

Math 10 - Homework 1 answers

1&2 Identify the following data by type (categorical, discrete, continuous) and level (nominal, ordinal, interval, ratio)

- Number of tickets sold at a rock concert **discrete, ratio**
- Make of automobile. **categorical, nominal**
- Age of a fossil. **continuous, ratio**
- Temperature of a nuclear power plant core reactor. **continuous, interval**
- Number of students who transfer to private colleges. **discrete, ratio**
- Cost per unit at a state University. **continuous, ratio**
- Letter grade on an English essay. **categorical, ordinal**

3. 1038 Americans were asked, "What is your favorite sport to watch?" The results were summarized into a pie graph.

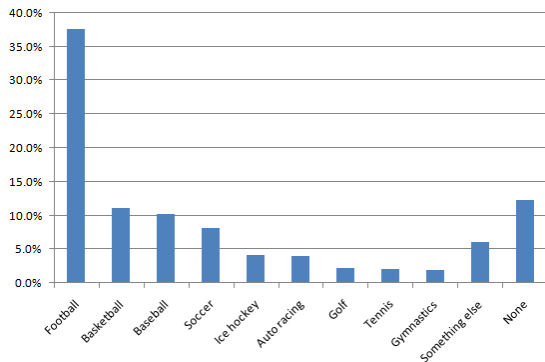
a. Interpret the pie graph.

Many answers possible. #1 was football.

b. Do you think a different graph would have a clearer way to show this data? Explain.

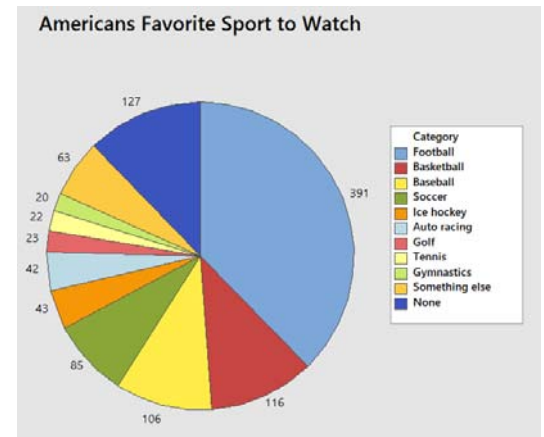
Bar Graph would be easier to read.

c. Using the same data create a bar graph. Instead of labeling each bar with counts, use percentages.



d. Compare the bar graph to the pie graph. In your opinion, which of these two graphs better explains the data?

Many answers possible, I prefer bar graph.



4. The two frequency histograms represent the ages of 78 Male US Senators and 22 Female US Senators. Ages were evaluated on October 20, 2017.

a. Estimate the center of each graph. Does there seem to be a difference in average age due to gender in the US Senate?
Men about 65, women about 60. Women seem younger.

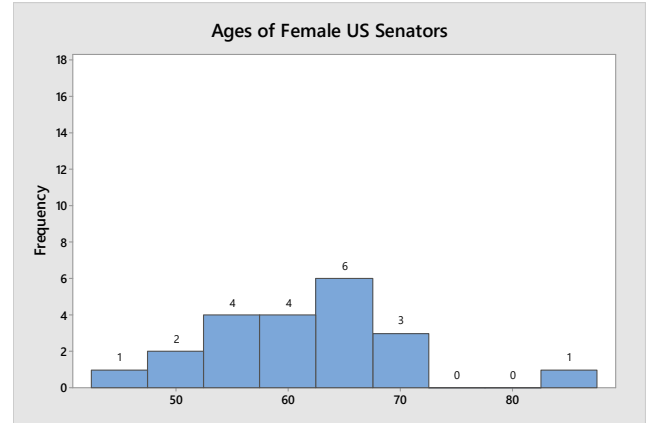
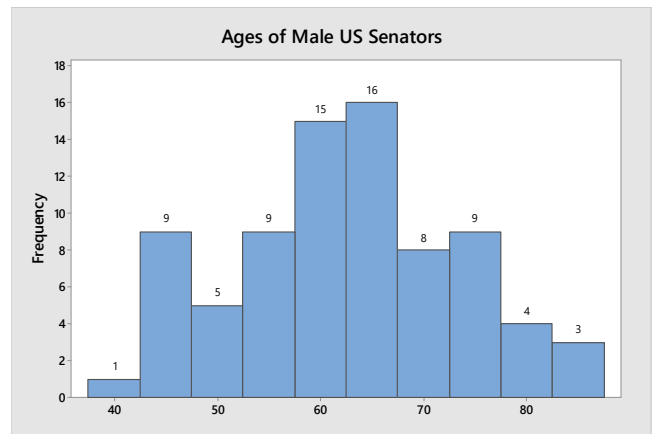
b. Estimate the range of each graph. there seem to be a difference in age spread due to gender in US Senate
Range is about 50 years for each gender, however, only one women is very old.

c. Is there a difference in shape between the two graphs?

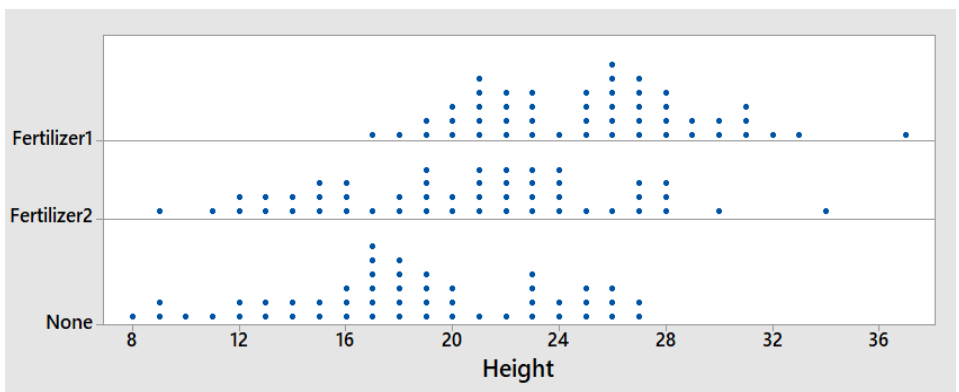
Both seem symmetric

d. Senator Diane Feinstein of California, who is 84 years old, represents an outlier among the females. Would your answers to parts a, b or c change if Senator Feinstein were removed from the data? Explain.

Range would decrease for women to about 35. Shape would be skewed negative for women.



5. An experiment was conducted on string bean plants. The plants were broken into three groups. The first group was given Fertilizer 1, the second group was given Fertilizer 2, and the third group was given no fertilizer. After 2 months, the heights in inches were measured with results shown in the dot plot. From the dot plots, describe the center, spread, shape and unusual features of each group, and then make an overall statement about the fertilizers.



	Fertilizer 1	Fertilizer 2	None
Center	26 inches	23 inches	18 inches
Spread (Range)	20 inches	26 inches	19 inches
Shape	Symmetric/skewed positive	symmetric	Skewed negative

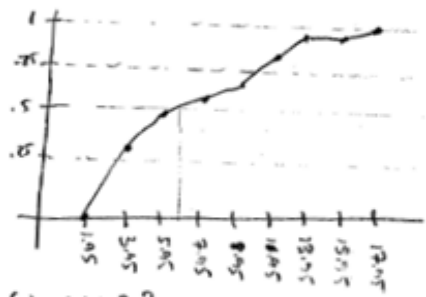
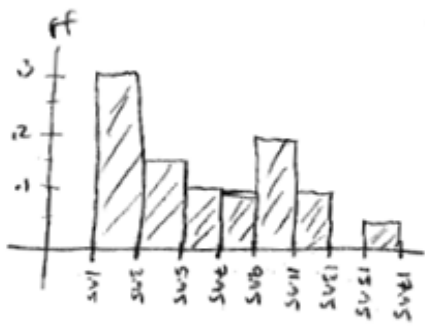
Tallest plants with fertilizer 1, most spread with fertilizer 2.

6. The February 10, 2017 Nielsen ratings of 20 TV programs shown on commercial television, all starting between 8 PM and 10 PM, are given below:

2.1 2.3 2.5 2.8 2.8 3.6 4.4 4.5 5.7 7.6
 7.6 8.1 8.7 10.0 10.2 10.7 11.8 13.0 13.6 17.3

- Graph a stem and leaf plot with the tens and ones units making up the stem and the tenths unit being the leaf. (for example, 2.1 would have a stem of 2 and a leaf of 1)
- Group the data into intervals of width 2, starting with the 1st interval at 2, and obtain the frequency of each of the intervals
- Graphically depict the grouped frequency distribution in part b by a histogram.
- Obtain the relative frequency, cumulative frequency and cumulative relative frequency for the intervals in part b.
- Construct an ogive of the data. Estimate the median and quartiles.
- Obtain the sample mean and median. Do you believe that the data is symmetric, right-skewed or left skewed?
- Determine the sample variance and standard deviation.
- Assuming the data are bell shaped, between which two numbers would you expect to find 68% of the data?

Stem	Leaf	freq	rf	crf	class interval	cf
2	1 3 5 8 8	6	.30	.3	1.5 - 3.5	6
4	4 5	3	.15	.45	3.5 - 5.5	9
5	7	2	.10	.55	:	11
6		2	.10	.65	:	13
7	6 6	2	.10	.75	:	15
8	1 7	4	.20	.95	:	19
9		2	.10	.95	:	19
10	0 2 7	3	.15	1.00	15.5 - 17.5	20
11	8	1	.05	1.00		
12						
13	0 6					
14						
15						
16						
17						
18						
19						
20						



(f) $\bar{x} = 7.465$ $med = 7.6$
 $med > \bar{x}$ which suggests left skewed

(e) $Q1 = 3.7$
 $med = 7$
 $Q3 = 10$

(g) $s^2 = 19.474$ $s = 4.413$

(h) 7.465 ± 4.413
 $(3.052, 11.878)$

7. The following data represents recovery time for 16 patients (arranged in a table to help you out)

count	Days (X)	$X - \bar{X}$	$(X - \bar{X})^2$	Z Score
#1	2	-4	16	-1.27
#2	3	-3	9	-0.96
#3	4	-2	4	-0.64
#4	4	-2	4	-0.64
#5	5	-1	1	-0.32
#6	5	-1	1	-0.32
#7	5	-1	1	-0.32
#8	5	-1	1	-0.32
#9	5	-1	1	-0.32
#10	6	0	0	0.00
#11	6	0	0	0.00
#12	7	1	1	0.32
#13	7	1	1	0.32
#14	8	2	4	0.64
#15	8	2	4	0.64
#16	16	10	100	3.18
Totals	96		148	

a. Calculate the sample mean and median

XBar=96/6 = 6 Median 5

b. Use the table to calculate the variance and standard deviation.

$s^2 = 148/(16-1) = 9.867$ $s = 3.141$

c. Use the range of the data to see if the standard deviation makes sense. (Range should be between 3 and 6 standard deviations)

Range/s=4.5, so the answer makes sense

d. Using the empirical rule between what two numbers should you expect to see 68% of the data? 95% of the data? 99.7% of the data?

Within 1s Within 2s Within 3s

2.86 -0.28 -3.42

9.14 12.28 15.42

e. Calculate the Z-score for each observation. Do you think any of these data are outliers?

See above – 16 seems to be an outlier more than 3 standard deviations above the mean.