

Study Session Week of 10/2

Objectives:

*you need
a calculator
and paper*

- I will examine and conclude what happens to shape, center & spread under linear transformations.
- I will apply that knowledge to free response and/or multiple choice questions.

Agenda:

- Together enter data.
- Perform some linear transformations and examine their effects.
- Practice in Free Response and/or Multiple Choice questions.

Data Transformations – step 1

- We will use the following data (just numbers) based on self-reports from the [National Survey of Student Engagement](https://thesocietypages.org/socimages/2012/05/30/how-much-do-students-study/) (NSSE) in a spreadsheet in your calculator

1006

In hours per week

Architecture	23.7	Accounting	15.1
Chemical engineering	21.6	History	15
Physics	19.7	Computer science	14.7
Chemistry	18.4	Political science	14.6
Art	18.1	Economics	14.4
Nursing	18	Psychology	13.9
Music	17.5	Sociology	13.8
Biology	16.7	Finance	13.3
Mathematics	16.4	Business administration	13.2
Philosophy	16.2	Journalism	12.8
Theater	16	Communications	12.5
Anthropology	16	Marketing	12.1
English	15.9	Physical education	11.8
Elementary education	15.2	Leisure studies	11.1
		Speech	10.8

Note: Majors shown are a sampling from a list of 85.

<https://thesocietypages.org/socimages/2012/05/30/how-much-do-students-study/>

Data Transformations – step 2

- Create the following table on your paper

1006

Measures of Center		Hours	Minutes (x 60)	Self-Report (- 30)	Rule 60x - 30
Mean	\bar{x}	15.4	927.9	897.9	both x and -
Median	Med	15.1	906	876	
Mode		16	960	930	
Measures of Spread		Hours	Minutes (x 60)	Self-Report (- 30)	Rule 60x - 30
Standard Deviation	$\frac{s_x}{\sqrt{6_x}}$	2.9	177.08	177.08	only multiply
IQR	$Q3 - Q1$	3.85	231	231	
Range	Max - Min	12.9	774	774	

- Complete the summary statistics for the hours \rightarrow 1 var Stats

Shape doesn't change except for "spread"

Data Transformations – step 3

- Set up column “b” to convert hours to minutes – label that column “minutes” and use the formula “=60*a”
 - Find those summary statistics
- Set up column “c” that adjusts the original data for self-reporting by subtracting ½ hour (.5) – label that column “self” and use the formula “=a-0.5”

Measures of Center	Hours	Minutes (x 60)	Self-Report (- .5)	Rule 60x - 30
Mean				
Median				
Mode				
Measures of Spread	Hours	Minutes (x 60)	Self-Report (- .5)	Rule 60x - 30
Standard Deviation				
IQR				
Range				

Data Transformations – step 4

- Set up column “d” to view the entire rule – label that “rule” and use the formula “=60*a – 30”
 - Find those summary statistics
- Create 4 box plots (one for each column) and sketch them on your paper side-by-side with the same scale

Measures of Center	Hours	Minutes (x 60)	Self-Report (- .5)	Rule 60x - 30
Mean				
Median				
Mode				
Measures of Spread	Hours	Minutes (x 60)	Self-Report (- .5)	Rule 60x - 30
Standard Deviation				
IQR				
Range				

Data Transformations – step 5

- Write a conclusion about the effects of multiplying (dividing too) and subtracting (adding too) on shape, center & spread.

Hours \rightarrow $Q_3 = 17.1$
 $IQR = 3.85 \Rightarrow 22.8$

$$17.1 + 1.5(3.85)$$

$Q_3 + 1.5 IQR$
 $Q_1 - 1.5 IQR$

What about extremes?

- *outlier?* I have seen studies that electrical engineers study about 30 hours per week, put that value into the data set and examine how your summary statistics changed.

- Add this information to your conclusion.



Suppose that the distribution of a set of scores has a mean of 47 and a standard deviation of 14. If 4 is added to each score, what will be the mean and the standard deviation of the distribution of new scores?

	<u>Mean</u>	<u>Standard Deviation</u>
(A)	51	14
(B)	51	18
(C)	47	14
(D)	47	16
(E)	47	18

51

A company wanted to determine the health care costs of its employees. A sample of 25 employees were interviewed and their medical expenses for the previous year were determined. Later the company discovered that the highest medical expense in the sample was mistakenly recorded as 10 times the actual amount. However, after correcting the error, the corrected amount was still greater than or equal to any other medical expense in the sample. Which of the following sample statistics must have remained the same after the correction was made?

- ~~(A)~~ Mean
- (B) Median
- ~~(C)~~ Mode
- ~~(D)~~ Range
- ~~(E)~~ Variance = σ^2