3.3 Measures of Spread

In the previous section, you studied measures of *central* tendency. It is important to note that distributions of continuous data can also be characterized and compared by considering their measures of *spread*.

PART A-What is the Interquartile Range?

To define these measures of spread, let's consider a familiar context—a class set of MCR 3U midterm marks. The marks are as follows:

90	90
87	74
80	77
75	90
83	67
75	77
78	92
76	80

Using this set of data, determine each of the following:

1. the range
2. the quartiles
3. the Interquartile range, and
4. Create a box-and-whisker plot (e.g., on p4 of this document) to represent the spread of data using *Fathom*. Copy the graphic as a picture and then paste into the space provided. Label the graphic with the following: *min, max, Q1, Q2, Q3, range, IQR.*

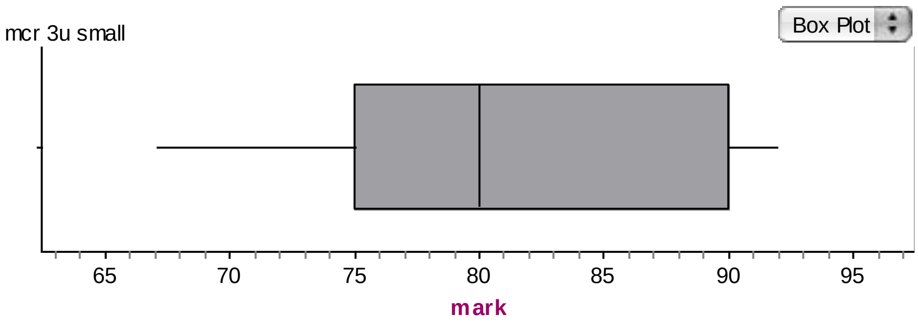
Solution:

1.

2.

3.

4.



Now consider the set of marks from another semester's MCR 3U class. Repeat steps 1 to 4 from above.

Solution:

79	75
79	77
76	79
78	82
72	81
85	72
79	79
74	77
74	81
76	77
81	63
63	69
86	67
72	84
78	
80	

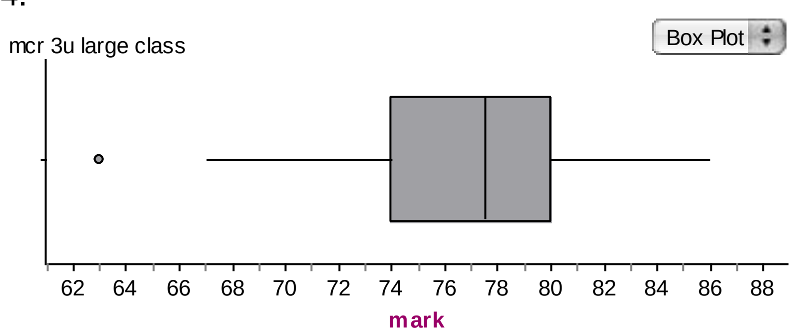
<u>Repeat steps 1 to 4 applied to first data set</u>

1.

2.

3.

4.



Task 1: Compare both 3U classes on the basis of their measures of spread. Write your ideas in the space provided below.

Part B—A Standard Measure of Spread

Problem 1: Let's say that both classes had equal measures of spread. How then could we compare them?

- The measure of spread we use should take into consideration the distance each piece of data is from the mean—these are called **deviations**.
    - Values that are less than the mean ($x \prec \bar{x}$) have negative deviations ($x - \bar{x}$ < 0); those greater than the mean, positive ($x - \bar{x}$ > 0).
    - As the deviations from the mean get larger, so does the spread of the data.
    - If you add up all the deviations for a data set, they will cancel out. Statisticians have shown that a root-mean-square quantity is a useful measure of spread. The **standard deviation** is the square root of the mean of the squares of the deviations.
    - The lowercase Greek letter sigma, σ, represents the standard deviation.

Task 2: Calculate the standard deviation for both classes. Compare these values and discuss the significance of the results.

Solution:

To move the task forward, I have calculated the standard deviation for the larger set of data. The values are as follows:

Mean = 76% Standard deviation: 5.72%

Continue the solution by calculating the standard deviation for the smaller set of data. Note: It is sometimes helpful to organize your data in table like the one provided below.

Mark	$x - \bar{x}$	$(x - \bar{x})^2$
90		
87		
80		
75		
83		
75		
78		
76		
90		
74		
77		
90		
67		
77		
92		
80		
	$\sum (x - \bar{x})^2 =$	

Standard deviation, σ =

Mean, $\bar{x}$ =

As mentioned before, the smaller the standard deviation or the variance, the more compact the data set.

To remain within the reach of this course, we will calculate the standard deviation for both **grouped** and **ungrouped population** data. The formulas are as follows:

Standard deviation for Ungrouped data:	Standard deviation for Grouped data:
$\sigma = \sqrt{\frac{\sum (x_i - x)^2}{n}}$	$\sigma = \sqrt{\frac{\sum f_i (x_i - x)^2}{n}}$

PART C-Consolidate Your Understanding

E.g. 1., For one week, the Math twins—Iluv and Uluv—record how many questions they successfully answer each day during math class. They find that Iluv answers 16, 17, 16, 15, and 14 questions. Uluv answers 14, 18, 22, 16, and 8 questions. Use statistical measures to determine which worker is consistent in answering questions. (For reference, see E.g., 2, p165)

x	$x-\bar{x}$	$(x-\bar{x})^2$
	$\sum(x-\bar{x})^2 =$	

x	$x-\bar{x}$	$(x-\bar{x})^2$
	$\sum(x-\bar{x})^2 =$	

E.g., 2, A railway line gives out bags of peanuts to its travelers, and each bag does not always contain the same number of peanuts. The following table represents a sample of 31 bags showing the number of peanuts per bag. Calculate the standard deviation for the sample of *grouped data* and describe its significance.

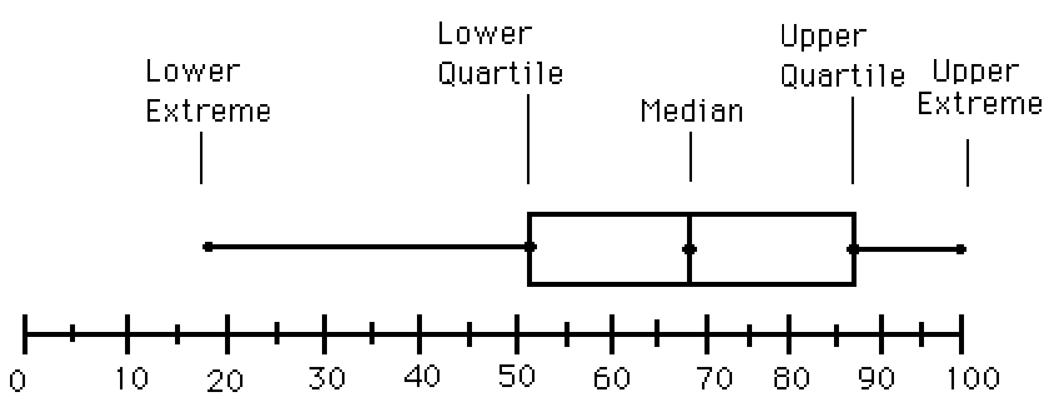
Number of Peanuts	28	29	30	31	32	33
Frequency	2	5	10	9	4	1

x	$x-\bar{x}$	$f(x-\bar{x})^2$
	$\sum f(x-\bar{x})^2 =$	

Part D—Constructing Visual Displays for Measures of Spread

Task 3:

- Using the statistical software program, *Fathom*, complete the assignment provided by your teacher—"Box Plots and Transformations"—up to and including Q13 on p27.
  - The goals of the assignment are to have you learn how to construct and interpret percentile plots, box-and-whisker plots (measures of spread; example below) and to introduce you to the idea of the number of standard deviations that seem to capture the majority of a distribution's data.
    - The assignment also requires that you use a pre-made *Fathom* file called "Census at School".
      - These files, along with the instructions for your assigned questions, are posted to the course website, Unit 3 sub-page.



Part E—Your Next Opportunity to Learn

Task 4:

Complete the following practice problems:

- p168 #1 (also create a box & whisker plot for this problem), 4, 6, 9, 10, 13
- Fathom will be sent home with students such that they have the capability of completing assignment and project work at home also.

Appendix

Part A Data Sets for Importing into Fathom

MCR 3U-Class 1 Data Set	MCR 3U-Class 2 Data Set
90	79
87	79
80	76
75	78
83	72
75	85
78	79
90	74
74	74
77	76
90	81
67	63
77	86
92	72
80	78
	80
	75
	77
	79
	82
	81
	72
	79
	77
	81
	77
	63
	69
	67
	84