Student-Centered Statistics for Algebra II

Kevin Michael Pierce Jr.
State University of New York (SUNY) Brockport

A thesis submitted to the Department of Education and Human Development of the State University of New York (SUNY) Brockport, in partial fulfillment of the requirements for the degree of Master of Science in Education
May 21, 2021

Abstract

Statistics is offered as a full course in some high schools while other schools focus on statistical topics integrated into courses such as Algebra. In the case of the later, teachers may feel unprepared to teach isolated statistical topics. This curriculum project was developed to provide support for teaching statistics as part of an Algebra II class. The four lessons were designed to provide an understanding of standard deviation, z-scores, and percentiles, which are then to be used in a culminating project based learning assignment. These lessons are aligned to the Common Core Standards for statistical reasoning.

Table of Contents

Introduction

One study shows that 18% of students entering an undergraduate program enter a science, technology, engineering, or mathematics (STEM) major (National Science Board, 2016). With many states only requiring Algebra I, Geometry, and Algebra II for high school graduation, many students will only briefly study statistics. For both STEM and non-STEM majors, statistics is an important area of study. The Common Core Standards, adopted and/or adapted by 41 states in 2010, thus designated an entire section of the mathematics standards to statistics (CCSS, 2021). In New York, these standards are incorporated within Algebra I and Algebra II. This curriculum project was designed to provide support instruction of standard deviation, z-score, and percentiles when statistics is part of an Algebra II course. The culminating problem based learning project was designed to create a student-centered classroom. The curriculum project is aligned to the common core standards in statistics and provides educators with additional resources to teach statistics. In this curriculum, instructors are encouraged to utilize student input and apply student generated data.

Literature Review

Current Climate of Statistics Education

Despite the rise in need and expectations in statistical literacy, teachers do not have the necessary preparation to support students in finding success in statistics (Lovett & Lee, 2017). Statistics poses a challenge to teachers and students alike with unique vocabulary, procedures, and techniques needed for success (Rossman et. al., 2007). If a teacher feels unprepared to teach a topic, their anxiety can be reflected in the attitudes of students (Fitzmaurice et. al., 2014). This may result in students who are not proficient in statistical reasoning and thus not prepared to navigate a data-driven world. The best practices for statistics education are formulating questions to solving authentic problems, collecting data, and formulating conclusions (Franklin et. al., 2007). This

curriculum project will do so through authentic problem solving, the project based learning model, technology, and assessment through performance tasks.

Authentic Problem Solving

Applying mathematics problems to a real-world context has become a standard teaching practice within most mathematics classrooms. Incorporating real world problems in mathematics has a positive impact on student understanding and engagement (Beswick, 2011). Statisticians and mathematicians look for patterns, but the meaning is grounded within context, which is heavily utilized in the statistics classroom (Franklin et. al., 2007). Students working in context can lead to a discovery of mathematical content in a realistic way. Beswick (2011) exemplifies this point in saying "there is also broad agreement that meaningful mathematics learning occurs in the process of engaging and grappling with not yet understood mathematics, rather than by learning procedures and then attempting to apply them" (p. 382). Project-based learning (PBL) is one way to integrate content and context in the statistics classroom.

Project Based Learning

One structure to support student learning in context is the project-based learning (PBL) model. PBL supports the engagement of students and their preparation in skills for the 21st century (Bell, 2010). Achieving engagement and developing 21st century skills can be seen as intangible, and one may argue that teachers who focus on these skills are neglecting mathematical conceptual fluency. However, the Project Based Learning structure has also shown benefits to student performance on standardized assessments (Craig & Marshall, 2019). The PBL model provides a more authentic environment to address the mathematical competencies needed for the 21st century. The integration of technology and assessments are a part of PBL.

Technology

Teaching mathematics in the 21st century requires technological literacy. The National Council of Teachers of Mathematics cites the effective use of technology as a pillar of mathematics education (2015). The use of calculators and other technological tools are cited as critical to the success in meeting standards that address what students are to learn in secondary statistics (CCSS, 2021). The use of technology has also been shown to benefit a student's perceived ability in statistics (Anastasiadou, 2012). Further, the National Council of Teachers of Mathematics states:

> It is essential that teachers and students have regular access to technologies that support and advance mathematical sense making, reasoning, problem solving, and communication. Effective teachers optimize the potential of technology to develop students' understanding, stimulate their interest, and increase their proficiency in mathematics. When teachers use technology strategically, they can provide greater access to mathematics for all students (2011, para. 1).

Technological tools allow for students to create representations of their data and present results in various ways. Further, correct visualization and representation of data is critical for instructional success in statistics education (Franklin et. al., 2017). Examples of effective technology use in statistics integrated into an Algebra II class include using a calculator to plot data, perform regression, and calculate measures of center and spread. The use of technology allows for students to spend less time in creating models, and more time in their analysis. Thus, technology is shown to be a useful pedagogical tool and develops technological savviness necessary for success in the future.

Assessment

Teachers who implement PBL in the classroom also need a different type of assessment. Thus, a summative assessment utilizing rubrics is a better suited assessment as opposed

to a traditional test (Bell, 2010). When designing a performance task, there are multiple key considerations: understanding expectations, authenticity, developing prompts, developing rubrics, and revision (Stoll & Schultz, 2019). There is a clear alignment between the best practices of statistics education and the utilization of a performance task. Beswick (2011) defines authenticity by being a common problem without a common solution. For a successful performance task, students must find or be given a problem which they have the ability to solve but will not repeat common procedures to do so. Authentic assessments designed to for use in PBL often require the use of rubric for scoring. By presenting a clear rubric to students, they can understand what they are expected to learn and know. Thus, performance tasks provide a unique opportunity for assessing student standards which melds perfectly with the PBL model and a unit on statistics.

Curriculum Design

This curriculum project focuses on the Common Core Algebra II topic of calculating and interpreting standard deviation. Algebra II presents the most complexity to students and teachers alike. One reason being the rigor of the Algebra II common core standards in statistics. In New York, students come into the Algebra II classroom with exposure to histograms, box plots, mean, median, range, and comparing data sets from introductory statistics being presented in Algebra I (CCSS, 2021). Thus, this curriculum project presents students with an opportunity to collect data and have them make observations to give an access point to the new material. The instructor should provide scaffolding during instruction of new content of standard deviation to provide the necessary supports. Lastly, the instructor should coach students on designing a research question which will adequately address standards while still engaging their interests.

The structure of these lessons will follow the best practices for PBL learning, including building on previous knowledge, providing necessary scaffolds, creating effective questions, utilizing

technology, and assessing effectively (Bell, 2010). When implemented well, this structure allows a teacher to work as a coach to guide students to understanding. Technology integrated into instruction well can support students to creatively explore data relationships (Bell, 2010). A PBL activity integrated into statistics can build on previously developed sociomathematical classroom norms and allow students to feel safe to explore, make mistakes, and ask questions.

The lessons in this curriculum were designed to be completed over four 40-minute classes to be followed by a performative assessment that will take a few classes to complete, depending on instructional goals. If teaching with a different schedule, these lessons may need to be modified. The lessons are numbered sequentially to show that they build on previous instruction. Therefore, no concept should be skipped. When following a PBL model, there can be different results depending on the students you work with and the data they collect. The following pacing guide provides suggested learning outcomes. The keys in the appendix provide suggestions, possible responses, as well as an example project.

Unit	**Standard Deviation, Z-Scores, and Percentiles**
Lesson 1	Objectives: • Access prior knowledge on creating histograms, mean, median, range, and inter-quartile range. • Describe the shape of data. • Define standard deviation and calculate using a calculator/software. Common Core Standards: • CCSS.MATH.CONTENT.HSS.ID.A.1 Represent data with plots on the real number line (dot plots, histograms, and box plots). • CCSS.MATH.CONTENT.HSS.ID.A.2 Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets. • CCSS.MATH.CONTENT.HSS.ID.A.4 Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there

	are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.
Lesson 2	Objectives: • Calculate standard deviation utilizing a calculator or computer software. • Estimate area under a curve given a number of standard deviations above or below the mean. • Compare two results given standard deviation. Common Core Standards: • CCSS.MATH.CONTENT.HSS.ID.A.4 Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.
Lesson 3	Objectives: • Define and calculate z-score. • Calculate a percentile using technology to analyze a point relative to other data. Common Core Standards: • CCSS.MATH.CONTENT.HSS.ID.A.4 Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.
Lesson 4	Objectives: • Calculate a value given a percentile using technology. • Design a research question which is meaningful and attainable. Common Core Standards: • CCSS.MATH.CONTENT.HSS.ID.A.4 Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.
PBL Assignment	Objectives: • Collect and plot data in response to a research prompt. • Calculate the mean and standard deviation of collected data. • Analyze a point within a data set using z-scores and percentiles. • Create a conclusion from data collected and calculations. Common Core Standards: • CCSS.MATH.CONTENT.HSS.ID.A.4 Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

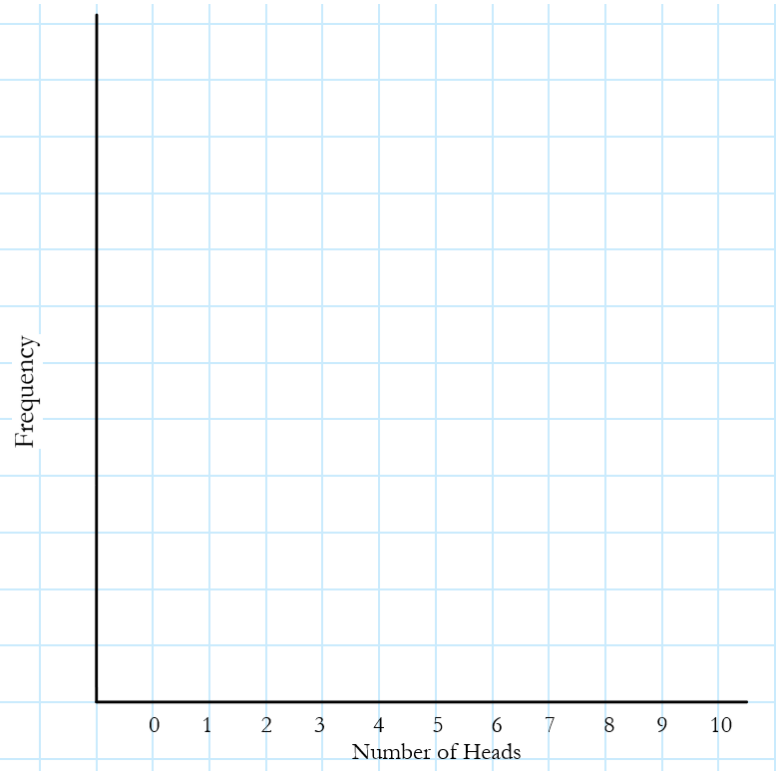
Lesson 1: Plotting Data, Measures of Shape, Center and Spread

Each person in class will flip a coin 10 times and count the number of times they get "heads". We will plot the classes data in the ***histogram*** below.

Histogram **–** a graphical display where data is displayed in ranges.

You can also use a random generator like the one found at https://www.random.org/coins/.

Num. of Heads	Frequency
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	



1. The shape of the data can help someone better understand our results. Let's describe the ***shape*** of the data from the choices below:

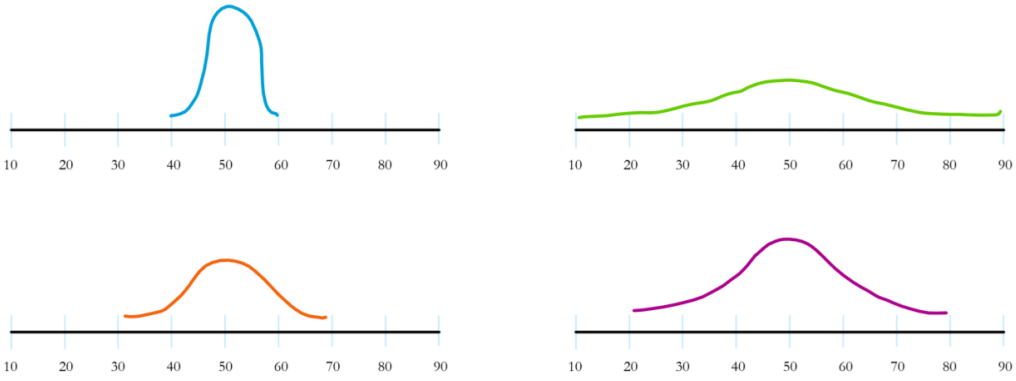
Oxford College, http://mathcenter.oxford.emory.edu/site/math117/shapeCenterAndSpread/

2. Beyond the shape of the data, what other ways can we describe the data we collected. In other words, what ways have we measured data in the past?

3. Let us sort those into two types of measures:

Measures of Center	Measures of Spread

4. Consider the four sketches of data sets below. Why can't we just give someone the Shape and a Measure of Center when describing data?



5. A new measure is called the ***standard deviation***.
   Have you heard of *deviation* or *deviant* before outside of math class? What does it mean?

   Deviation is difference from the mean, how far you are from the norm. A ***Standard Deviation*** is how much data typically differs from the mean. It is calculated by this formula, where μ (mu, pronounce myoo) is the mean.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

   Basically, it is the average difference from the average! Do not worry you will not need to calculate it by hand.

   Take out your calculator and let us input the class data from the first page.

6. Calculate the mean and standard deviation of the introductory data set.

7. Is the standard deviation a measure of center or measure of spread?

8. If a data set has a large standard deviation, what does that mean? A small standard deviation?

Lesson 2: Applying Standard Deviation

9. Using the coin flip data collected yesterday, find…
    a. The value which is one standard deviation above the mean.

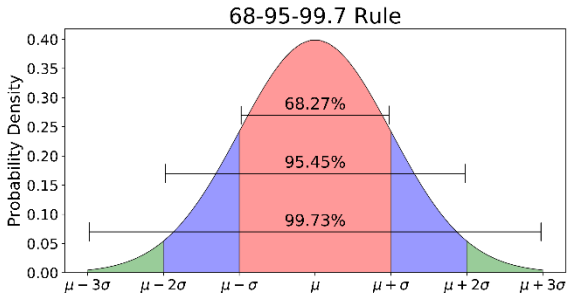
    b. The value which is one standard deviation below the mean.

    c. What percentage of the data was in between the values you found in part *a* and *b*?

    d. Based on how we described the standard deviation, what do the values you found in *a* and *b* tell us?

10. Follow procedures *a-d* above but now for 2 standard deviations.

11. Your work above should reflect what mathematicians have found over time. This is the power of the standard deviation! No matter what data you use, you will find…
    a. Most data (68%) of data falls within one standard deviation of the mean.
    b. 95% of data is within 2 standard deviations.
    c. Almost all (99.7%) of data is within 3 standard deviations.



When you know the standard deviation, you know the average difference from the mean.

Practice:

12. Miguel is the mean height for his age group. Carl is one standard deviation above the mean. What percentage of people in the age group are between Miguel and Carl?

13. Clarise was told her height is two standard deviations above the mean. What percentage of people in her age group is Clarise taller than?

14. Mean height for 10-year-olds is 54 inches with a standard deviation of 2.4. How would someone who is 51.6 inches tall compare to their peers?

15. You ask your teacher how you scored on the statistics test. They state that the mean score was 80 points, and the standard deviation was 6 points. You are told that your score was better than 97.5% of your peers. What did you score?

Lesson 3: Z-Scores and Percentiles

Last class we discovered what a standard deviation was and what it tells us about where data lies. To summarize…

~68% of data falls within 1 standard deviation of the mean

~95% of data falls within 2 standard deviations of the mean.

~99.7% of data falls within 3 standard deviations of the mean.

But what if a data point does not fall perfectly on ±1, ±2, or ±3 standard deviations from the mean?

That is referred to as the **z-score.** A **z-score** tells you how many standard deviations a score is above or below the mean. Let's break the formula down:

$$z = \frac{x - \mu}{\sigma}$$

Examples:

The mean length of a hammerhead shark is 99.6 cm with a standard deviation of 8.7.

(Click here to access data used: https://www.researchgate.net/figure/Average-size-fork-length-FL-and-standard-deviation-SD-of-sharks-measured-for-all_tbl2_273143398)

a. On a boat trip you spot a shark which is 103.4cm long. What is the z-score for this shark?

b. What is the z-score of a shark that is 80cm long?

c. How do the values from *a* and *b* compare? What does this tell you?

Though it is helpful to know z-score, it may not be the most useful way to explain how a score relates to other data. A more common way to compare is with **percentiles**. A **percentile** is the percentage of data which is at or below the given point.

16. Have you heard of percentiles before? Where?

Example:

If I am at the 95^{th} percentile for my height in my age group, I am taller than 95% of people in my age group.

17. Look at your work on numbers 13 and 14. Explain those answers in terms of percentiles.

As you can see, there is a close tie between z-scores and percentiles. Using this technique, your calculator can find a percentile given a data point, a mean, and a standard deviation.

Using your calculator to practice the following:

18. The mean IQ is 100 with a standard deviation of 15. Find the percentile of a person with an IQ of…
    a. 81

    b. 113

    c. 100

Lesson 4: Finding Values from Percentiles

Similarly, your calculator can find a value at a given percentile. Use your calculator to find the following:

19. The mean IQ is 100 with a standard deviation of 15.

    a. A person is a the 62^{nd} percentile for their IQ. What is their score?

    b. A person in the 50^{th} percentile?

    c. A person who only had 2% of people score higher than them?

Review:

20. In 2018, the mean batting average in Major League Baseball (MLB) was the highest in the modern era at .248 with a standard deviation of .015.
    a. Find the z-score of a batter with a batting average of .230 in 2018.

    b. Find the percentile of that batter compares to the rest of the MLB.

Project Based Learning Assignment

Over the next few classes, you will design your own experiment to collect or find data and analyze it using the information you have learned in class.

A rubric will be given so you know what is expected of your experiment. To start, we need to select a topic and develop a focus.

Here is an example:

I am interested in promoting mental health. I notice students are very stressed and not getting enough sleep. Students should get 8 hours of sleep a night. I will ask students 'How much sleep did you get last night?' and analyze the data collected.

Thus, my research question could be: *Are high school students getting enough sleep?*

21. What are 3-5 topics which interest you?

22. What are questions you have revolving around these topics?

23. Pick one question which you could study further and write a research question that you could use to collect data.

24. Run your question by a partner so they can edit it, be careful you are not promoting bias!

25. Have the teacher check your research question. They will initial your rubric when your topic and question are approved.

Name: **Statistics Project**

My topic is: Approved:

	3	2	1	0
Plotting Data **(3 points)**	Student can accurately and appropriately plot data.	Student may choose an inappropriate graph type **OR** make 1-2 plotting errors.	Student makes many mistakes in plotting data.	No evidence in student work
Calculating Mean and Standard Deviation **(2 points)**		Student correctly calculates the mean and standard deviation.	Student makes one calculation error in calculating mean and standard deviation.	Student does not calculate the mean and standard deviation.
Explanation in Context **(2 points)**		Student describes the mean and standard deviation in the context of their experiment.	Student can describe the mean **OR** standard deviation in context **OR** explanation lacks clarity.	Student does not correctly explain mean and standard deviation in context.
Calculation of z-score **(3 points)**	Student selects one point from their data and accurately calculates its z-score with work shown.	Student selects one point from their experiment and calculates the z-score with one calculation error and work shown.	Student selects one point from their data and makes a conceptual error in calculating its z-score.	Student does not calculate the z-score.
Calculation of Percentile (**2 points)**		Student selects one point from their data and accurately calculates its percentile with work shown.	Student selects one point from their data and calculates its percentile with one or two errors.	Student does not calculate the percentile.
Conclusion **(2 points)**	Student designs an effective research question and writes a valid conclusion based on the data collected and calculations performed.	Student writes an acceptable research question and the conclusion loosely based on the data collected and calculations performed.	Student writes ineffective research question and has a conclusion not backed by data collected.	Student does not write a valid conclusion or research question.

My Turn

I was initially uncomfortable with the statistics curriculum. However, through using a curriculum like what is presented in this curriculum project I learned that incorporating a PBL assignment can provide unprecedented student engagement, enthusiasm, and learning. Not only did I see increased student success, but students also grew to develop a love of statistics. Other teachers who use this curriculum may need to modify the curriculum to better suit their instructional and student learning styles.

I used these lessons and project in a high school in Western New York. Though brief, the physical collection of data engaged students, allowed a low entry point for students of all abilities, and created a concrete base for students to build their understanding. By simply creating an experiment of heads-or-tails, students were be drawn in and excited to learn.

After this hook, students utilized the data to answer prompts. The prompts were designed to be open-ended and allowed students to create their own understanding. I acknowledge that giving up classroom control to teach to gain a student-centered learning approach can be daunting. However, with sociomathematical classroom norms where collaborative learning is expected I was pleased with the excitement students brought to their project. I believe if teachers can effectively use their energy to guide learning, then students can develop a deeper understanding.

I have found that students prefer to engage in practice exercises that are applicable to the real-world. Most mathematical concepts can be applied in the abstract easily and successfully. However, for statistics that is not the case. Statistical question design and use thrives in the concrete. This is a best practice for engaging students by bringing them in through their interests and relevance. I chose questions around baseball from their experiences with a baseball fan in class.

I likewise encourage teachers to cater the material to their classroom population. There is abundant data on a plethora of subjects!

Lastly, I left the project rubric intentionally open-ended as I was impressed and surprised by what students come up with for their projects. Many students were excited with the opportunity to be assessed in a way that is not a traditional test. It can be enlightening for teachers to look through the lens of *I want to know what my students know*. Though some students may thrive with this type of assessment, teachers should note to check in on students and give guidance to those who need more direction to effectively complete tasks. Teachers should also steer students toward topics which are appropriate and seek data which is attainable. The author notes that the beauty of a rubric is students should know exactly what to put in their project, and how in depth it should be (point values).

Conclusion

The curriculum should provide deeper conceptual understanding opportunities for students, increase student engagement, provide learning skills applicable for the 21st century, and assessments which allow all students to share their understanding through performance tasks. The curriculum accomplishes the following while also being aligned to the common core curriculum standards for statistical reasoning as used in Algebra II in New York State. It was created by a teacher for teachers. I hope you can use these resources and ideas in your own classroom.

References

Anastasiadou, S. D. (2012). *Diversifications between Expected and Perceived Attitudes toward Learning Statistics with Technology: A Multilevel Statistical Analysis*. 17.

Bell, S. (2010). Project-Based Learning for the 21st Century: Skills for the Future. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *83*(2), 39–43. https://doi.org/10.1080/00098650903505415

Boaler, J. (2002). Exploring the Nature of Mathematical Activity: Using Theory, Research and 'Working Hypotheses' to Broaden Conceptions of Mathematics Knowing. *Educational Studies in Mathematics*, *51*(1/2), 3–21.

Beswick, K. (2011). Putting Context in Context: An Examination of the Evidence for the Benefits of 'Contextualized' Tasks. *International Journal of Science and Mathematics Education*, *9*(2), 367–390. https://doi.org/10.1007/s10763-010-9270-z

Catch and Bycatch in U.S. Southeast Gillnet Fisheries, 2008 - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Average-size-fork-length-FL-and-standard-deviation-SD-of-sharks-measured-for-all_tbl2_273143398 [accessed 29 Mar, 2021]

Common Core State Standards (CCSS) (2021). National Governors Association Center for Best Practices, Council of Chief State School Officers. Common Core State Standards (Mathematics). Washington, D.C. http://www.corestandards.org/

Craig, T. T., & Marshall, J. (2019). Effect of project-based learning on high school students' state-mandated, standardized math and science exam performance. *Journal of Research in Science Teaching*, *56*(10), 1461–1488. https://doi.org/10.1002/tea.21582

Fitzmaurice, O., Leavy, A., & Hannigan, A. (2014). Why is statistics perceived as difficult and can practice during training change perceptions? Insights from a prospective mathematics teacher. *Teaching Mathematics and Its Applications*, *33*(4), 230–248. https://doi.org/10.1093/teamat/hru010

Franklin, C., Kader, G., Newborn, D., Moreno, J., Peck, R., Perry, M., Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K-12 curriculum framework*. American Statistical Association.

Kramarski, B., Mevarech, Z. R. & Arami, M. (2002). The effects of metacognitive instruction on solving mathematical authentic tasks. Educational Studies in Mathematics, 49, 225–250.

Leinwand, S., Brahier, D. J., & Huinker, D. (2014). *Principles to Actions: Ensuring Mathematical Success for All*. Reston, VA: NCTM, National Council of Teachers of Mathematics.

Lovett, J. N., & Lee, H. S. (2017). New Standards Require Teaching More Statistics: Are Preservice Secondary Mathematics Teachers Ready? *Journal of Teacher Education*, *68*(3), 299–311. https://doi.org/10.1177/0022487117697918

National Council of Teachers of Mathematics (NCTM). (2011, October). *Strategic Use of Technology in Teaching and Learning Mathematics*. National Council of Teachers of Mathematics. https://www.nctm.org/Standards-and-Positions/Position-Statements/Strategic-Use-of-Technology-in-Teaching-and-Learning-Mathematics/.

National Science Board. (2016, January). *Science and Education Indicators 2016*. National Science Foundation. https://www.nsf.gov/statistics/2016/nsb20161/#/report/chapter-1/transition-to-higher-education/preparation-for-college

Oxford College. (n.d.). *Shape, Center, and Spread of a Distribution*. The Department of Mathematics and Computer Science. Retrieved February 16, 2021, from http://mathcenter.oxford.emory.edu/site/math117/shapeCenterAndSpread/

Rossman, A., Chance, B., & Medina, E. (2007.). *Some Key Comparisons between Statistics and Mathematics and Why Teachers Should Care*. 14.

Stoll, L., & Schultz, S. (2019). How to Design a Performance Task. *Science Scope*, *042*(07). https://doi.org/10.2505/4/ss19_042_07_40

Sun, K. L. (2018). Brief Report: The Role of Mathematics Teaching in Fostering Student Growth Mindset. *Journal for Research in Mathematics Education*, *49*(3), 330–335. https://doi.org/10.5951/jresematheduc.49.3.0330

Appendix

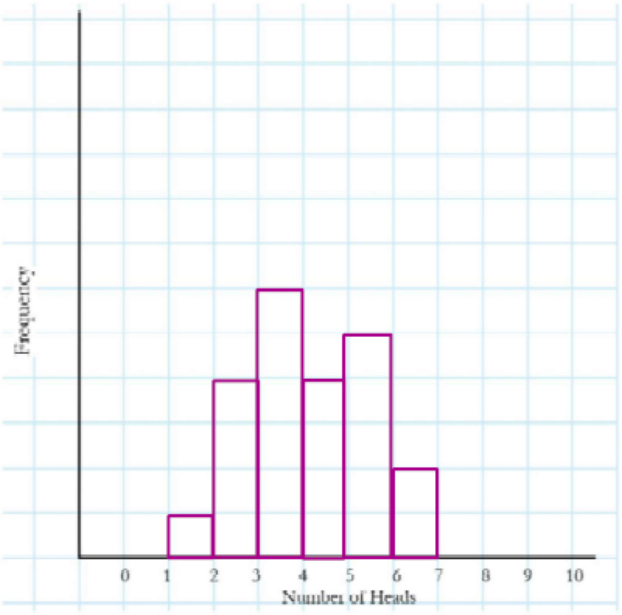
Day 1: Introduction and Review

Each person in class will flip a coin 10 times and count the number of times they get "heads". We will plot the classes data in the ***histogram*** below.

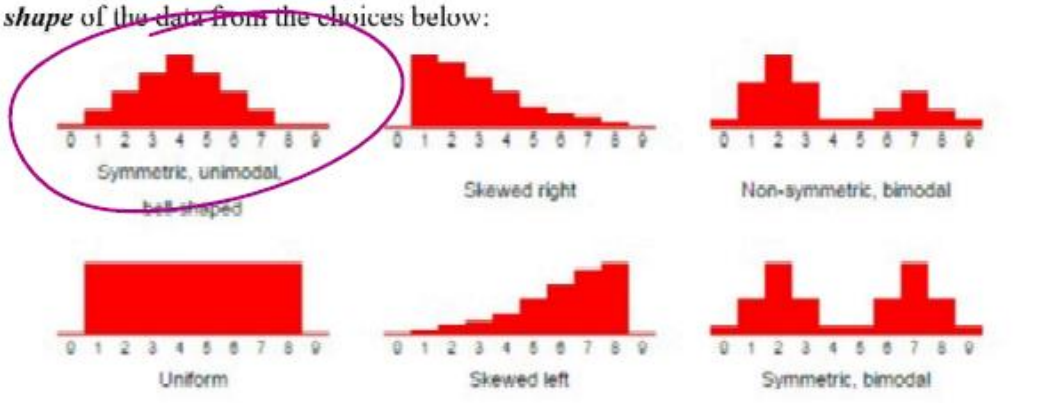
Histogram – a graphical display where data is displayed in ranges.

You can also use a random generator like the one found at https://www.random.org/coins/.

Num. of Heads	Frequency
1	
2	|
3	||||
4	~~|||||~~
5	||||
6	~~||||~~
7	||
8	
9	
10	



1. The shape of the data can help someone better understand our results. Let's describe the ***shape*** of the data from the choices below:



(Oxford College)

2. Beyond the shape of the data, what other ways can we describe the data we collected. In other words, what ways have we measured data in the past?

Examples include:
Mean, median, mode, range, 1st or 3rd Quartile, IQR

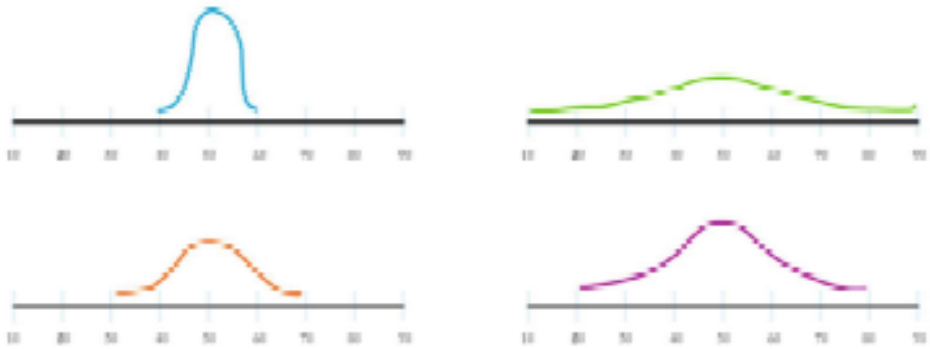
*This is a brainstorm, allow any and all ideas.

3. Let us sort those into two types of measures:

*I would have a brief discussion of what the words center and spread mean to make sorting easier, you do not need to sort them all!

Measures of Center	Measures of Spread
Mean Median	Range Interquartile range (IQR)

4. Consider the four sketches of data sets below. Why can't we just give someone the Shape and a Measure of Center when describing data?

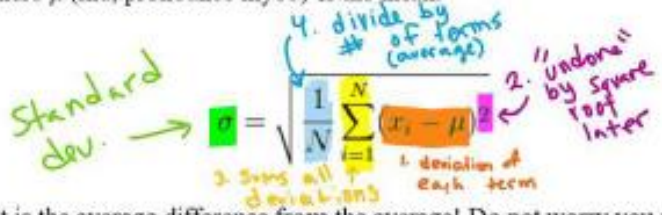


Example: All of these data sets have a mean around 50. You need spread to show how the data lies around the mean.

5. A new measure is called the ***standard deviation***.
Have you heard of *deviation* or *deviant* before outside of math class? What does it mean?

Examples: someone who is an outsider.
*meant as a brainstorm

Deviation is difference from the mean, how far you are from the norm. So a ***Standard Deviation*** is how much data typically differs from the mean. It is calculated by this formula, where μ (mu, pronounce myoo) is the mean.



Basically, it is the average difference from the average! Do not worry you will not need to calculate it by hand.

Take out your calculator and let us input the class data from the first page.

6. Calculate the mean and standard deviation of the introductory data set.

*I typically create calculator references for my students to show how to input data into my school's version of calculator/software and complete calculations

7. Is the standard deviation a measure of center or measure of spread?

Spread

8. If a data set has a large standard deviation, what does that mean? A small standard deviation?

A large standard deviation: data is typically far from the mean
Small: data is typically close to the mean

Day 2: Practice with Standard Deviation

9. Using the coin flip data collected yesterday, find…
   a. The value which is one standard deviation above the mean.

   b. The value which is one standard deviation below the mean.

   c. What percentage of the data was in between the values you found in part *a* and *b*?

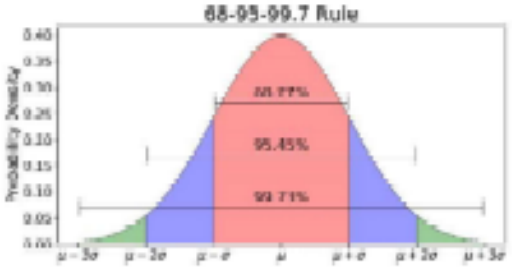
      *should be around 68%

   d. Based on how we described the standard deviation, what do the values you found in *a* and *b* tell us?

10. Follow procedures *a-d* above but now for 2 standard deviations.

    *should be around 95%

11. Your work above should reflect what mathematicians have found over time. This is the power of the standard deviation! No matter what data you use, you will find…
    a. Most data (68%) of data falls within one standard deviation of the mean.
    b. 95% of data is within 2 standard deviations.
    c. Almost all (99.7%) of data is within 3 standard deviations.



When you know the standard deviation, you know how data TYPICALLY behaves.

Practice:

12. Miguel is the mean height for his age group. Carl is one standard deviation above the mean. What percentage of people in the age group are between Miguel and Carl?

68/2=34

34%

13. Clarise was told her height is two standard deviations above the mean. What percentage of people in her age group is Clarise taller than?

95/2 = 47.5

50+47.5=97.5%

14. Mean height for 10 year-olds is 54 inches with a standard deviation of 2.6. How would someone who is 51.6 inches tall compare to their peers?

51.6-54=-2.4
So 1 SD below the mean
50-47.5=2.5
They are shorter than 97.5%, and taller than 2.5%

15. You ask your teacher how you scored on the statistics test. They state that the mean score was a 80 points and the standard deviation was 6 points. You are told that your score was better than 97.5% of your peers. What did you score?

97.5-50=47.5
47.5*2=95
So you scored 2 standard deviations above the mean.

80+6+6=92
You scored a 92.

Day 3: Z-Scores and Percentiles

Last class we discovered what a standard deviation was and what it tells us about where data lies. To summarize…

~68% of data falls within 1 standard deviation of the mean

~95% of data falls within 2 standard deviations of the mean.

~99.7% of data falls within 3 standard deviations of the mean.

But what if a data point does not fall perfectly on ±1, ±2, or ±3 standard deviations from the mean?

That is referred to as the **z-score.** A **z-score** tells you how many standard deviations a score is above or below the mean. Let's break the formula down:

$$z = \frac{x - \mu}{\sigma}$$

Examples:

16. The mean length of a hammerhead shark is 99.6 cm with a standard deviation of 8.7 (Catch and Bycatch, 2008).
    a. On a boat trip you spot a shark which is 103.4cm long. What is the z-score for this shark?

103.4-99.6=3.8/8.7=.437

    b. What is the z-score of a shark that is 80cm long?

80-99.6= -19.6/8.7=-2.2529

    c. How do the values from *a* and *b* compare? What does this tell you?

A positive z score is above the mean, negative below

    d. Estimate the percentiles for the values found in *a* and *b*.

Though it is helpful to know z-score, it may not be the most useful way to explain how a score relates to other data. A more common way to compare is with **percentiles**. A **percentile** is the percentage of data which is at or below the given point.

17. Have you heard of percentiles before? Where?

*brainstorm: example: doctor's office

Example:

If I am at the 95th percentile for my height in my age group, I am taller than 95% of people in my age group.

18. Look at your work on numbers 13 and 14. Explain those answers in terms of percentiles.

As you can see, there is a close tie between z-scores and percentiles. Using this technique, your calculator can find a percentile given a data point, a mean, and a standard deviation.

Using your calculator to practice the following:

19. The mean IQ is 100 with a standard deviation of 15. Find the percentile of a person with an IQ of…

a. 81

normCdf(-9.E999, 81, 100, 15) = 0.10263

10th percentile

b. 113

normCdf(-9.E999, 113, 100, 15) = 0.8069377

81st percentile

c. 100

normCdf(-9.E999, 100, 100, 15) = 0.5

50th percentile

*once again, I create references for the calculator for students. There are many great video references online.

Day 4: Working Backward

Similarly, your calculator can find a value at a given percentile. Use your calculator to find the following:

20. The mean IQ is 100 with a standard deviation of 15.

    a. A person is a the 62nd percentile for their IQ. What is their score?

    invNorm(.98, 100, 15) = 104.5822 ~ 105

    b. A person in the 50th percentile?

    invNorm(.98, 100, 15) = 100

    c. A person who only had 2% of people score higher than them?

    invNorm(.98, 100, 15) = 130.8 ~131

Review:

21. In 2018, the mean batting average in the MLB was the highest in the modern era at .248 with a standard deviation of .015.

    a. Find the z-score of a batter with a batting average of .230 in 2018.

    .230-.248 = -0.018 / .015 = -1.2

    b. Find the percentile of that batter amongst the MLB.

    normCdf(-9.E999, .230, .248, 0.015) = 0.115

    The 12th percentile

Example Project:

Research Topic: Sleep

Research Question: How much sleep did you get last night?

Hours	Tally
0	1
1	0
2	0
3	2
4	1
5	4
6	8
7	4
8	2
9	1
10	1
10+	0

Mean: 5.875

Standard Deviation: 2.02715

On average, a student sleeps 5.875 hours a night. The average student sleeps between 3.84 and 7.90 hours a night.

I slept 7.5 hours last night.

$$z = \frac{7.5 - 5.875}{2.02715}$$

$$z = .8016$$

My amount of sleep is .8016 standard deviations above the mean.

NormCdf (7.5, 5.875, 2.02715) = 0.788613

I know that I am in the 78th percentile for sleep. I slept more than 78% of my peers.

Teenagers should get 8 hours of sleep a night. My analysis shows that most students are not getting enough sleep at night. I believe that this has had a large effect on student mental health in our student body. We should have a program to promote healthy sleep.