale:

1 = poor 2 = fair 3 = good 4 = very good 5 = outstanding

	1	2	3	4	5
Introduction					
Hypothesis					
Description/justification of sampling technique and population					
Data analysis: • central tendency • spread or dispersion • outliers • *z*-scores • regression					
Visuals (graphs and slides)					
Organization					
Communication					
Technology					
Conclusions					

Highlights:	Things to improve or add: