

Week 5

Activity 1:

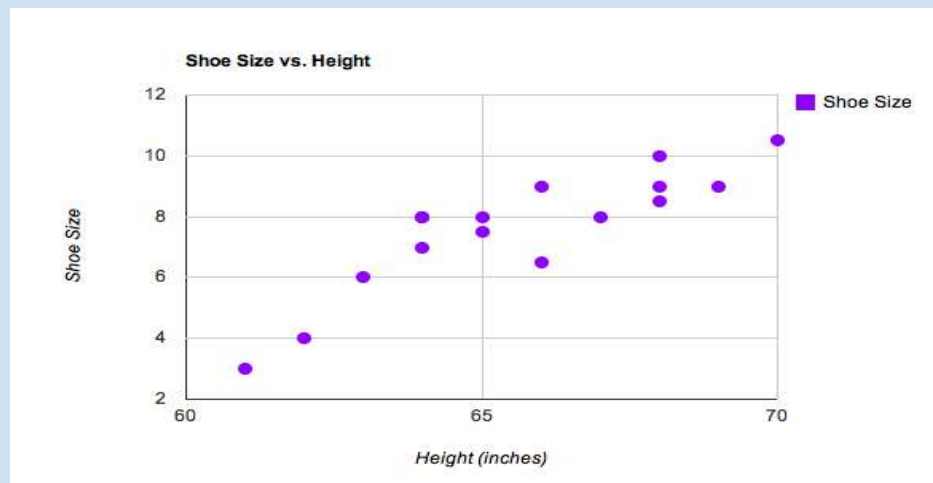
Students study the linear association between variables by analyzing two sets of data. They learn to **analyze data** in a spreadsheet as well as in a graph, and to **generalize patterns** of positive and negative linear association based on the shape of the graph. Students also learn how to **represent data** in a spreadsheet and relate the “slope” of their graphs to the association between two sets of data.

	A	B	C
1	Name	Height	Shoe Size
2	Sasha	63	6
3	Hector	68	9
4	Kayla	67	8
5	Adriana	65	7
6	Sean	64	7
7	Kay	61	3
8	Hillary	69	9
9	Thomas	68	9
10	Luis	64	8
11	Mayra	66	6
12	Michael	68	10
13	Tonya	64	8
14	Philip	65	8
15	Raymond	66	9
16	Jalen	70	10
17	Samantha	62	4

1: Create a scatter plot and analyze data

Activity:

Have students create a scatter plot chart to visualize the collected data like the one below. Ask them the following



questions:

Q1: What piece of information can we get from the spreadsheet that we cannot from the graph?

Q2: Describe any advantages of representing the data graphically.

Q3: Describe any patterns in the trends between height and shoe size.

Q4: What term do we use to describe trends represented by graphs with “positive slope”?

2: Comparing data patterns

Activity Overview: In this activity, students will represent data in an additional graph. Students will compare two patterns and their graphical representations to strengthen their understanding of positive and negative association.

Breed	Lifespan (years)	Male Weight (lbs)
Afghan Hound	12	60
Airedale Terrier	11	60
American Stafford. Terrier	12	62
Basset Hound	13	42
Beagle	13	23
Bearded Collie	12	42
Bedlington Terrier	14	20
Bernese Mountain Dog	7	92
Border Collie	13	37
Border Terrier	13	14
Boxer	10	65
Bull Terrier	13	62
Bullmastiff	9	120
Cairn Terrier	13	16
Cav. King Charles Spaniel	11	14
Chihuahua	13	4
Chow Chow	14	57
Cocker Spaniel	13	23
Dachshund	12	15
Dalmatian	13	68
Doberman Pinscher	10	78
English Cocker Spaniel	12	33
English Setter	11	67
English Springer Spaniel	13	50
English Toy Spaniel	10	11
Flat-Coated Retriever	10	65
German Shepherd	10	80
German Shorthaired Pointer	12	63
Golden Retrievers	12	70
Gordon Setter	11	68

Great Dane	8	160
Greyhound	13	68
Irish Setter	12	70
Irish Wolfhound	6	110
Jack Russell Terrier	14	16
Labrador Retriever	13	72
Miniature Poodle	15	16
Norfolk Terrier	10	11
Old English Sheepdog	12	62
Pekingese	13	8
Pomeranian	15	5
Rhodesian Ridgeback	9	110
Rottweiler	10	115
Samoyed	11	55
Scottish Deerhound	10	100
Scottish Terrier	12	21
Shetland Sheepdog	13	18
Shiba Inu	14	25
Shih Tzu	13	13
Siberian Husky	14	47
Soft Coated Wheaten Terrier	13	40
Staffordshire Bull Terrier	14	35
Standard Poodle	12	40
Tibetan Terrier	14	24
Toy Poodle	14	7
Vizsla	12	53
Weimaraner	10	63
Welsh Springer Spaniel	12	40
West Highland White Terrier	13	18
Wire Fox Terrier	13	18

Activity:

Present students with the [sample data](#) on the average weight of various breeds of dogs compared to the average lifespan of those breeds. Ask them to repeat the process of creating a scatter plot from [Activity 1](#). Ask your students the following questions:

Q1: Compare the trend in the graph in weight vs. lifespan to that in the graph of shoe size vs. height. Describe the difference between the two trends.

Q2: Do bigger breeds tend to live longer or shorter lives than smaller breeds?

Activity:

Ask your students the following questions to encourage thinking about the relationship between hard data and graphical representations:

Q3: If we draw a straight line that passes through the “middle” of the plotted data, would the line have a positive slope or a negative slope?

Q4: What term do we use to describe trends represented by graphs with negative slope?

Application and Modeling of Standard Deviation

Lesson Overview

When looking for patterns in data, it is helpful to look at the **central tendency** to see if patterns can be discovered. Students will be able to examine graphs of data using the measures of **central tendency** to determine when to use each in order to get a true overview of the data. Students will use the **data collection** to gather data, **decomposition** to regroup collected data, **abstraction** and **data analysis** to make real world insights based on the calculation of statistics about a population or data set. Finally they will **simulate** a dice-throwing game and alter the **algorithm design** to reflect changes to the game.

Activity Overview: Grades alone can be given to the students or they could be asked to generate the grades. Then, the students can be asked to group the data and find the average and standard deviation, similar to how it is done below:

Grades	Results	Totals
69	<60%	13
97	61-70%	1
37	71-80%	3
87	81-90%	6
79	91-100%	7
80		
88		
50		
85		
91		
20		
30		
99		
17		
81		
88	Average	60.83333333
97		
31	Standard Deviation	32.65688629
16		
92		

34		
80		
24		
0		
85		
94		
10		
94		
59		
11		

Activity 1: Measuring distribution (30 to 45 minutes)

Activity Overview: In this activity, students will analyze collected data to calculate various statistics and make meaningful insights about the data or population. Students **abstract** the data to make real world insights using **data analysis**. Students **decompose** the data into ranges to more easily determine the most effective drug. This activity engages students through student-teacher interaction.

Activity:

Have your students go through the following scenarios and questions:

1. **Standard deviation is used to show how spread out the data is from the mean, which is useful for understanding how closely related the data are.**
2. **In the sheet given above, a teacher is looking at the results of the most recent test.**
3. **The average in the class is 61%, but very few students actually received a 61%.**

Q1: What is the standard deviation for this data?

Q2: Would it be helpful to grade this assessment on a curve, that is, adjust the scores so that a 100% is based on the highest grade in the class?

4. The “Drug Trials” sheet given below gives data from a fictional drug trial to cure persistent hiccuping. The data records how many hours it took for the drug to cure the hiccups.

Drug 1	Results	Totals	Drug 2	Results	Totals	Drug 3	Results	Totals
4	0-1	1	0	0-1	34	2	0-1	12
5	2-3	17	0	2-3	31	6	2-3	64
5	4-5	57	5	4-5	34	3	4-5	0
5	6-7	25	4	6-7	1	6	6-7	24
6			1			2		
6			1			6		
5			3			3		
3			5			6		
4			2			6		
5			5			6		
6			3			2		
5	4.54		4		1.516	0		1.3872
5			3			2		
6	1.20349491066643		1		1.720465053	2		1.810497169

6			0			6		
6			2			3		
3			4			3		
4			3			6		
4			4			6		
3			4			2		
6			2			6		
6			5			3		
6			3			3		
6			2			6		
4			1			6		
6			3			6		
4			5			6		
3			0			2		
4			5			0		
5			1			3		
6			0			2		
5			3			2		
5			1			2		
5			4			3		
3			1			3		
5			3			6		
3			0			6		
5			2			3		
3			6			6		
2			3			6		
4			1			1		
6			4			6		
4			5			2		
4			4			6		
4			1			3		
6			5			3		
4			2			3		
4			2			2		
6			5			3		
5			0			2		
6			1			2		
4			5			2		
4			3			2		
5			5			3		
6			2			3		
4			1			2		
4			0			1		
4			2			2		
2			2			3		
3			4			3		
3			4			2		

5			2			0		
4			5			2		
6			5			3		
5			5			3		
2			4			1		
5			1			2		
5			4			0		
4			0			3		
3			0			3		
6			1			3		
4			3			6		
6			3			2		
6			5			6		
5			4			2		
0			0			1		
5			3			2		
4			5			3		
5			1			0		
5			3			3		
6			5			6		
6			3			6		
4			1			0		
3			1			2		
5			1			3		
4			1			0		
5			4			3		
5			3			3		
4			0			2		
3			2			2		
4			2			1		
5			2			3		
5			0			2		
2			3			3		
6			4			3		
5			4			3		
4			1			3		
3			1			2		
6			1			3		
4			5			2		

Q3: What is the mean and standard deviation for each drug?

A placebo (<http://wikipedia.org/wiki/Placebo>) is a pill that does nothing. It is given to some patients in a drug trial (the “control group”) to make sure that the drug has more effect than no treatment at all.

Q4: Which of the three drugs would likely be the placebo?

Q5: What might be an effective way to advertise Drug 2 over Drug 3? Most people do not understand “standard deviation,” so you will need to convey your recommendation in non-mathematical terms.