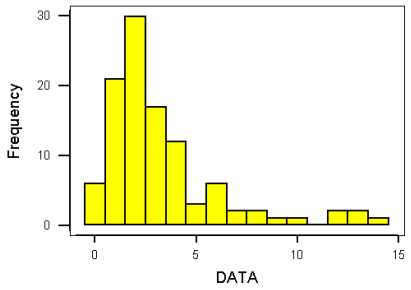
[redacted]

Assignment Assignment-01 due 09/20/2022 at 11:59pm PDT

Problem 1. **(2 points)**

Consider the histogram shown below.



(a) Is this data skewed to the right, symmetric, or skewed to the left? _____
(Enter "SKEWED RIGHT", "SYMMETRIC", or "SKEWED LEFT" without the quotations.)

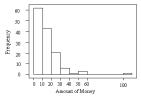
(b) Which is bigger for this data set, the mean or the median, or are they about equal? _____
(Enter "MEAN", "MEDIAN" or "ABOUT EQUAL" without the quotations.)

Correct Answers:

- SKEWED RIGHT
- MEAN

Problem 2. **(3 points)**

In a statistics class with 136 students, the professor records how much money each student has in their possession during the first class of the semester. The histogram below is of the data collected. (**Click on the image for a larger view.**)



(a) The histogram

- A. is skewed right.
- B. is asymmetric.
- C. has an outlier.
- D. all of the above.

(b) The percentage of students with over 20 dollars in their possession is

- A. over 40%.
- B. about 30%.
- C. about 20%.
- D. about 10%.

(c) The number of students with under 10 dollars in their possession is closest to

- A. 70
- B. 60
- C. 50
- D. 40

Correct Answers:

- D
- C
- B

Problem 3. **(2 points)**

The table below summarizes the number of surface flaws found on the paintwork of new cars following their inspection after primer paint was applied by a new method:

No. of flaws	No. of cars
0	3
1	7
2	12
3	11
4	3
5	2
6	2

Part a) Find the mean number of flaws per car. Please give your answer to two decimal places.
The mean number of flaws per car is: ___

Part b)
Find the variance of the number of flaws per car. Please give your answer to two decimal places.

The variance of the number of flaws per car is: ___

Solution:

Here $n = 40$ and the sample mean is

$$\bar{x} = \frac{3\times 0+7\times 1+12\times 2+11\times 3+3\times 4+2\times 5+2\times 6}{40} = 2.45.$$

The sample variance is

$$\frac{3\,(0-2.45)^2+7\,(1-2.45)^2+12\,(2-2.45)^2+11\,(3-2.45)^2+3\,(4-2.45)^2+2\,(5-2.45)^2+2\,(6-2.45)^2}{39}$$

which is 2.151.

Correct Answers:

- `2.45`
- `2.151`

Problem 4. **(2 points)**

A North American company manufactures rivets for use in car production. A sample of forty rivets from the production line had mean 6.709 and standard deviation 0.282 (both in 1/100 of an inch). On communicating these results to the company headquarters in Europe, a request is made for the summary statistics to be converted into millimeters. Given that one inch is 2.538 centimeters, find the mean and standard deviation of the sample in millimeters.

What is the mean in millimeters? (Provide your answer to 3 significant figures.) ___

What is the standard deviation in millimeters? (Provide your answer to 3 significant figures.) ___

Solution: If $\bar{x} = 6.709$ is the mean of the original data, then data in mm are

$$y_i = \frac{x_i}{100}\times 2.538\times 10 = 0.2538\,x_i$$

and the mean $\bar{y}$ is $0.2538\times 6.709 = 1.7027442$ mm.

Similarly the variance of the y_i's will be

$$s_y^2 = 0.2538^2\times s_x^2,$$

which gives a standard deviation of

$$s_y = 0.2538 \times 0.282 = 0.0715716$$

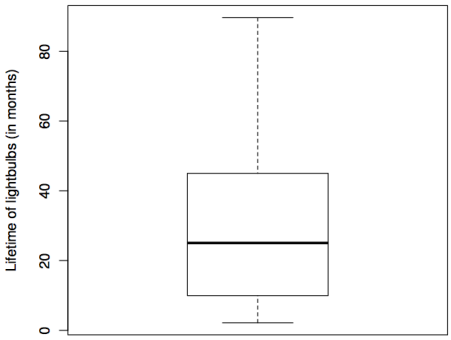
in mm, when original standard deviation was 0.282 (in 1/100 in.)

Correct Answers:

- 1.70274
- 0.0715716

Problem 5. **(1 point)**

The following diagram shows a boxplot for the lifetimes (in months) of a sample of 30 light- bulbs.



What is the median lifetime?

- 10.22
- 25.04
- 30.34
- 43.71
- 89.64

Solution: The correct answers are:
25.04

Correct Answers:

- 25.04

Problem 6. **(3 points)**

For each problem, select the best response.

(a) What percent of the observations in a distribution lie between the first quartile and the third quartile?

- A. 50%.
- B. 25%.
- C. 75%.
- D. None of the above.

(b) Which of the following is least affected if an extreme high outlier is added to your data?

- A. The median.
- B. The standard deviation.
- C. The mean.
- D. None of the above.

(c) To make a boxplot of a distribution, you must know

- A. the five-number summary.
- B. the mean and the standard deviation.
- C. all of the individual observations.
- D. None of the above.

Correct Answers:

- A
- A
- A

Problem 7. **(1 point)**

You are doing a survey on UBC students. Participating students come from different majors and years. The participants are asked how much they pay for tuition this academic year. What is the variable of interest?

- A. UBC students
- B. Study year of a student
- C. Major of a student
- D. Tuition fee paid by a student

Correct Answers:

- D

Problem 8. **(3 points)**

Some students conducted a survey on some people's weekly income:

$200, $600, $178200, $400, $2000, $300, $700, $400

The mean of these people's weekly income is _____.

The median of these people's weekly income is _____.

Which number, mean or median, is a better way to represent these people's weekly income?

- mean
- median

Solution:
Find the mean

To find the mean of a group of numbers, we first add up all numbers and find their sum:

$$\text{sum} = 200 + 600 + 178200 + 400 + 2000 + 300 + 700 + 400 = 182800$$

Next, we divide the sum by how many numbers there are:

$$\text{mean} = \frac{182800}{8} = \$22{,}850.00$$

Solution: The mean of these people's weekly income is $22,850.00.

Find the median

To find the median of a list of number, we first need to order these numbers from smallest to largest:

200, 300, 400, 400, 600, 700, 2000, 178200

The number in the middle is the median. In this case, there are two numbers in the middle: $400, 600$. The median is the average of these two numbers:

$$\text{median} = \frac{400 + 600}{2} = \$500.00$$

Solution: The median of these people's weekly income is $500.00.

Mean or median?

This group of data has an outlier, a number significantly bigger or smaller than the other numbers in the group. The outlier is $178200. This number significantly increased the mean.

If we say these people's mean income is $22,850.00, it gives people a false impression that most people in this group make a lot of money.

Median can "cancel out" outliers. In this case, the median is $500.00. It's true that most people in this group make that amount of money.

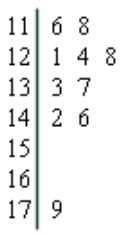
When there are outliers in a group of number, the median is a better way to represent the group of data. When there are no outliers, either mean or median can be used.

Correct Answers:

- `\$22,850.00`
- `\$500.00`
- `median`

Problem 9. **(3 points)**

For a Physics course containing 10 students, the maximum point total for the quarter was 200. The point totals for the 10 students are given in the stemplot below. (**Click on the image for a larger view.**)



(a) To which of the following data sets does this stemplot correspond?

- A. all integers between 116 and 179
- B. 16, 18, 21, 24, 28, 33, 37, 42, 46, 79
- C. 1, 2, 3, 4, 6, 6, 7, 8, 8, 9
- D. 116, 118, 121, 124, 128, 133, 137, 142, 146, 179

(b) This stemplot is most similar to

- A. a time plot of the data with the observations taken in increasing order.
- B. a histogram with class intervals between 110 and 120, between 120 and 130, etc.
- C. a boxplot of the data.
- D. reporting the 5-point summary for the data, with the mean.
- E. None of the above.

(c) The lowest score in the class as a percentage of the total possible points is _____%.

Correct Answers:

- D
- B
- 58