

Booklet 5 answers

Page 2-4 Population Based Rates:

- 1a. 50 per 100,000 b. 468.75 per 100k ; 2a. 3 per 100k b. 145 per 100k
3a $2/1200 = 166.67$ per 100k b. 1250 per 100k; 4. 10 per 260,000 = 3.85 per 100k
5. 2 per 100k; 6. population is of Kingston, Ontario – prevalence is 9.7 per 100k
7. 16.9 per 100k; 8. 146.7 per 100k; 9. 404.66 per 100k
10. 55% or 550 per 1000 - 55,000 per 100k is not very meaningful, but correct.

11. 248.79 (or simply 249-rounded up)

12. 80975.76 (in this case rounding up does not make sense. If you don't have a whole person then rounding up can be considered a mistake.) report the actual or 80975

13. 219.13 per 100,000 14. 0.0083 per 100,000 15. 476.17 per 100,000

page 5 1a it could be in a simple survey: “are you willing to sacrifice...etc) and get a % who say yes... i.e. count of objects (Canadians) and their response to yes/no question.

b. similar to a... survey response ‘did you play board games and get a % who say yes. i.e. count of objects (Canadians) and their response to yes/no question.

c. this is a yes/no question (again) but you can't just survey people to get an answer – one would need to conduct an experiment measuring the duration (# of days of cold symptoms) and comparing those who took echinacea to those who did not.

d. ask Canadians and French (yes/no question) are you a tea drinker then count how many say yes and calculate a rate of ‘yes’ as %

e. This is the only one that cannot be quantified simply... ‘improve the lives of the poor’ is hard to think of as a ‘quantity’ or a measurement or count of objects etc.

f. more tea per day.. means this is a bit different from d in that you are asking individuals to quantify how much tea they drink (in #cups?) rather than just yes/no – do you drink tea

page 6. 1a yes as it gets across the options as described

b. good again... heart rate is what is being measured... all the other stuff (walking up stairs etc) is not what is being measured

c. rate of preference is not the quantity being studied. Preference belongs there for sure.. perhaps ‘classroom type preference’ is better... rate of preference is what you will examine after data collection here

d. yes... it is good

page 7: 1a measurement b. measurement c. categorical d. categorical e. categorical f. measurement

page 8. 1a head circumference (meas) b. urbanvs rural (cat) bicycle ownership (yes/no) cat

c. change in GPA (measurement) (meaning subtraction of after – before) so only one variable... but it may mean collecting two and generating the third from the first 2 😊

d. minutes of exercise (meas) GPA(meas)

e. over/under (cat) doctor burnout yes/no (cat)

page 9: 1a hard to say what is typical except that there are not too many nurses making more than 2 calls... with a very high number of 'zero' calls. The distribution is slanted with a large hump to the left and a tail to the right. It is interesting that so many are 0 and that there are a few that are really high including one patient who needed over 10 calls.

1b. this is a setup that is easier on the eye as there is a centre to the distribution of teacher marks.

Typical marks is in the centre... about 75 or so What is really weird is that there are so many with percentage grades over 100... there must be some sort of mistake here!!!

Page 10: 1a. There are 5 categories in the pie chart to the left (Foreign born, Canadian born non-indigenous, Canadian born Indigenous status unknown, Unknown and Canadian born indigenous). But the Canadian born indigenous category got split up into 3 categories Metis, Inuit and First Nations. The correct answer is then either 5 or 7 😊 . b 70% of those Canadians with tuberculosis are born outside of Canada. c Canadian born indigenous (19% of cases of tuberculosis cases) are further broken down as Metis, Inuit and First Nations; here we see that 63% indigenous TB cases are in First Nations individuals etc. d. 8% of tuberculosis cases as Canadian born non-indigenous is very low as non-indigenous make up over 90% of the Canada wide population.
e I would also like to know how many total cases there are – if there are a million folks with TB then I am very worried, if there are only 100, then I would not be worried at all.

page 11: 1a 36, b 48, c 0, d 94, e x_0 does not exist; 2a $33+36+37 = 106$ b 637 c $0 + 21 + 33+36+37 = 127$
d $(0-2)^2 + (21-2)^2 + (33-2)^2 + (36-2)^2 = 2482$

page 12-13 Working with Data

1. $N=8$, $\min=82$, $\max=188$, $x_4=156$, $\text{mean} = 154.5$; $\text{median}=161.5$, $\text{mode}=\text{none exists}$; $\text{range}=106$, $\text{std.dev}=30.21$
2. $N=7$, $\min=75$, $\max = 186$, $x_4 = 144$, $\text{mean} = 137.43$, $\text{median} = 144$; $\text{mode}=\text{none exists}$; $\text{range}=111$, $\text{stdev}=42.62$
3. $N=7$, $\min=62$, $\max = 177$, $x_4 = 155$ (or 62 in unordered set), $\text{mean} = 132.86$, $\text{median} = 155$; $\text{mode}=166$; $\text{range}=115$, $\text{stdev}=42.95$
4. $N=10$, $\min=44$, $\max = 201$, $x_4 = 115$ (or 136 in unordered set), $\text{mean} = 126.5$, $\text{median} = 126.5$; $\text{mode}=\text{none}$; $\text{range}=157$, $\text{stdev}=51.44$
5. The Gulas have the highest mean height. Their mean height is $(154.5 - 126.5 =) 28\text{cm}$ more than the shortest mean height (the Zingers) as a raw difference and thus we can say that the mean height of Gulas is $(28/126.5 =)$; 22.13% higher than the mean height of Zingers.
6. The Gulas have the highest median height as well. Their median height is 35cm taller than the median Zinger height which translates to a 27.67% difference.
7. Since Gulas are taller in mean and median – it is easy to argue that they are the tallest of the given families. But you can also look at max heights (assuming that there are a few very young children in the Zinger and other families) and argue that the measures of centre aren't appropriate for these comparisons
8. Variability is easier – Range is not very useful – std devs are the best and here there is much more variability within the Zinger family than within all other families.

Page 14. 1a 25th percentile ≈ 50 ; b 75th percentile ≈ 75 ; c. 90th percentile ≈ 91 ; 95th percentile ≈ 94
2a 70th b. 30% had a higher mark than you. C. 44 is roughly at the 15-20th percentile, so 15-20% of the students had a mark lower than yours.