QUIZ 1 - SOLUTIONS

(LESSONS 1-10: INTRO and DESCRIPTIVE STATISTICS)
MATH 119 – SPRING 2023 – KUNIYUKI

1) (3 points). A department store gives a gift card to every hundredth customer who makes a purchase. What sampling method is being used here? Box in the best answer:

   • simple random sampling
   • systematic sampling
   • cluster sampling
   • stratified sampling

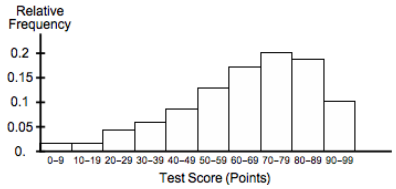
2) (5 points). A poll randomly selects 805 American adults. They are asked:

   "Do you believe computer scientists' ability to develop AI [artificial intelligence] would do more good than harm to society, more harm than good, or equal amounts of harm and good?" ('Don't know' was also an option.)

   Based on the observed frequencies below, find the corresponding relative frequencies. Write the answers in **percent form**. If you round off, round off to three significant digits. It may help to show some work for partial credit. (Note: These figures are based on a January 2023 Monmouth University poll.)

Response	**Frequency**	**Relative Frequency (as a percent)**
More good than harm	75	$\frac{75}{805} \approx 0.0932$ or 9.32%
More harm than good	333	$\frac{333}{805} \approx 0.414$ or 41.4%
Equal amounts	373	$\frac{373}{805} \approx 0.463$ or 46.3%
Don't know	24	$\frac{24}{805} \approx 0.0298$ or 2.98%
	Sum = n = 805	

3) (7 points). 70 students take an exam. A relative frequency histogram for their scores is below.



• a) (3 points). Estimate the relative frequency of scores in the 60s (between 60 and 69 points), as in class.

The relative frequency is between 0.15 and 0.20.

• b) (4 points). Describe the distribution shape. Consider modality and skewness.

This distribution is unimodal and left-skewed.

4) (3 points). Based on the (correct) answer to Problem 3) above, which statement below is more likely to be true for the distribution in Problem 3)? Box in one:

• The mean is less than the median.

• The median is less than the mean.

The mean tends to be pulled towards the "tail" of a skewed distribution, away from the median.

5) (40 points). Five adults are randomly selected in a town in Japan. They are asked for the number of times they have visited their doctor in the past year. The data below are their responses.

11 14 16 11 12

(Note: This problem was inspired by an article in *Forbes* magazine in September 2014. By comparison, Americans visit their doctor on average four times a year.) https://www.forbes.com/sites/niallmccarthy/2014/09/04/americans-visit-their-doctor-4-times-a-year-people-in-japan-visit-13-times-a-year-infographic/?sh=52f8486ce347

(Show work as we have in class. *Treat as sample data.*) Based on this data …

• a) (4 points). Find the **mean** number of doctor visits. Write the exact answer to one decimal place.

The mean is: $\frac{11+14+16+11+12}{5} = \frac{64}{5} =$ 12.8 visits.

Note: According to the *Forbes* magazine article, the mean is about 13 visits.

• b) (2 points). Find the **median position number**.

The median position number is: $\frac{n+1}{2} = \frac{5+1}{2} =$ 3.

• c) (4 points). Find the **median** number of doctor visits.

First, sort the values: 11 11 **12** 14 16
The median is the **third** lowest value: 12 visits.

• d) (4 points). Find the **mode** of the numbers of doctor visits.

The mode is the most frequent value, 11 visits.

• e) (4 points). Find the **midrange** of the numbers of doctor visits. Write the exact answer to one decimal place.

The midrange is: $\frac{Min + Max}{2} = \frac{11 + 16}{2} =$ 13.5 visits.

• f) (4 points). Find the **range** of the sample data values.

Range = Max – Min = 16 – 11 = 5 visits.

• g) (3 points). Box in the most appropriate sentence:

- The midrange and the range are measures of center.
- The midrange is a measure of center but the range is a measure of spread.
- The midrange is a measure of spread but the range is a measure of center.
- The midrange and the range are measures of spread.

• h) (6 points). Fill out the following table. For the sample mean, use your answer to Part a). Write exact answers; do not round off.

	Data (x) values in years	**Deviations** $(x - \bar{x})$ values	**Squared Deviations** $(x - \bar{x})^2$ values
Adult #1	11	11 – 12.8 = – 1.8	3.24
Adult #2	14	14 – 12.8 = 1.2	1.44
Adult #3	16	16 – 12.8 = 3.2	10.24
Adult #4	11	11 – 12.8 = – 1.8	3.24
Adult #5	12	12 – 12.8 = – 0.8	0.64

• i) (1 point). What do the deviations from the sample mean add up to? 0.

• j) (4 points). Find the **sample variance**. Write the exact answer, to one decimal place.

The sample variance is: $s^2 = \frac{\sum (x - \bar{x})^2}{n - 1} = \frac{18.8}{4} =$ 4.7 visits^2.

• k) (4 points). Find the **sample standard deviation**. Round your answer off to one decimal place.

The sample SD is: $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} = \sqrt{4.7} \approx$ 2.2 visits.

6) (11 points). A student's grade report for a term is below. Write the exact GPA for the term in decimal form. As stated in class, grades of "A," "B," "C," "D," and "F" are worth 4, 3, 2, 1, and 0 grade points, respectively. A "+" modifier adds 0.3, while a "-" subtracts 0.3. Show work, as in class!

Grade Report		
Course	**Number of Units**	**Grade**
Math	5	B+
Biology	3	C-
Psychology	5	A

Number of units taken by the student = $\sum w = 5 + 3 + 5 = 13$ units.

$$\text{GPA} = \frac{\sum w \cdot x}{\sum w} = \frac{(5)(3.3) + (3)(1.7) + (5)(4)}{13} = \frac{41.6}{13} = \boxed{3.20 \text{ grade points}}.$$

7) (8 points). So far, your grade record in a class looks like this:

Exam	**% of overall grade**	**Your score (out of 100 points)**
Quiz 1	15%	85
Quiz 2	15%	92
Quiz 3	15%	75
Midterm	25%	80
Final	30%	c

What must you get on the Final to get at least 70% in the class overall? (What kind of score do you need c to be?) Show work, as in class!
Note: You may be surprised at how low the answer is!

$$\frac{(0.15)(85) + (0.15)(92) + (0.15)(75) + (0.25)(80) + (0.30)c}{1} \geq 70$$

$$57.8 + 0.3c \geq 70$$

$$0.3c \geq 12.2$$

$$c > 40.\overline{6}$$

To be safe ("conservative"), let's round up.
You must get at least 41 points on the Final.

8) (14 points). A company claims that the amounts of sugar in its new brand of candy bar have mean 4.00 grams and standard deviation 0.30 grams.

• a) (4 points). Use the "Two SD" (2σ) Rule for Usual Values to give an appropriate interval of usual amounts of sugar in the new brand.

$$(\mu - 2\sigma,\ \mu + 2\sigma) = (4.00 - 2(0.30),\ 4.00 + 2(0.30))$$
$$= \boxed{(3.40\text{ g},\ 4.60\text{ g})}$$

• b) (2 points). Assume for now that the amounts of sugar in the new brand are approximately normally distributed. According to the **Empirical Rule**, about what percent of the amounts of sugar in the new brand are within **one** standard deviation of the mean?

About 68% of the amounts of sugar in the new brand are within one SD of the mean.

• c) (2 points). Assume for now that the amounts of sugar in the new brand are approximately normally distributed. According to the **Empirical Rule**, about what percent of the amounts of sugar in the new brand are within **two** standard deviations of the mean?

About 95% of the amounts of sugar in the new brand are within two SDs of the mean.

• d) (4 points). A scientist studies the new brand of candy bar and randomly selects one bar. The bar has 4.73 grams of sugar. What is the ***z* score** for this bar? Round off your answer to two decimal places. Note: The bar's amount of sugar would be considered **"unusual"** based on the company's claim and the "Two SD" (2σ) Rule for Usual Values.

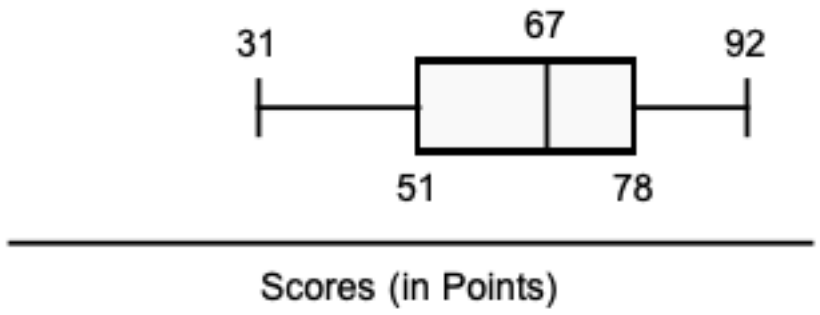
$$z = \frac{x - \mu}{\sigma} = \frac{4.73 - 4.00}{0.30} \approx \boxed{2.43}$$

Observe that this z score is above 2.00, which reflects the fact that 4.73 grams was **not** in the interval of usual values from a).

• e) (2 points). **Interpret the *z* score** from d), as in class.

The amount of sugar in the selected candy bar is **about 2.43 standard deviations (SDs) above the mean** (based on the company's claim).

9) (9 points). The scores on a test (in points) in a large class are summarized by the boxplot (also known as a "box-and-whisker" plot) below. The minimum score is 31 points. The maximum score is 92 points. There are no extreme outliers.



- a) (2 points). A score of 51 points is at which **quartile**?

  51 points is at $\boxed{Q_1}$, the first quartile.

- b) (2 points). A score of 78 points is at which **percentile**?

  78 points is at $\boxed{P_{75}}$, the 75^{th} percentile (which is the third quartile).

- c) (2 points). What is the **median** of the class scores?

  The median is at $\boxed{67 \text{ points}}$.

- d) (3 points). What is the **IQR (Interquartile Range)** of the class scores?

  IQR = $Q_3 - Q_1 = 78 - 51 = \boxed{27 \text{ points}}$.