

STAT 157 – Test 1

March 24, 2020

Name: Solution Set

Instructor: Patricia Wrean

Total: 30 points

1. (5 points) In a statistics class, each student had their height measured along with their arm span, which is the distance measured from fingertip to fingertip when arms are spread wide. Some of the results are shown in the table below. All measurements have been rounded to the nearest centimetre.

height	arm span
176	170
165	166
160	162

For the following questions, circle the best answer or fill in the blank.

(a) This data is

①

(i) univariate

(ii) bivariate

(iii) multivariate

← two measurements, height and arm span

(b) Name a variable being measured.

①

height or arm span

(c) For your answer to part (b), this variable is

②

(i) qualitative

(ii) quantitative and discrete

(iii) quantitative and continuous

← for either answer in (b), because quantity is rounded

(d) The experimental unit is

①

(i) a student

(ii) the height and arm span of a student

(iii) the height and arm span of all students in the class

(iv) the height and arm span of all statistics students

← -1 if selected (iii)

2. (6 points) The gas prices (in cents per litre) for regular gas at a sample of gas stations in Victoria are as follows:

132.9, ~~120.9~~, ~~117.9~~, ~~127.9~~, ~~117.9~~, ~~114.9~~, ~~128.9~~, ~~117.9~~
~~114.9~~, ~~117.9~~, ~~117.9~~, ~~117.9~~, ~~120.9~~, ~~127.9~~, ~~128.9~~, ~~132.9~~

(a) State the mean, median, and range of this data set. If appropriate, round to two decimal places.

3

$$\bar{x} = \frac{\sum x_i}{n} = \frac{979.2}{8} = 122.4$$

$$\text{median} = \frac{117.9 + 120.9}{2} = 119.4$$

$$\text{range} = \text{max} - \text{min} = 132.9 - 114.9 = 18$$

mean: 122.4
 median: 119.4
 range: 18

(b) The standard deviation of this data set is 6.57 cents per litre. Calculate the z-score of the lowest data point and indicate whether that value is likely or unlikely.

$$z_{\text{low}} = \frac{x - \bar{x}}{s} = \frac{114.9 - 122.4}{6.57} = -1.14155$$

z-score: -1.14
 likely / unlikely

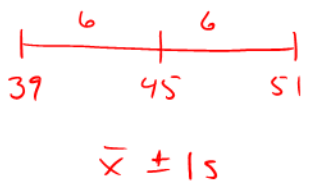
$|z| < 2$ is likely

(-1) if dropped negative sign

3. (4 points) The mean of a distribution is 45 and the standard deviation is 6.

(a) If you know that the distribution is unimodal and symmetrical, what can you say about the percentage of values that fall between 39 and 51?

2



Empirical rule says $\sim 68\%$ of values in this interval

no approx $(-\frac{1}{2})$
 only Tcheby (-1)

(b) If you know nothing about the shape of the distribution, is it possible that there are no values that fall between 39 and 51? Explain briefly.

2

yes, because Tchebysheff's theorem says $\geq 1 - \frac{1}{k^2} \geq 1 - \frac{1}{1^2} \geq 0\%$ i.e. within this interval

(-1) 's yes but no reasoning
 $(-\frac{1}{2})$ no \geq

4. (5 points) The online video game store Steam is having its annual sale. Each game has a different price and you may assume that there are many games for sale.

- (a) If every game has its price reduced by 25%, what happens to the following quantities? Be as specific as you can!

mean: decrease by 25%

standard deviation: decrease by 25%

- (b) If, instead, every game has its price reduced by \$15, what happens to the following quantities? Be as specific as you can!

median: decrease by \$15

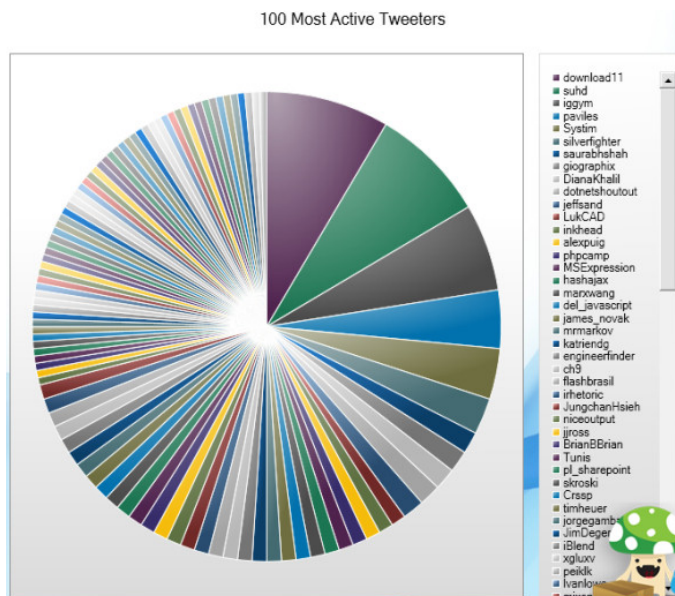
range: stay same

- (c) Only one game is on sale: the most expensive game will have its price reduced so that it is now the same price as the second-most expensive game. All other items keep their current prices. Does the following quantity increase, decrease, or stay the same?

median: stay same

5. (2 points) The following graph is a pie chart showing the top 100 Twitter users. There is one main reason that this graph is badly designed. Give that reason.

Please note that you do not need to be able to read the legend of the graph to answer this question.



too many
categories!

⊖ if only said
colours reused

6. (4 points) A sports researcher is studying the resting heart rate for current NHL hockey players. For the following situations, identify the sampling plan used to pick a sample of these hockey players.
- (a) All of the hockey players are sorted into groups based on the position that they play (center, left wing, goalie, etc.), and then a random selection of players from each group are measured. stratified
- (b) All hockey players currently playing in the NHL are put on a list, and then the 3rd player is selected and then every 15th player after that. 1-in-15 systematic
- (c) A certain number of teams are selected at random and then every player on that team is measured. cluster
- (d) All hockey players currently playing in the NHL are put on a list and a certain number are randomly drawn from the list. simple random
7. (2 points) State whether the following studies are experimental or observational by circling the correct choice.
- (a) A study took a group of adults and randomly divided them into two groups. One group was told to drink tea every night for a week, while the other group was told not to drink tea that week. Researchers then compared when each group fell asleep. Expt / Observ
- (b) Another study took a random sample of people and examined their social media habits. Each person was classified as either a light, moderate, or heavy social media user. The researchers looked at which groups tended to be happier. Expt / Observ
8. (2 points) Students who were studying in the Math Lab on a particular day were asked to fill in a questionnaire about whether the Math Lab hours are suitable for them. You may assume that a large number of students were sampled.

Do you think that this will give a representative sample of students who might want to use the Math Lab? Briefly explain.

no, you are only measuring students who are there already and not the ones who cannot make the current hours

this is a convenience sample