Name: NOTES KEY
Date: _____

PROBABILITY & STATISTICS

Calendar of Chapter: See the 'Homework' link on the webpage

What You'll Learn:

8.5 & 8.6: Probability Pts 1&2 –probability basics & simple probability calculations
8.1/8.2 Part A – difference between population & sample, & different sampling methods
8.1/8.2 Part B – how to conveniently collect unbiased, useful data
8.4: Misleading Graphs – how data can be presented creatively to support different points of view

What is the definition of '**probability**'?

What is a **survey**?

8.5 - Probability – Part 1

Focus: To learn what probability is, and how it's calculated.

Warmup:
i) TACO BELL!
ii) After research & discuss define the term 'probability' in your own words? Be ready to share with the class.

Ex1 – Find the probability of each:
a) Having a coin land on heads after flipping
b) Rolling a die and getting a six

What are three ways you can represent a probability mathematically? How can you represent it in a diagram?

Theoretical Probability
Using our examples above, how can we write a mathematical formula for probability?

Ex2 – Answer the following in fraction, decimal, & percent, and in a tree diagram:
a) When rolling a die, what is the probability of rolling a 5 OR a 6?
b) What about a 4 OR 5 OR 6?
c) When flipping a coin, what is the probability of it landing on heads OR tails?

i) How many different combinations are there?

ii)

PROBABILITY: the extent to which an event is likely to occur, measured by the ratio of favourable cases to the number of cases possible

a) Equal chance of heads or tails so 50%

b) equal chance of 1, 2, 3, 4, 5, 6 so $\frac{100\%}{6} = 16.7\%$

fraction: $\frac{1}{2}$

decimal: 0.5

percent: 50%

Diagram: TREE

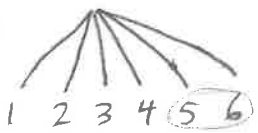
FLIP A COIN:

Theoretical Probability:

$$P(E) = \frac{n(E)}{n(A)}$$

probability of an event occurring $=$ $\frac{\text{number of cases for the event to occur}}{\text{number of all possible cases}}$

a)



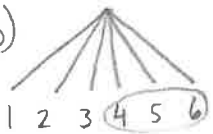
$P(5 \text{ or } 6) = \frac{2}{6} = \frac{1}{3}$

Decimal: $1 \div 3 = 0.\overline{3}$

Percent: $0.3333 = 33.\overline{3}\%$

$\times 100$

(b)



$P(4, 5, \text{or } 6) = \frac{3}{6} = \frac{1}{2}$

Decimal: 0.5

Percent: 50%

(c)

$P(H \text{ or } T) = \frac{2}{2} = 1$

Decimal = 1

Percent = 100%

Ex3 – (a) Flip the coin 6 times and record the result of each flip:
b) Do your results equal the theoretical probability of flipping heads?
c) Would your result happen every time?

Flip #	1	2	3	4	5	6
Result						

a) Answers will vary

b) Answers will vary

c) No, the results can differ

Experimental Probability
How can experimental probability become more accurate?

Experimental Probability: the ratio of the number of times an event occurs to the total number of trials or times the activity is performed.
The more trials done, the closer the experimental probability comes to the theoretical probability

Ex 4 – (a) If you asked 8 people in Canada what colour their hair was, and 6 said blond, what would the experimental probability be?
b) Could we conclude that probability is accurate for Canada?
c) How could we improve the accuracy?
d) What other factors in the survey matter?

a) $\frac{6}{8} = \frac{3}{4} = 0.75 = 75\%$

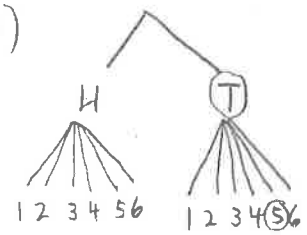
b) No, it is well known that less than 75% of Canadians have blond hair

c) Ask more people!

d) Who is being asked? Is it random?

Back to Theoretical:
Ex5 – If you flipped a coin & then rolled a die, what is the probability of getting 'tails' and then getting a 'five'?
a) Draw a tree diagram
b) Find the probability as fraction, decimal, and %.
c) Write the **sample space** for all possible outcomes (this is the denominator of your fraction).

a)



b) 1 pathway in 12 possible

$P(T \text{ and } 5) = \frac{1}{12}$

dec $= 0.08\overline{3}$

% $= 8.\overline{3}\%$

{ H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6 }

Ex 6 – Suppose you roll a six-sided die, and then a four-sided die. What is the probability of rolling a 5 OR 6, and then a 3 OR 4? Use the **sample space** for assistance.

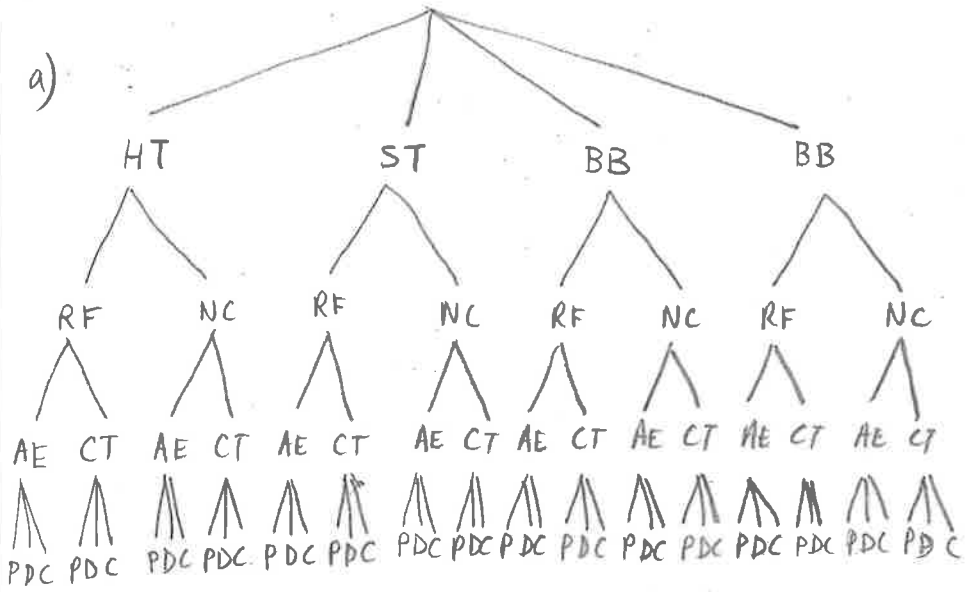
$$P(5 \text{ or } 6 \text{ AND } 3 \text{ or } 4) = \frac{4}{24} = \frac{1}{6}$$

$$= 0.1\overline{6}$$

$$= 16.\overline{6}\,\%$$

Ex 7 – Solve the TACO BELL problem:
a) How many possible different combos are there?
b) What is the probability of having a combo with a burrito & NOT Dr. Pepper?

a)



P = pepsi
D = diet pepsi
C = Dr Pepper (cherry!)

how many pathways?
- count the bottom row!

48

OR $4 \times 2 \times 2 \times 3 = 48$

b)

$$P(\text{burrito \& } \overline{\text{Dr. Pepper}}) = \frac{16}{48} = \frac{1}{3} = 0.\overline{3} = 33.\overline{3}\,\%$$

('not')

8.6 - Probability Part 2

Focus: To understand subjective probability and learn probability notation.

Warmup:
Estimate the probability that it will rain sometime in the next two days.

Answers will vary

Explain what you based your probability on.

weather today ; season ; forecast

What assumptions did you make to help you decide on a probability?

Answers will vary

(ex) forecast is correct, weather today, weather last few days

Subjective Probability (Subjective Judgement)

Subjective Probability: an estimated probability that describes a person's belief in an event. It can be based on prior experimental probability, educated guesses, prior experience, beliefs, and/or bias.

Ex1 – Classify each as theoretical, experimental, or subjective probability:
a) Jon scored 42 baskets out of 50 attempts, so made 84% if shots
b) Jane thinks her team has a 60% chance to win
c) Ben says there is a 1 in 6 chance he can roll a 4 with a die.

a) experimental

b) subjective

c) theoretical

Sometimes, prior experimental probability is used to come up with subjective probability.

(ex) Ben made 84% of his shots today, so he is sure he will make most of his shots in the next game

Ex2 – Joe scored 2 goals on 6 shots in hockey. How many goals will he score today if he gets 3 shots?

$\frac{2}{6} = \frac{1}{3} = 0.\overline{3} = 33.\overline{3}\%$

3 shots today, so 1 goal

this is subjective. He could get 3 goals, 2, 1, or none. Who knows for sure!?

Mathematicians love to abbreviate.

Ex3 – (a) 'the probability of rolling a 5 or 6' can be written as:

$P(5 \text{ or } 6)$

b) We can also write the 'probability of not rolling a 5 or 6' like this:

'not'

$P(\overline{5 \text{ or } 6})$

What is the relationship between the two probabilities above?

$P(\overline{5 \text{ or } 6}) = 1 - P(5 \text{ or } 6)$

Ex4 – Look at the picture of the door. What questions come to mind?

How many different possible 4 number combinations are there?

Ex5 – If you rolled a die and then flipped a coin, find each without a diagram:
a) *P(6 and heads)*
b) *P(1-3 and tails)*

a) $P(6 \text{ and heads}) = P(6) \times P(H) = \frac{1}{6} \times \frac{1}{2} = \frac{1}{12} = 8.\overline{3}\%$

b) $P(1\text{-}3 \text{ and tails}) = P(1\text{-}3) \times P(T) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 25\%$

Ex6 – Suppose you flipped cards in a deck. What's the probability of flipping:
a) two hearts?
b) a 4 and then a 7?
c) three straight aces?
d) a 9 and 10 in any order

a) $P(2 \text{ hearts}) = P(\text{1ST heart}) \times P(\text{2nd heart}) = \frac{13}{52} \times \frac{12}{51}$

$= \frac{156}{2652} = 0.059 = 5.9\%$

b) $P(4 \text{ then } 7) = P(\text{1ST } 4) \times P(\text{2nd } 7)$

$= \frac{4}{52} \times \frac{4}{51} = \frac{16}{2652} = 0.006 = 0.6\%$

c) $P(3 \text{ aces}) = \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50} = \frac{24}{132\,600} = 0.000181 = 0.0181\%$

d) $P(9 \text{ and } 10 \text{ OR } 10 \text{ and } 9) = \frac{8}{52} \times \frac{4}{51} = \frac{32}{2652} = 0.012 = 1.2\%$

8.1/8.2 Part A - Populations & Samples

Focus: To learn the differences, advantages, and disadvantages of populations and samples.

Warmup: If you wanted to find out what the most popular social media site amongst students at school, how would you go about it?

Answers will vary

What is a **survey**?

an investigation of the opinions or experience of a group of people, based on a series of questions

What is a **sample?**

a portion drawn from a population, which could lead to estimates of the attributes of the whole population

What is a **population?**

a particular section, group, or type of people

What is a **census**?

an official count or survey of a population

What are the advantages & disadvantages of using a **sample** vs. using a **population** for a survey?

	Advantages	Disadvantages
Sample	- less people, less work - cheaper	less people to base results on. Results may not be representative of the population
Population (Census)	conclusions drawn from the survey will be representative of population	huge task to complete survey, in most cases - time AND money

Ex1 – Would you use a sample or do a census to find out the most popular cell phone brand in our class?

Census - wouldn't be overly difficult to ask EVERY member in a classroom.

Ex2 – Write a scenario where a sample would be the best choice.

(ex) want to find the favourite video game in Canada

Be ready to share with the class.

What are different **'Sampling Methods'**? (How do we decide who to ask in the school about social media?)

What are the advantages/disadvant of each sampling method?

see p.276 in text for descriptions

1) Convenient Sampling - 'just ask who's around'
2) Voluntary Response Sampling - individuals choose to be involved
3) TV/Online Survey - limited to people who watch a show or go to a website
4) Random Samples - everyone in population has equal chance
   *also Stratified Random Sampling: everyone in each defined group of a population has equal chance

#4 - gives least bias, but challenging to set up
#1,2,3 - huge chance of bias, but easier to get survey results

Ex1 – If you wanted to see how many teenagers in Victoria liked country music:
a) What would be your population?
b) Would you do a census? Explain.
c) How would you sample the population? What, specifically, would you do? Explain.
d) How many people do you think is enough to get a **representative** sample?

e) What would your question be so that it was unbiased?
f) Write a question that would be biased.

a) all teenagers in Victoria
b) No, too difficult and expensive
c) Answers will vary. Is there a chance of bias?
d) Answers will vary, but the more, the better!

e) Answers will vary, but language should be as neutral as possible (no bias in language)
(ex) Q: What is your opinion about country music?
Circle One: Like it Don't Like it

f) Do you hate country music?
Circle One: Yes No

8.1/8.2 Part B - Question Crafting & Data Collection

Focus: To understand bias and other factors when building a survey & analyzing results.

Warmup:
What does **bias** mean?
Discuss and/or research.

in favour of, or against, something

How could the sampling method, size, survey question(s), or other factors possibly bias results?

Potential Problem	What It Means	Example
bias	the question influences responses	Do you think biking is lame?
use of language	how a question is crafted can influence an answer	Do you like ← country music?
timing	when the data is collected could affect results	Do you like your job? - after a long, rough day at work
privacy	if data collected is personal	health issue questions
cultural sensitivity	avoid asking questions that could be offensive to cultures	How do you cook ham? (not every culture eats ham)
ethics	data not used for purposes other than the survey	Giving consumer info to private corporations
cost	how much will the survey cost?	- how many people? - cost of printing etc?
time	how much time will the survey take for people	- A 30 min survey may be too long for most people

For each scenario, explain any problem(s) and the effect it could/would have on results.

Ex1

a) Jim asks 20 teens at Mayfair if they like the mall. 18 say 'yes'. He concludes 90% of teens in Vic like the mall.

a) if they're at the mall, they probably like it ⇒ bias
- results shouldn't be assumed for the population

b) A survey is conducted to find out the level of school spirit. A random sample is polled right after the school wins the Provincial Soccer Championship.

b) Timing → try to ask on a more 'regular' day
results will likely misrepresent a higher school spirit

c) Mike wants to ask every teenager in Toronto if they want a skate park downtown.

c) Time + Cost to find and ask EVERY TEEN in Toronto
- survey may never get completed

d) Jane does survey downtown, asking people if they have ever shoplifted.

d) people may not want to admit this ⇒ privacy, possibly ethics
- data may be inaccurate

e) Thomas wants to ask a random sample of people how much income they make, and how much they spend on Xmas presents.

e) Privacy, Timing
Privacy: -people may not want to divulge income
-people may be reluctant to share consumer info
Timing: people may over or underestimate spending depending on the time of year of the survey
- incomplete or inaccurate data

f) Carmen asks people 'You wouldn't ever throw out recyclables, would you?'

f) Use of Language could lead people to feel pressure to say 'no'
- lead to overestimation of recycling tendencies

Ex2 – Craft a question for students in your grade that is neutral (non-biased).

Answers will vary

8.4 - Misleading Graphs

Focus: To recognize and understand graphs that have misleading traits.

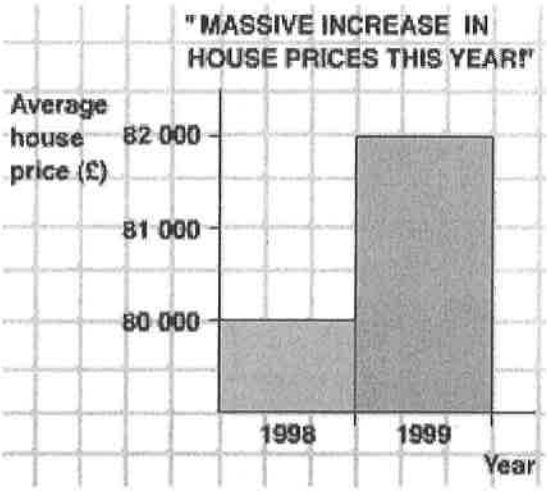
Warmup:
A real estate agent is trying to convince a family to sell their house. He shows them this bar graph.
a) Can you describe anything **misleading** about the graph?
b) Why might the agent use the graph the way it is?
c) Is the agent being dishonest by using this graph?
d) How could the graph be fixed so that it's more accurate looking?

a) The y axis starts at 79 000 instead of zero

b) The second bar is much higher than the first, suggesting a big increase

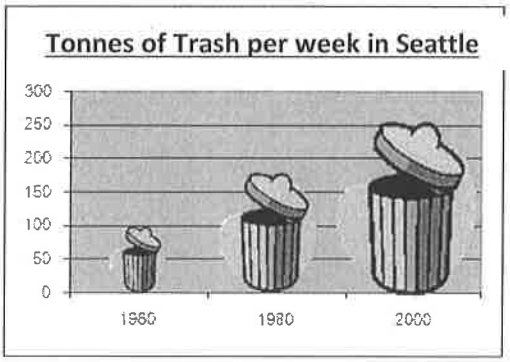
c) Sort of - the info is accurate but not presented in an honest context

d) Have the y axis start at 0



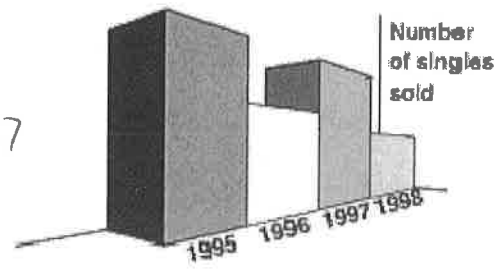
Ex1 – A group of social activists are lobbying for cities to increase their recycling efforts. They use this graph to support their cause. What is **misleading** about the graph?

The trash cans not only get taller, but they get wider. This distorts the differences, making them seem larger than they actually are



Ex2 – The British rock band Oasis insists 1995 was their best year, and show this bar graph as support. What is **misleading** about the graph?

It's a 3-D perspective, so even though 1995 + 1997 are equal, 1995 appears larger at a glance



Describe different ways that graphs can be made to look misleading.

- not starting axes at zero
- 3D perspectives
- size (width) of pictures
- no numbers on axes
- not including all data

Why would somebody construct a graph that is misleading?

to help support and advance their interests

Would you consider this unethical?

opinion - the data is not fake or incorrect, it is just presented in a misleading fashion

What do we, as citizens have to be wary of whenever we use data and/or graphs to help us form opinions?

- scrutinize the data and how it's presented as there is always an agenda from the source

Ex3 – Suppose you want to bring attention to poverty in our city and you get an advertising spot online. Use the information provided to make a **misleading graph** that will support your cause so that people may take notice.

Number of people in poverty in Victoria 4 years ago: 95 000
2 years ago: 94 000
this year: 96 500

leave this out!

use a y axis that doesn't start at zero!