

Calculating Central Tendency & an Introduction to Variability

Student/Class Goal
Students will be able to make better decisions based on averages and variances.

Outcome (*lesson objective*)
Students will use mean, median and mode; calculate range, quartiles, inter-quartiles, variance and standard deviation for a data set.

Time Frame
6 hours

Standard *Use Math to Solve Problems and Communicate*

NRS EFL 6

Number Sense	Benchmarks	Geometry & Measurement	Benchmarks	Processes	Benchmarks
Words to numbers connection		Geometric figures		Word problems	6.26
Calculation	6.1	Coordinate system		Problem solving strategies	
Order of operations		Perimeter/area/volume formulas		Solutions analysis	
Compare/order numbers		Graphing two-dimensional figures		Calculator	6.29
Estimation		Measurement relationships		Math terminology/symbols	
Exponents/radical expressions		Pythagorean theorem		Logical progression	
Algebra & Patterns	Benchmarks	Measurement applications		Contextual situations	
Patterns/sequences		Measurement conversions		Mathematical material	6.33
Equations/expressions		Rounding		Logical terms	
Linear/nonlinear representations		Data Analysis & Probability	Benchmarks	Accuracy/precision	
Graphing		Data interpretation		Real-life applications	
Linear equations		Data displays construction		Independence/range/fluency	
Quadratic equations		Central tendency	6.23		
		Probabilities			
		Contextual probability			

Materials

white board, chalk board or smart board
paper and pencils, calculators
Students to fly away with paper airplane contest in the Twin Cities Media Article
Paper Airplane Contest Problem Statement Handout

Learner Prior Knowledge

Students will need to know the meanings of mean, median and mode and how to correctly calculate mean, median and mode (M.3.18); calculate exponents (M.3.5); and order of operations (M.4.3).

Instructional Activities

Step 1 - Discuss the differences in calculating average. Review the computation involved in determining mean, median and mode and when each is used.

- Mean** - add all of the numbers and divide by how many numbers you have.
- Median** - order the numbers from least to greatest then find the middle number.
- Mode** - determine the most frequent number.

Give some examples of the appropriate use of mean, median and mode and some pros and cons of each. For example: The median may be a better indicator if a set of numbers has an outlier. An **outlier** is an extreme value that differs greatly from the other values.

The mean may be a better indicator if a set of numbers is large and does not include outliers. Examples of when mode is best used would be when we are looking for the most common eye color in a group of people, the most popular breed of dog in the country or the song that was downloaded the most. However, if all numbers in a set occur the same number of times, there is no mode.

Step 2 - Using age, weight, income etc. as examples of data, work several problems for students so they can see the different outcomes of mean, median and mode on the same data.

Example Eight children have the following weight in pounds: 102, 114, 98, 130, 111, 94, 80, 98. Find the mean, median and mode.

$$\text{Mean} = \frac{102+114+98+130+111+94+80+98}{8}$$

$$= \frac{827}{8} = 103.375$$

Median = 80, 94, 98, 98, 102, 111, 114, 130

$$= \frac{98 + 102}{2} = 100$$

To find the median, we must put the children's weight in order from least to greatest; then, because there is an even number of children, we must find the mean of the middle two numbers.

Mode = 98 It is the only number that occurs more than one time.

Step 3 - Using an example in Step 2, introduce the concept of range in a set of numbers. **Range** is the difference between the largest and smallest number in a set.

Example The set of numbers above is 80, 94, 98, 98, 102, 111, 114, 130
The range = 50 (130 – 80 = 50)

Use several other examples to demonstrate range in a set of numbers and work them as a class.

Example Gas prices at area stations were: \$3.59, \$3.64, \$3.69, \$3.75, \$3.83, \$3.89.
The range in price is \$3.89 - \$3.59 = \$0.30. We say the gas price had a range of \$0.30.

Step 4 - Do the following examples with student participation.

1. 19, 3, 10, 7, 19, 15, 80, 11, 4, 8, 16, 17

For this set of data, find the (a) range, (b) mean, (c) median, (d) mode, (e) outliers (f) which measure of central tendency best describes this data? Have students explain why they picked their answer.

Solution: First, put the data into ascending order: 3, 4, 7, 8, 10, 11, 15, 16, 17, 19, 19, 80

- (a) Range = 80 – 3 = 77
- (b) Mean = 209/12 = 17.4, rounded to the nearest tenth
- (c) Median = (11 + 15)/2 = 13 (even amount of numbers, mean of two middle numbers)
- (d) Mode = 19
- (e) Outlier = 80 because it is considerably larger than each of the other values
- (f) The median is the best measure of central tendency for this data. The mode, 19, is too large. The mean is too large because the outlier is so much bigger than the other values.

2. 44, 39, 11, 44, 42, 38, 39, 44, 41

For this set of data, find the (a) range, (b) mean, (c) median, (d) mode, (e) outliers (f) which measure of central tendency best describes this data? Have students explain why they picked their answer.

Solution: First put the data into ascending order: 11, 38, 39, 39, 41, 42, 44, 44, 44

- (a) Range = 44 – 11 = 33
- (b) Mean = 342/9 = 38
- (c) Median = 41
- (d) Mode = 44 since it occurs most often
- (e) Outlier = 11 because it is considerably smaller than the other numbers.
- (f) The Median is the best measure of central tendency because the mode is too large and the mean is too small.

Have students do this problem on their own or with a partner.

3. Bowling scores: 188, 194, 230, 203, 199, 231, 224, 210, 197, 220

For this set of data, find the (a) range, (b) mean, (c) median, (d) mode, (e) outliers (f) which measure of central tendency best describes this data? Have students explain why they picked their answer.

Solution: First put the data into ascending order: 188, 194, 197, 199, 203, 210, 220, 224, 230, 231

- (a) Range = $231 - 188 = 43$
- (b) Mean = $2096/10 = 209.6$
- (c) Median = $(203 + 210)/2 = 206.5$
- (d) Mode = there is no mode because each number occurs just once
- (e) Outlier = there is no outlier because the values are close together
- (f) Since there are no outliers, the mean is probably the best measure of central tendency. However, the median would also be a good measure of central tendency.

Step 5- Introduce the concept of inter-quartile range (IQR). **Quartiles** divide a rank-ordered data set into four equal parts. The values that divide each part are called the first, second, and third quartiles; and they are denoted by Q1, Q2, and Q3, respectively.

The **inter-quartile range** (IQR) is a measure of variability, based on dividing a data set into quartiles.

- Q1 is the "middle" value in the *first* half of the rank-ordered data set.
- Q2 is the median value in the set.
- Q3 is the "middle" value in the *second* half of the rank-ordered data set.

The inter-quartile range is equal to Q3 minus Q1.

Example Consider the series of numbers in Step 3: 3, 3, 4, 5, 6, 9, 10, 10, 13, 14, 15

 ↑ ↑ ↑
 lower median upper
 quartile quartile
 Q1 Q2 Q3

Q1 - Lower quartile (middle value of the lower half) = 4
Q2 - Median (middle value) = 9
Q3 - Upper quartile (middle value of the upper half) = 13
Inter-quartile range = $9 (Q3 - Q1) (13 - 4)$
Range = $12 (15 - 3)$

If there is an even number of data items, then use the average (mean) of the middle numbers.

Example Find the median, lower quartile, upper quartile, inter-quartile range and range of the following numbers.

17, 3, 40, 6, 31, 9, 1, 13, 16, 19, 7, 24

Solution: Make sure the data in ascending order:

1, 3, 6, 7, 9, 13, 16, 17, 19, 24, 31, 40
 ↑ ↑ ↑
 Q1 Q2 Q3

Q1 (lower quartile) = $(6 + 7) / 2 = 6.5$
Q2 (median) = $(13 + 16) / 2 = 14.5$
Q3 (upper quartile) = $(19 + 24) / 2 = 21.5$
Inter-quartile range = $Q3 - Q1 = 21.5 - 6.5 = 15$
Range = largest value - smallest value = $40 - 1 = 39$

Step 6 - Introduce the concept of variance and related terms. **Variance** is a measure of variability, which describes the spread or dispersion of a set of data. **Population** refers to an entire group of numbers, people or things (sometimes called a universe). A population can be *infinite* or *finite*. **Set** is a well-defined collection of objects. **Sample** is a subset of a population.

Example In a study of weights of students at Kent State University, the entire student body is the population which is a finite number. Rather than study every student in the population, perhaps we will work with a *sample* of only 200 students.

In a **population**, **variance** is the average squared deviation from the population mean, as defined by the following formula:

$$\sigma^2 = \frac{\sum (X_i - \mu)^2}{N}$$

where σ^2 is the population variance, μ is the population mean, X_i is the i th element from the population, and N is the number of elements in the population. The capital Greek letter sigma Σ is commonly used in mathematics to represent a summation of all the numbers in a grouping.

Example A population consists of four observations: {2, 3, 6, 7}. What is the variance?

Solution:

First – Computer the population mean: $\mu = (2 + 3 + 6 + 7) / 4 = 4.5$

Second – Put all know values into the formula for the variance of a population.

$$\begin{aligned}\sigma^2 &= \sum (X_i - \mu)^2 / N \\ \sigma^2 &= [(2 - 4.5)^2 + (3 - 4.5)^2 + (6 - 4.5)^2 + (7 - 4.5)^2] / 4 \\ \sigma^2 &= [(-2.5)^2 + (-1.5)^2 + (2.5)^2 + (3.5)^2] / 4 \\ \sigma^2 &= [6.25 + 2.25 + 6.25 + 12.25] / 4 \\ \sigma^2 &= 27 / 4 \\ \sigma^2 &= 6.75\end{aligned}$$

The variance of a **sample**, is defined by slightly different formula, and uses a slightly different notation:

$$s^2 = \frac{\sum (x_i - x)^2}{n - 1}$$

where s^2 is the sample variance, x is the sample mean, x_i is the i th element from the sample, and n is the number of elements in the sample. The sample variance can be considered an unbiased estimate of the true population variance. Therefore, use this formula if you need to estimate an unknown population variance, based on data from a sample.

Example A sample consists of four observations: {2, 3, 6, 7}. What is the variance?

Solution:

First – Computer the sample mean: $x = (2 + 3 + 6 + 7) / 4 = 4.5$

Second – Put all know values into the formula for the variance of a population.

$$\begin{aligned}s^2 &= \sum (x_i - x)^2 / (n - 1) \\ s^2 &= [(2 - 4.5)^2 + (3 - 4.5)^2 + (6 - 4.5)^2 + (7 - 4.5)^2] / (4 - 1) \\ s^2 &= [(-2.5)^2 + (-1.5)^2 + (2.5)^2 + (3.5)^2] / (4 - 1) \\ s^2 &= [6.25 + 2.25 + 6.25 + 12.25] / (4 - 1) \\ s^2 &= 27 / (4 - 1) \\ s^2 &= 9\end{aligned}$$

Step 7 - Introduce the concept of standard deviation. Once the variance of a set is determined, it is a simple matter to get the **standard deviation**. It is a numerical value used to indicate how widely individuals in a group vary. The standard deviation is the positive square root of the variance. Thus, the standard deviation of a **population** is:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum (X_i - \mu)^2 / N}$$

where σ is the population standard deviation, σ^2 is the population variance, μ is the population mean, X_i is the i th element from the population, and N is the number of elements in the population.

Example Using the numbers in Step 5, a population consists of four observations: {2, 3, 6, 7}. The variance is 4.5.

$$\sigma = \sqrt{\sigma^2}$$

$$\sigma = \sqrt{4.5}$$

$$\sigma = 2.12$$

It is important to distinguish between the standard deviation of a population and the standard deviation of a sample. They have different notation, and they are computed differently. The standard deviation of a population is denoted by σ and the standard deviation of a sample by s .

The standard deviation of a **sample** is:

$$s = \sqrt{s^2} = \sqrt{\sum (x_i - \bar{x})^2 / (n - 1)}$$

where s is the sample standard deviation, s^2 is the sample variance, \bar{x} is the sample mean, x_i is the i th element from the sample, and n is the number of elements in the sample. Using this equation, the standard deviation of the sample is an unbiased estimate of the standard deviation of the population. And finally, the standard deviation is equal to the square root of the variance.

Example Using the numbers in Step 6, a sample consists of four observations: {2, 3, 6, 7}. The variance is 4.5.

$$s = \sqrt{s^2}$$

$$s = \sqrt{9}$$

$$s = 3$$

Step 8 – This culminating activity has students determine how to create a fair judging scheme for a paper airplane contest while considering both the most accurate paper airplane and the best floater. Students individually read the media article, *Students to fly away with paper airplane contest in the Twin Cities*, to become familiar with the context of the problem. Then complete the questions that follow.

In teams of three or four, students are given the *Paper Airplane Contest Problem Statement* handout and work on the problem for 30-45 minutes. Students are given a sample of data that includes multiple flights of paper airplanes by three different pilots. Each team writes their solution in a letter or memo to the client. Then, each team presents their solution to the class. Whole class discussion is integrated with these presentations to discuss the different solutions, the statistics involved, and the effectiveness of the different solutions in meeting the needs of the client.

Assessment/Evidence (based on outcome)

Completion of problems
Paper airplane group project

Teacher Reflection/Lesson Evaluation

This lesson has not yet been field tested.

Next Steps

Additional central tendency problems from classroom or online resources can be given for students to practice until they have mastered the concepts.

Technology Integration

Average or Central Tendency: Arithmetic Mean, Median, and Mode

<http://www.khanacademy.org/math/algebra/ck12-algebra-1/v/average-or-central-tendency--arithmetic-mean--median--and-mode>

Range, Variance and Standard Deviation as Measures of Dispersion

<http://www.khanacademy.org/math/algebra/ck12-algebra-1/v/range--variance-and-standard-deviation-as-measures-of-dispersion>

Purposeful/Transparent

Students will use the concepts of central tendency to make decisions about a real-life situation. The teachers presents central tendency terminology and give examples of each as they work the problems together.

Contextual

The culminating activity asks students to work in teams to create judging criteria for a paper airplane contest, then to present their plan to a client.

Building Expertise

Students are introduced to central tendency terminology – mean, median, mode, range, quartile, population, variance sample, standard deviation – and then given opportunity to practice the concepts to gain mastery.

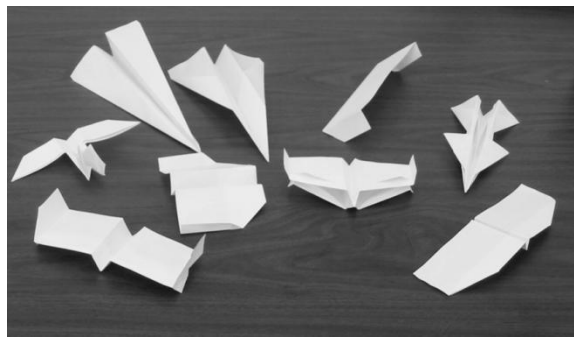
Students to fly away with paper airplane contest in the Twin Cities

St. Paul, MN –If you stop by Amy Frank’s eighth grade science classes this week you are likely to find a very busy group of kids. Ms. Frank’s students will follow in the footsteps of the Wright Brothers, engineers, and pilots as they design and fly paper airplanes in the Twin Cities Annual Paper Airplane Rodeo held in the Metrodome.

Frank’s students will be designing, creating and flying paper airplanes throughout the week. The students are learning how engineers work as they plan, create, test and redesign their paper airplanes. They won’t be using aluminum parts or jet engines for these planes. All they will need are pieces of paper – and a whole lot of imagination.

Students will need to design planes that are able to fly long distances as well as stay in the air for a long period of time. Each contestant will design a plane to try to win prizes in one of two categories: Best Floater and Most Accurate. Said Frank, “The contest is designed to require the students to be very thoughtful about making their planes, so students who want to enter the paper airplane contest must follow a few rules.” The rules are as follows: each plane must be made using a single sheet of 8.5”x11” paper. No cuts can be made in the paper; and, no tape, staples, glue or paper clips can be used to hold the plane together or to change the plane’s weight or balance. Also, each entry must qualify as being able to fly. For example, last year, a spitball and a dart were disqualified because they didn’t really fly - even though it was possible to throw them so that they stayed in the air for a long time. Parachutes and helicopters also were disqualified because they don’t go anywhere. For each throw, the judges will measure the time spent in the air, the distance the plane lands from the starting point, the distance the plane lands from the target, and the angle the plane lands from the target.

Because all paper airplanes are minutely different, it is difficult to make decisions about which plane is best. In order to make the competition as fair as possible, the judges are implementing two new processes for the contest. First, to minimize thrower advantages, the contest will have three neutral pilots to throw all planes in the contest. Second, the judges are designing a new scoring system to fairly judge the two winners.



“Some students are really getting into this contest – I’ve heard a couple who said they’re bringing in-flight refreshments, crash helmets, and parachutes,” said Frank. “It will be lots of fun.”

Questions to Get You Started

1. What are the categories for which the airplanes will be judged?
2. What types of measurements do you believe should be taken for each throw to fairly judge the contest?
3. How would you decide which airplane is the best floater?

4. How would you decide which airplane is the most accurate?
5. What are the judges doing differently this year than in years past? Why are they doing it?

Paper Airplane Contest

Problem Statement In past competitions, the judges have had problems deciding how to select a winner for each award (Most Accurate and Best Floater). They don't know what to consider from each path to determine who wins each award. Some sample data from a practice competition and a description of how measurements were made have been included. To make decisions about things like being the best floater, the judges want to be as objective as possible. This is because there usually are only small differences among the best paper airplanes and it seems unfair if different judges use different information or different formulas to calculate scores. So, this year, when the planes are flown, the judges want to use the same rules to calculate each score.

Write a brief 1 or 2-page letter to the judges of the paper airplane contest. Give them a rule or a formula which will allow them to use the kind of measurements that are given in Table 1 to decide which airplane is: (a) the most accurate flier, and (b) the best floater. Table 1 shows a sample of data that were collected for four planes last year. Three different pilots threw each of the four planes. This is because paper airplanes often fly differently when different pilots throw them. So, the judges want to "factor out" the effects due to pilots. They want the awards to be given to the best airplanes – regardless who flies them.

Use the data in Table 1 to show exactly how your rule or formula works – because the judges need to use your recommendation for planes that will be flown during the actual competition this year.

Note: The paper airplanes were thrown in a large 40-foot by 40-foot area in the arena. Each paper plane was thrown by a pilot who was standing at the point that is marked with by the letter S in the lower left hand side of each graph in Figure 1. So, this starting point is located at the point (0,0) on the graph. Similarly, the target is near the center of each graph, and it is marked with the letter X. So, the target is located at the point (25,25) on the graph.

In Table 1, the angles are measured in degrees. Positive angles are measured in a counter clockwise direction – starting from a line drawn from the lower left hand corner of the graphs to the upper right hand corner of the graphs (or starting from the point S and passing through the point X). Negative angles are measured in a clockwise direction starting from this same line.

Table 1: Information about Four Paper Airplanes Flown by Three Different Pilots

Plane	Flight	Pilot F				Pilot G				Pilot H			
		Distance from Start	Time in Flight	Distance To Target	Angle from Target	Distance from Start	Time in Flight	Distance To Target	Angle from Target	Distance from Start	Time in Flight	Distance To Target	Angle from Target
A	1	22.4	1.7	15.2	16	30.6	1.6	14.5	23	39	1.8	7.5	-10
	2	26.3	1.7	16.7	26	31.1	1.6	11.9	19	36.3	1.7	4.3	-6
	3	31.6	1.7	7.1	10	26.7	2.2	8.9	-4	35.9	2.2	9	-14
B	1	32.1	1.9	7.6	-11	35.9	1.9	14.3	-23	43.7	2.0	9.5	6
	2	42.2	2.0	9.2	-9	39	2.1	11.1	16	29	2.0	7.6	7
	3	27.2	2.1	10.2	-11	25.6	2.0	11.7	12	36.9	1.9	12.4	19
C	1	19.2	1.8	16.6	-8	42.9	2.0	9.8	9	35.1	1.6	2.8	4
	2	28.7	1.9	9.3	11	44.6	2.0	9.3	-1	37.2	2.2	2	-1
	3	23.6	2.1	17.3	-25	35.7	2.2	3.2	-5	42	2.1	9.8	10
D	1	28.1	1.5	8.9	9	37.2	2.1	20.2	-32	41.7	2.2	10.1	11
	2	31.6	1.6	14.8	-24	46.6	2.0	11.4	-2	48	1.9	14.1	-8
	3	39.3	2.3	9.1	12	34.7	1.8	22.2	-36	44.7	1.7	11.5	-9