

High School – Algebra 1

Intro to Statistics and Interpreting Data

This learning progression is for a classroom with 9th and 10th graders in an algebra 1 class. The class will be introduced to statistics concepts or interpreting data. Students will be able to read and create graphs and tables, calculate key components of a data set (median, mean, interquartile range, standard deviation), and contextualize the shape of the data set (skew, distribution). The tasks used in this learning progression are of my own design, but the progression follows like that of South Western's *Algebra 1 – An Integrated Approach* textbook. The goal of the progression is for students to meet the standards (given on column at right).

The content the progression will cover will be brand new to the students, as the title of the progression suggests. It begins with a worksheet with tasks about concepts of reading data graphs and tables, as well as creating graphs and tables from a given data set. The next worksheet will contain tasks cover concepts of data middles (median, mean, mode), interquartile ranges, and standard deviation. Furthermore, this task will have students understand what information each calculation finds and when to apply said calculation. The final worksheet will have tasks will have students analyzing the shape (skew), distributions, and centers (median, mean, mode) of data sets with the same quantifiers and outliers in data. The concepts in the task will be introduced to students by inquiry. Materials will be given to students (formulas and definitions), but how to use the materials and what their contextual meaning is will be for the students to discover. The pacing for each worksheet will be one, as well as one exit slip, for one instructional period.

Instruction will be based on inquiry, asking students to figure out how to complete a task, then reporting out their solutions. Using inquiry in instruction provides students a chance to use their current knowledge and – through scaffolding and assisted learning – build upon that

Common Core Standard

CCSS.Math.HS.ID – Interpreting categorical and quantitative data.

Summarize, represent, and interpret data on a single count or measurement variable.

HS.ID.A.1. – Represent data with plots on a real number line (dot plots, histograms, and box plots).

HS.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.

HS.ID.A.3. – Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

Mathematical Practices

MP2 – Reason abstractly and quantitatively.

MP3 – Construct viable arguments and critique the reasoning of others.

MP7 – Look for and make use of structure.

knowledge. This use of scaffolding is like that of Lev Vygotsky's Zones of Proximal Development (ZPDs) Theory. To help promote student thinking, students will pair up and work together to find solutions to inquiry items. Students will be monitored for understanding and discourse. The monitoring will mostly take the form of listening to pairs and checking in with pairs from time to time, but a portion of the monitoring will be with the whole class, asking pairs to report out where they are in the worksheet by show of hands and what ideas or solutions pairs have. I will purposefully avoid asking questions that related to the questions on the worksheet, as I want students to experience their own discoveries and not be lead to one solution or another. If there are any contradictions between student responses, they will be dealt with in such a way that elicits discourse and mathematical reasoning. For example: if a pair of students disagrees with how another pair chose to represent their data, I will ask for each group's reasoning and invite the rest of the class to give input – with reasoning – as well. This will promote mathematical thinking, while still allowing for development of conceptual understanding through discovery.

Among this group, there are a few students who are English learning (EL) and one student who has Emotional-Behavioral Disorder (EBD). During instruction, the EL students will be purposefully paired with students whose first language is English. I will use the pre-pair and think-pair-share strategies, suggested by Jodi Reiss' *120 Content Strategies for English Language Learners*. These strategies are meant to reduce stress due to second-language use and promote academic language acquisition/use. The pairs with EL students will be monitored, and I will elicit mathematical discourse by asking "why?", or "how do you know?", questions and eliciting visuals by asking students to "show me [what you're thinking]." This will help promote the use of the language functions and language acquisition/use. The EBD student has exhibited lack of attention and patience, which can lead to loud outbursts. However, the student is Tier – 1 for RTI and is included in general education. To accommodate the needs of the student, clear guidelines about when students may talk have been given, and the student is aware of them. Furthermore, the EBD student will be provided clear instructions on the task worksheets, monitored for pacing of progress, and given re-teaching when need be. These tactics are best practice suggestions from Rena Lewis' *Teaching Students with Special Needs in General Education Classrooms*.

Formative assessment of conceptual understanding will ask students will report out the class their findings after working in pairs. These reports will require components of oral expression and visual aid to represent the students' mathematical reasoning and conceptual understanding. During these moments of sharing, and even in smaller discussions, students use of mathematical language (vocabulary, syntax, discourse) will be assessed. Included in the formative assessments are exit slips for every task. Each exit slip will ask students to demonstrate their conceptual understanding by giving a written explanation, as well as use a real-world example with a visual. Furthermore, the exit slips will ask students to use the same real-world example (with modifications if the original example is not applicable) so that students are building information and knowledge upon the same foundation.

Worksheet 1: Reading and Writing Tables and Graphs – Representing Categorical and Quantitative Data

Standard HSS.ID.A.1 - Represent data with plots on a real number line (dot plots, histograms, and box plots).

Students will work towards MP2 as they analyze and, contextualize, and decontextualize data sets with reasoning; MP3 as they give reasonings for critiques on graph/table use and peer solutions.

Hinge Question: Why do some graphs/tables better represent some data and not others? Give an example with a graph or table.

Assessment: answer hinge question on exit slip.

Student will be given the standard, written with “I can – ” prefixing the quoted text, at the beginning of class to promote mathematical reasoning. Students will then pair up and work together on a worksheet with tasks relating to the standard. The worksheet will introduce several types of graphs and tables, but will not address what type of data they are best for.

The first set of tasks on the worksheet will ask student to read and contextualize graphs and tables. Students must give reasoning for how they interpret the graphs/tables. For example: a student will be given a histogram with data on human heights. The student may respond with “this graph shows the how many people are of each height.” The next set of tasks will ask students if the graph/table that is used best represents the data. Students must give mathematical reasoning to why, or why not, they think a graph/table best represents the data set given. For example: students will be given a box-and-whiskers plot with categorical data rather than quantitative data. A student might respond to this by stating “the box-and-whiskers plot does not best represent this data set because categorical data cannot be represented on a continuum/number line.” The last set of tasks will ask students to create graphs/tables for data sets. They must then give mathematical reasoning to why they chose to represent the data in that way. For example: students will be given a set of data that shows how many college students are of a given major. A student response may be to represent the data as a histogram and reason “this is the best graph because it shows how many students are in what major.”

At the end of class, students will be complete an exit slip with a hinge question: “Why do some graphs/tables better represent some data and not others? Give an example with a graph or table”. The hinge question will be separate from the worksheet. An example of a student’s response could be “because of the type of data, or what you want to show, some graphs/tables are better than others.” The student may then create a set of data and represent it with two different graphs, detailing what each one highlights about the data.

Worksheet 2: Give a Prediction – Using Statistics Appropriate to the Shape, Distribution, Center, and Spread of Data

Standard 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.

Students will work towards MP2 as they contextualize calculations for centers, standard deviations, distributions, and interquartile ranges; MP3 as they give reasonings for critiques on peer solutions.

Hinge Question: What are the different pros and cons (what makes each one good or bad) of using median, mean, mode, interquartile range, and standard deviation? Give one pro (good thing) or con (bad thing) for each item. Build upon the example you used for the previous exit slip.

Assessment: answer hinge question on exit slip.

At the beginning of class, the standard, presented in “I can – ” form, will be given to students. Students will also pair up and work through a worksheet together. This worksheet will introduce definitions for left/right-skewed, median, mean, mode, interquartile ranges, and standard deviation, and distribution. The definitions will be technical, and will offer no contextual examples. Equations for mean and standard deviation will be given. Throughout the worksheet, students must analyze a data set, or the representation of one, and decided how to contextualize it with the information they are given.

The first set of tasks will ask students to contextualize the skew of different data sets. Students must give reasonings to their thinking. For example, if given a right-skewed graph that represents test score averages for a biology class, a student’s response may be “the biology class did not do very well on the test because most of the data is on the low end of the graph.” Students will next have tasks asking them to analyze different centers of data sets and graphs/tables, which center the set is representing, and what the contextualization of that center is. For example: a student is given a stem-and-leaf table of the biology class from the previous example. Their response may be “this table is highlighting the mode of the data because it shows which score was most frequent.” The next tasks will ask students to calculate interquartile ranges, standard deviation, and distribution. Students will also have to contextualize their meaning. For example: a student will be given a data set whose standard deviation is 2 (and for the sake of repetition let the data set be the scores of the same biology class, yet again). The student may respond that “the average distance between student scores is 2 points.” The last set of tasks will give students a description of a data set, and students must draw a graph or table that represents that set with reasoning to support their decision. For example: a description of a set would be “The running speeds of house cats is collected. The average of the set is 5 miles per hour, with 25% of the data between 1 and 21.” The student may draw a box-and-whiskers plot with the first quartile between 1 and 21, with the reasoning that “the first quartile must be between 1 and 13,

because that is where 25% of the data is. Furthermore, the next quartile must be between 21 and 25, because 21 is the end of the first quartile and 25 is the median.”

Like for the previous worksheet, at the end of class students will complete an exit slip with a hinge question: “What are the different pros and cons (what makes each one good or bad) of using median, mean, mode, interquartile range, and standard deviation? Give one pro (good thing) or con (bad thing) for each item. Build upon the example you used for the previous exit slip.” The reason for giving a laymen’s term for pro and con is because they are not words used as part of the vocabulary being assessed in the mathematical discourse or language functions. There may be an EL student who may not know what pro and con mean. By providing the laymen’s definition, the EL student can still understand the meaning of the hinge question, and it provides some new vocabulary for them. An example of a student’s response could be a table with the vocabulary in one column and a “pros” and “cons” columns.

Worksheet 3: Interpreting Differences – Comparing Data Sets with the Same Quantifiers

Standard HSS.ID.A.3 - Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

Students will work towards MP2 as they compare data sets and reason the contextual difference of what the sets are showing; MP3 as they give reasonings for critiques on peer solutions; MP7 as students compare the structures (shapes, centers, distributions) of different data sets.

Hinge Question: Why does the shape of data matter? When can some data be ignored? Give examples and graph, or table, for both questions. Build upon previous exit slip examples.

Assessment: answer hinge question on exit slip.

The beginning of class will begin the same way it had for the previous worksheets. This final worksheet will have students looking at different data sets with the same quantifiers, defining their abstract differences (e.g. median A is higher than median B, there is more distribution in set A than in set B, etc.), then contextualizing these differences.

The first set of tasks will have students identify outliers, and reason how they affect calculations. For example: a student is given a data set with a value that is 2 standard deviations above the value under it. The student reasons “the outlier will affect the median and mean of the data, pulling it away from the center of the rest of the data.” The final tasks will ask students to create graphs/tables for two data sets with the same quantifiers. Students then must reason what the key differences are, and calculate them, and contextualize the differences. They must do the same thing for similarities. For example: a student is given two data sets that represent the growth rate of two populations of pea plants. The student chooses to represent the data sets as box-and-whisker plots. They see that pea-population A is left-skewed, while pea-population B is not skewed. They reason “the peas in population A grow faster than the peas in population B. This is represented by the left-skew of population A.”

The variety of tasks in this worksheet is noticeably less than in the previous two. However, creating graphs/tables, calculating different points, and reasoning contextualization gives these tasks more complexity. This may happen in the previous worksheets, but I assume for this last worksheet there will be a large discrepancy of pacing through the worksheet. I predict some students will finish before others by a large margin. If this is the case, those students will have the opportunity to begin drafting a portfolio entry for what they learned thus far in the *Statistics and Probability* unit. These entries will be finalized after the unit's summative assessment has been graded and returned to the students, but they will contain the content of the unit in the student's own understanding, a list of concepts they needed to know before progressing through the unit, how the content in the unit connects with the concepts in said list, and how they feel about the content leading into and out of the unit. Along with their written explanations, students will provided visuals to demonstrate their understanding.

At the end of class there will be another exit slip with a hinge question: "Why is it important to compare the shape of data? When can some data be ignored? Give examples and graph, or table, for both questions. Build upon previous exit slip examples." An example of a student's response could be "It is important to compare the shape of data sets to understand the differences between groups." The student may then proceed to give an example with an outlier, and contextualize the differences in the data sets' means, and why they choose not to include the outlier in their calculations.