 How do we analyze and then make conclusions from a data set? Or from a scenario in which probabilities are being considered? How do I visually present my data and the outcomes of my analysis? How can we visualize events and outcomes when considering probability events? How do I use data & statistics & probabilities to make decisions? How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?

Mr S has provided you with some data from his shot put throwers from the 2017 season. In this Problem Set, you will be required to complete a statistical analysis of the data to determine which thrower is the "best" thrower on the team. Your statistical analysis will include the following components:

- 1. An appropriate data table and a frequency histogram of the data for each thrower
- 2. Calculation of the mean, median, mode for each thrower
- 3. A five number summary (min, Q1, median, Q3, max), including a box-whisker plot for each thrower
- 4. An appropriate graph that allows you to compare the data from the three throwers on the same graph

Thrower #1	8.74	8.94	9.66	10.01	10.01	8.43	10.25	10.14	9.04	9.30	8.69
	8.85	9.25	9.46	10.23	8.95	9.65	8.79	10.62	9.78	9.26	9.39
Thrower #2	10.39	10.86	10.94	9.00	9.15	9.35	9.35	8.25	8.85	8.95	9.10
	10.20	9.53	8.76	8.03	8.96	9.25	9.98	10.82	10.10	8.96	9.68
Thrower #3	8.79	9.39	9.94	11.47	9.72	8.49	9.63	9.49	9.83	8.82	9.24
	9.13	9.56	9.94	9.75	9.12	8.96	8.83	9.25	9.38	9.62	9.98

THE THROWERS' DATA

As every good mathematician/statistician knows, there are always QUESTIONS you can ask about the data set!!!!!! List some questions you may have about the data set.

1. Prepare a Frequency Table for each thrower.

Distance Thrown	Thrower #1	Thrower #2	Thrower #3
$8.00 \le x \le 8.50$			
$8.50 < x \le 9.00$			
$9.00 < x \le 9.50$			
$9.50 < x \le 10.00$			
$10.00 < x \le 10.50$			
$10.50 < x \le 11.00$			
$11.00 < x \le 11.50$			
$11.50 < x \le 12.00$			

- 2. Prepare a frequency histogram for each thrower.
- 3. Record AT LEAST 3 OBSERVATIONS from your histograms and/or from the data.
- 4. How probable is it that Thrower #1 gives us a throw over 10 m? Under 9 m? How about the other 2 throwers?

PART 1 - STATISTICAL ANALYSIS - Comparing Central Tendencies

- 4. WHY do we calculate measures of central tendency when analyzing data sets? (Give at least 2 reasons)
- 5. Calculate the three measures of central tendency (mean, median and mode) of each of the throwers

C/T	C/T Thrower #1	C/T Thrower #2	C/T Thrower #3
Mean			
Median			
Mode			

PART 1 - STATISTICAL ANALYSIS - Comparing Dispersions/Spread

- 7. WHY do we calculate measures of dispersion/spread when analyzing data sets? (Give at least 2 reasons)
- 8. Prepare a 5 number summary for each of the three throwers.

C/T	C/T Thrower #1	C/T Thrower #2	C/T Thrower #3
Minimum			
Q ₁			
Q ₂			
Q ₃			
Maximum			
Range			

9. Record AT LEAST 3 OBSERVATIONS from your measures of dispersion/spread.

PART 1 - STATISTICAL ANALYSIS - Comparisons

10. Prepare a graph that allows you to **compare** the data from the three throwers on the SAME graph.

11. Record AT LEAST 3 OBSERVATIONS from your graph.

PART 2 – DECISION MAKING & JUSTIFYING

Now that you have completed the required statistical analysis, you must make a decision as to which thrower is the best. First, you must explain what it means to "be the best" thrower. Secondly, you will now tell me who is the best and WHY you think that they are the best (your reasoning must be STATISTICALLY based!)

- 12. Explain your ideas about what it means to be the BEST shot put thrower. Present at least 3 ideas.
- 13. In your opinion, the BEST thrower is _____.
- 14. Explain WHY you selected this thrower as the best. Your explanation MUST be STATISTICALLY based, with references to some of your analysis results from PART 1.

Part 1 - Skills Review - HOMEWORK

- 1. Here are the weights of 7 students in class (measured in pounds): 150, 160, 173, 155, 160, 175, 170. Calculate the mean, median, mode and range of student weights.
- 2. Samuel is trying to determine the average height of high school male students. Because he is on the basketball team, he uses the heights of the 14 players on the team, which are given below in inches.

69, 70, 72, 72, 74, 74, 74, 75, 76, 76, 76, 77, 77, 82

- a. Calculate the mean, median, mode and range for this data set. Round any non-integer answers to the nearest tenth.
- b. Is the data set above a fair sample to use to determine the average height of high school male students? Explain your answer.
- c. Determine the values of Q_1 (first quartile) and Q_3 (third quartile) and hence, determine the interquartile range of this data set.
- 3. The high temperatures for a 7-day week during December in Chicago were:

29°, 31°, 28°, 32°, 29°, 27°, and 55°.

- a. Find the mean high temperature for the week.
- b. In this example, is the mean temperature a good representation of the data?
- c. Find the median temperature for Chicago during this week.
- d. Which measure of central tendency is a better indicator of the "central tendency" of temperatures in Chicago this week.
- 4. Here are some simple die (number cube or dice) probability questions. How probable it is that you:
 - a. Roll an even number on one roll of a die.
 - b. Roll an odd number on one roll of a die.
 - c. Roll a prime number on one roll of a die.
 - d. Roll an odd prime number on one roll of a die.
 - e. Roll an odd prime number greater than 4 on one roll of a die.
 - f. Roll an even prime number greater than 4 on one roll of a die.
 - g. Roll an odd prime number or a multiple of 2 on one roll of a die.
 - h. Roll an odd prime number and a multiple of 2 on one roll of a die.

5. To help you understand how to work with Venn diagrams, answer the following questions:

Numbers are written on cards, the cards are shuffled and one is selected at random. The outcomes for this experiment and events A and B are shown in the Venn diagram. Complete the following.

- a The sample space S = {____}} (outcomes inside the rectangle)
- b Event A = { _____} (outcomes inside the circle A)
- c Event 'not $A' = \{ ___ \}$ (outcomes outside the circle A)
- d Event $B = \{ _ \}$ (outcomes inside circle B)
- e Event 'not $B' = \{ ___ \}$ (outcomes outside circle B)
- f Event 'A and B' = $\{__\}$ (outcomes in the intersection of the circles)
- g i Event 'A or B or both' = { ____} (outcomes within both circles, including the intersection)
- ii Is this an 'inclusive or' event or an 'exclusive or' event?
- h i Event 'A or B but not both' = { _____} (outcomes within both circles, excluding the intersection) ii Is this an 'inclusive or' event or an 'exclusive or' event?
- i Event 'neither A nor $B' = \{ __ \}$ (outcomes outside both circles)
- 6. At a wedding, the guests may have ice cream or custard with their dessert. The Venn diagram shows information about the choices the guests made.
 - a. How many guests had custard?
 - b. How many guests had ice cream and custard?
 - c. How many guests went to the wedding?
 - d. How probable is it that a randomly selected guest had ice cream for dessert?
 - e. How probable is it that a randomly selected guest had neither ice cream nor custard?
- 7. A set of numbers is given as {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16}. Let Set A be the multiples of 3 and let set B be the multiples of 5.
 - a. Complete the Venn diagram
 - b. One of the numbers is selected at random. Write down
 - The probability that you picked a number from set A i.
 - ii. The probability that you picked a number from set B
 - $P(A \cap B)$ iii.





