|  |  |
| --- | --- |
| **Unit** | **S1** |
| **Title** | Statistics 1 |
| **Mathematical Goals** | Students will…   * Create dot plots, histograms and boxplots S.ID.1 * Learn methods for creating graphs and calculating measures using available classroom technology S.ID.1 * Develop language for describing data distributions S.ID.2 * Analyze and compare multiple data sets S.ID.3 * Understand the relationship between mean and median for symmetrical and skewed data distributions S-ID.2 * Recognize outliers when they exist, and know to investigate their source -- that bit of data is way out there, why is that? Is there something weird about it that means we should disregard it? S-ID.3 * Know that outliers affect the mean, but not the median of a data set S.ID.3 * Describe variability by calculating deviations from the mean S.ID.2 * Compare two datasets with the same means but different variability, and contrast them by calculating the deviation of each data element from the mean S.ID.2 * Learn to create a 5-number summary, box plot, and compute the interquartile range using technology S.ID.2 * Interpet IQR as a description of variability better-suited to a skewed distribution S.ID.3 * Interpret a table that divides data into different categories S.ID.5 |
| **Focus Standards** | **S.ID.1, S.ID.2, S.ID.3, S.ID.5** |
| ***The story before this unit (including prior knowledge)*** | In middle school, students are introduced to data sets and different ways to represent data (histograms, dot plots, box plots). Statistics is introduced as a tool to answer questions about a population that have variability in the answer. Students first see measures of center (median, mean) and measures of variability (interquartile range, mean absolute deviation). Students use these measurements to draw informal comparative inferences about two populations. |
| ***The part of the story happening in this unit*** | Students build on and expand their understandings of statistics in this unit. The key characteristics (measures of shape, center and spread) are again seen and in addition, students may further describe the shape of a data distribution (symmetric, skewed, flat or bell shaped) and summarize by a statistic measuring center and a statistic measuring spread. Instead of creating representations of data, the emphasis in high school is on *interpreting* representations and judiciously interpreting measures of center and spread.  Students will get precise with measures of center. They’ll learn that mean and median are equal for symmetrical distributions, explain why mean and median are not equal for skewed distributions, select median as the better measure of center for skewed distributions, and generalize what kinds of distributions have mean > median vs. median > mean.  Students will learn that standard deviation is a measure of spread, that a larger standard deviation means the data is more variable or spread-out, and the meaning of standard deviation as “typical distance from the mean” for a symmetrical distribution. While developing their understandings, students will calculate standard deviation by hand for a small data set at least once and will be able to interpret the meaning of standard deviation.  Given different visual representations of data (box plots, histograms, dot plots, frequency graphs) students can draw and justify significant and meaningful conclusions about the given situation. As every SP standard is a modeling standard, most work should be done in a specific context, and the work done should habitually be related back to the context.  Students are introduced to two-way frequency tables and understand how to interpret relative frequencies in the given context. |
| ***The story after this unit*** | In S2 (which could take place either before or after this unit), students will also build their statistics foundation by learning ways to determine whether two sets of data are correlated, and how strongly. Students will identify linear association and will interpret slope and intercept in the context of the data. Given different visual representations of data (linear models) students can draw and justify significant and meaningful conclusions about the given situation. Students begin to use technology as a means to plot data and generate correlation coefficients.  In S3, students will revisit two-way frequency tables from a probability standpoint, and use them as a tool for conceptualizing and finding conditional probabilities. Later on in S4, students will combine the ideas of distributions and probability. They will define normal distributions and use normality to solve problems, and be able to use the distributions of probability models to find the likelihood of a particular outcome.  In S4, students will further examine standard deviation and will be expected to generate standard deviation using technology. |

**UNIT FLOW SUMMARY**

|  |  |
| --- | --- |
| **UNIT S1** (17 - 20 days) | **Statistics 1** |
| **Section 0** (1 day) | **Diagnostic Pre-Unit Assessment** |
| **Section 1** (2 days) | **How can data be represented and summarized meaningfully? (hook/umbrella activity)** |
| **Section 2** (3 - 4 days) | **Explore Data Distributions** |
| **Section 3** (2 - 3 days) | **Measures of Center** |
| **Section 4** (1 day) | **Mid-Unit Assessment** |
| **Section 5** (3 days) | **Describing Variability with Deviations from the Mean** |
| **Section 6** (1 day) | **Variability for Skewed Distributions** |
| **Section 7** (1 - 2 days) | **Bringing it All Together** |
| **Section 8** (2 days) | **Two-way Frequency Tables** |
| **Section 9** (1 day) | **Summative Assessment** |

|  |  |
| --- | --- |
| **Section 0:** 1 day | **Diagnostic Pre-assessment** |
| **Pre-Unit Assessment Targets** | Diagnose students’ recall of middle school statistics, specifically their ability to...   * recognize a statistical question 6.SP.1 * describe the distribution of data collected to answer a statistical question by its center, spread, and overall shape 6.SP.2 * interpret statistical plots like dot plots and box plots 6.SP.4 * summarize numerical data sets in relation to their context 6.SP.5 * informally assess the degree of visual overlap of two numerical data distributions with similar variabilities 7.SP.3 |
| **Sample Activity** | * [*Pre-assessment*](https://docs.google.com/document/d/1k_Tp94VE3_i8lK9biN8jp8tUhM0aGWkYyGIN8GDgWIQ/edit?usp=sharing) |



|  |  |
| --- | --- |
| **Section 1:** 2 days | **How can data be represented and summarized meaningfully? (hook/umbrella activity)** |
| **Mathematical Goal** | Students will...   * Revisit various ways to represent data: dot plots, histograms, and box plots S.ID.1 * Interpret representations within their context S.ID.1 * Introduce the language of “symmetric” and “skewed” to describe distributions S.ID.2 * Establish expectations and norms for mathematical practices |
| **Narrative overview of section**  (and how the standards are achieved) | The goal is to revisit the methods for representing and summarizing data students learned in middle school. Students should be looking at and interpreting dot plots, histograms, and box plots, attending constantly to context. Not just what are the quartile values, but what do they mean? Not just whether the graph is symmetric or skewed, but what does that tell us about the population? |
| **Sample Activity 1.1** | [Wealth of Nations](http://mathalicious.com/lessons/wealth-of-nations), *Mathalicious*  **WHAT:** How is wealth distributed in America? Students interpret frequency graphs and calculate the mean and median income for citizens of three hypothetical countries. They recall that mean and median are measures of center, and while each tells you different information, neither gives you a complete picture of the data. We can better understand a set of data by analyzing how it is distributed. Then, they construct S.ID.1 and compare box plots of the wealth distribution in the three countries. Students must use appropriate tools (by hand or using technology) strategically when creating the box plots MP.5. Finally, they confront the reality that the way wealth is distributed in the United States today is way more skewed than any of these.  **WHY:** Statistics are useful because they help us describe and compare data collected in our world. It is important to start this unit to use statistics in a meaningful way. Given the news in recent years about income inequality, and powerful entities rigging the system to skew wealth distribution even further, the utility of statistics for describing and modeling our reality is apparent MP.4. |
| **Sample Activity 1.2** | [M&M’s Activity](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsOFE4T3FWalg5S3M/edit), *High Tech High, Jade White*  **WHAT:** Students analyze the M&M’s in their personal bag and expand their data collection to that of the whole class. This requires collaboration, communication, and attention to precision MP.6. Students represent their data graphically in at least one way and are challenged to represent it multiple ways if they can S.ID.1.  **WHY:** The M&M’s Activity does not have heavy algebraic demands so it gives *all* students an entry point into data collection, sample size and data representation. |
| **Focus Standards** | S.ID.1 Represent data with plots on the real number line (dot plots, histograms, and box-plots). |
| **Mathematical Practices** | MP.4, MP.5, MP.6 |

|  |  |
| --- | --- |
| **Section 2:** 3-4 days | **Explore Data Distributions** |
| **Mathematical Goals** | Students will...   * Create dot plots, histograms and box plots S.ID.1 * Learn methods for creating graphs and calculating measures using available classroom technology S.ID.1 * Develop language for describing data distributions S.ID.2 * Analyze and compare multiple data sets S.ID.3 |
| **Narrative overview of section**  (and how the standards are achieved) | Much of this section will review ideas developed in sixth and seventh grade (but not eighth grade), so the teacher can take the opportunity to address any gaps in understanding revealed in the pre-assessment. The addition to the previous knowledge will be the words used to describe distributions (flat, skewed, bell-shaped, symmetric). With any task used, students should always be asked to interpret the statistical terms and measures in the context of the measurements at hand.  Additionally, the teacher may opt to use these activity to demonstrate using the class’s available technology to create histograms, box plots, and calculate measures of center and spread, so that students are more equipped to choose the appropriate tool MP.5 whenever they must analyze data in the future. |
| **Sample Activity 2.1** | [Dice Activity](https://docs.google.com/a/hightechhigh.org/document/d/1t9UuteNR2vqrQDOEICgPdZf1KPIrdBGW1zFfs5A5M7I/edit), *High Tech High, Jade White*  **WHAT:** The class creates a frequency graph of the outcomes of rolling 2 dice. This sets up a formal lesson about the spread, shape, center, mean, median, mode and quartiles, including formal definitions and examples of what each looks like. The context provides an opportunity to academically discuss different data distributions and connect a frequency graph with a corresponding box-plot.  **WHY:** Before introducing formal academic language, it’s important to provide a context, so that the vocabulary can be attached to something. Also, before being able to interpret a given graph, students should have an opportunity to create their own S.ID.1 from data they collect, so they have a concrete basis for understanding what frequency graphs mean MP.4. |
| **Sample Activity 2.2** | [Penny Weight Activity](https://docs.google.com/a/hightechhigh.org/document/d/1z-TF39lNwbpbv5DunxO9ZWa2ERJh5lFtDVC4TybYwWk/edit), *High Tech High, Jade White*  **WHAT:** The class collects the weights and years from a batch of pennies. They create frequency graphs S.ID.1, and analyze them using language developed previously. The distribution of the years should be pretty flat but may be skewed toward recent years, assuming a random sample of pennies in circulation, and the distribution of weights should show high frequencies around one value -- a steep bell shape.  **WHY:** In order to understand the meaning of vocabulary describing the shapes of distributions S.ID.2, it is important to compare and contrast several examples. It also provides students an opportunity to independently, but with teacher support, practice using technology MP.5 (if used) to create histograms and box plots, and calculate measures of center and spread. |
| **Sample Activity 2.3** | [Speed Trap](http://www.illustrativemathematics.org/illustrations/1027), *Illustrative Mathematics*  **WHAT**: Are the speed distributions similar for cars traveling northbound and for cars traveling southbound on an interstate highway? Students draw box plots of two data sets S.ID.1, and then use the plots and appropriate numerical summaries of the data to write a few sentences comparing S.ID.2 the speeds of northbound cars and southbound cars at this location.  **WHY**: Speed Trap could be used if students need more practice finding quartiles, creating box plots, and analyzing and comparing statistical measures in context S.ID.3. |
| **Sample Activity 2.4** | [Using Frequency Graphs](http://map.mathshell.org/materials/download.php?fileid=1230), MARS  **WHAT:**  In this activity students are asked to read and interpret multiple data distributions. Students are given eight different distributions with eight different written interpretations and they must connect the distributions with their descriptions based on the shape, spread and center of the given graphs S.ID.2 S.ID.3. Students must look for and express regularity in repeated reasoning by working through the eight examples; understanding the differences and how this affects the graphs MP.8.  **WHY:** The purpose of this activity is to provide students with an opportunity to connect multiple distributions with specific outcomes in a real-world context MP.2. Students work in pairs to discuss and construct arguments to support their choices of pairings MP.3. |
| **Sample Activity 2.5** | [Representing Data Using Box Plots](http://map.mathshell.org/materials/download.php?fileid=1243), *MARS*  **WHAT:** This activity builds off of [Using Frequency Graphs](http://map.mathshell.org/materials/download.php?fileid=1230) above. In this extension, students are given eight data distributions (the same eight as above) and eight box-plots and they must connect the two different data representations by analyzing differences in shape, center and spread S.ID.2 S.ID.3. If the Using Frequency Graphs activity was also completed, students can connect the two data representations here with the written descriptions, combining the two activities. Students must look for and express regularity in repeated reasoning by working through the eight examples; understanding the differences and how this affects the graphs MP.8.  **WHY:** The purpose of this activity is to provide students with an opportunity to connect two different ways of representing data (data distribution and box-plot) in the same context MP.2. Students work in pairs to discuss and construct arguments to support their choices of pairings MP.3. |
| **Target Standards** | S.ID.1 Represent data with plots on the real number line (dot plots, histograms, and box plots).  S.ID.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.  S.ID.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** | MP.2, MP.3, MP.4, MP.5, MP.8 |

|  |  |
| --- | --- |
| **Section 3:** 2 - 3 days | **Measures of Center** |
| **Mathematical Goals** | Students will...   * Recall how to calculate mean and median * Understand mean and median as a “typical value” that can answer a statistical question. * Know that mean = median for symmetrical data distributions S-ID.2 * Explain why mean ≠ median for skewed data distributions S-ID.2 * Select mean as the better measure for symmetrical data, and median as the better measure for skewed data S-ID.2 * Generalize what kinds of distributions have mean > median, and what kinds have median > mean S-ID.2 * Recognize outliers when they exist, and know to investigate their source -- that bit of data is way out there, why is that? Is there something weird about it that means we should disregard it? S-ID.3 * Know that outliers affect the mean, but not the median of a data set S-ID.3 |
| **Narrative overview of section**  (and how the standards are achieved) | Students will be able to describe mean and median both as “typical values” that describe the center of a set of data. Knowing what they are is important for being able to use them to compare the center of data sets S.ID.2. Students will better understand the differences in using median vs mean to summarize data, and be familiar with cases where each is more appropriate. This is important for being able to choose a statistic appropriate to the shape of the data S.ID.2. They will be able to identify distributions where mean = median, mean > median, and mean < median, and explain why this occurs. This involves interpreting differences in shape in the context of data sets, S.ID.3. With any task used, students should always be asked to interpret the statistical terms and measures in the context of the population measured. Teachers may continue to use the class’s available technology to calculate measures of center. |
| **Sample Activity 3.1** | [Haircut Costs](http://www.illustrativemathematics.org/illustrations/942), Illustrative Mathematics  **WHAT:** In this problem, students are given data (minimum/maximum/quartiles/median/mean) in a table and are asked to graph the data in a box plot S.ID.1 and draw conclusions from it.  **WHY:** The focus is off computation, and rather on using attributes to compare two data sets S.ID.2. By asking which measure, mean or median, is more appropriate, an opportunity is presented for students to describe the effect of outliers S.ID.3 (some extreme high values “pull” the mean to the right). |
| **Sample Activity 3.2** | [Identifying Outliers](https://www.illustrativemathematics.org/illustrations/1888), Illustrative Mathematics  **WHAT:** In this task, students look at given data about the distances traveled to get to school. Students examine what happens as one of the data points moves; specifically they look at how the change of one data point can affect the mean and median S.ID.2. Additionally, this problem provides a good example of an outlier; students learn how to identify an outlier and examine how an extreme value affects an entire data set S.ID.3.  **WHY:** The purpose of this problem is to help students develop their understanding of extreme data points (outliers) and how they affect the measures of center of a data set. This task provides students a great opportunity to develop their skills in constructing arguments and critiquing the reasoning of others MP.3 as students must provide an opinion based on the given data to fully complete parts of this task. Lastly, this task provides an opportunity for students to use technology to create box plots and calculate measures of center if the teacher would like to use it to do so MP.5. |
| **Sample Activity 3.3** | [Describing Data Sets with Outliers](https://www.illustrativemathematics.org/illustrations/1875), Illustrative Mathematics  **WHAT:** In this task, students examine and model the mean and median in six very different yet distinct real-world situations MP.4. Students reason through the meanings and interpretations of the different measures of center in different contexts S.ID.2 MP.2. Students look at how outliers in each situation can affect the calculations for the mean and median S.ID.3.  **WHY:** The purpose of this problem is to allow students an opportunity to fully develop their understanding of the different measures of center in given real-world contexts. In order to fully explain and support their work, students must demonstrate sound understanding of the mean and median and how outliers affect the two values in calculations. |
| **Target Standards** | S.ID.1 Represent data with plots on the real number line (dot plots, histograms, and box plots).  S.ID.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.  S.ID.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** | MP.2, MP.3, MP.4, MP.5 |

|  |  |
| --- | --- |
| **Section 4:** 1 day | **Diagnostic Mid-assessment or Summative Assessment** |
| **Pre-Unit Assessment Targets** | Assess students’ ability to   * describe a set of data given a frequency distribution S.ID.2 * identify and calculate spread, center, shape, outliers, quartiles, mean, median, mode etc. S.ID.2 * construct and interpret a box plot S.ID.1 * compare, contrast and draw conclusions when given two data sets S.ID.3 |
| **Sample Activity** | [Mid-Assessment](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsR0V4ZDlyZ3dMZGs/edit) ([Answers](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsYmlfMWJTSU5UME0/edit)) |

|  |  |
| --- | --- |
| **Section 5:** 3 days | **Describing Variability with Deviations from the Mean** |
| **Mathematical Goals** | Students will...   * Describe variability by calculating deviations from the mean S.ID.2 * Compare two datasets with the same means but different variability, and contrast them by calculating the deviation of each data element from the mean S.ID.2 * Interpret sets with greater deviations as having greater variability S.ID.2 * Walk through a calculation of standard deviation for a small data set, and understand it as a typical deviation from the mean S.ID.2 |
| **Narrative overview of section**  (and how the standards are achieved) | In the previous section, students interpreted the meaning of the various measures of center. Measures of center are important, as they tell you with a single number what value is typical for a population. However, the nature of data involves variation, and *how much* data varies is an important question. In this section, students grapple with variability, the other major feature of measurements taken to answer a statistical question.  In middle school, they learned that IQR and mean absolute deviation are ways to describe spread. In this section, they’ll learn to calculate its more sophisticated cousin, standard deviation. They will start by looking at how much each data point deviates from the mean, and use these calculations to describe different data sets as more or less variable. Then, the process for calculating standard deviation should be stepped through. (Although the standards do not insist they learn to calculate it, doing so a few times will illustrate its meaning.) They’ll understand standard deviation as “typical distance from the mean,” and understand that higher values for standard deviation imply that a distribution is more spread out, whereas lower values imply that data is more clustered around the mean. |
| **Sample Activity 5.1** | [Measuring Variability in a Data Set](https://www.illustrativemathematics.org/illustrations/1887), Illustrative Mathematics  **WHAT:** In the first part of this task students compare and contrast the mean absolute deviation (MAD) calculations with the standard deviation calculations. Students build on this to calculate the standard deviation for a small data set about Jim’s test scores. In the last two parts of this task, students look at the test scores of other students and, after comparing them to Jim’s scores, draw conclusions about them based on their understanding of how to use the standard deviation as a measure of variation in data sets S.ID.3.  **WHY:** This is a great introductory problem for standard deviation calculations because:   * it connects and builds on students prior knowledge (MAD) * provides students with a small data set so the calculations are manageable by hand * provides an opportunity to compare standard deviation values S.ID.2 of multiple data sets in a real-world context |
| **Sample Activity 5.2** | [Understanding the Standard Deviation](https://www.illustrativemathematics.org/illustrations/1886), Illustrative Mathematics  **WHAT:** This task is conceptual rather than computational, and has students explore their understanding of the standard deviation as a measure of variability S.ID.2. There are four different parts to this task and each explores a different way of looking at/understanding the standard deviation. Students connect box plots to the standard deviation in the first part and create distribution graphs with specific standard deviation values in the second part. The third part has students examine two histograms, creating an explanation/argument as to which has the greater standard deviation MP.3. Lastly, the fourth part, has students reason abstractly and quantitatively to create number sets to fit specific givens about a standard deviation MP.2. The third and fourth parts are more difficult so students are encouraged to work in partners here.  **WHY:** From the IM website, “the purpose of this task is to deepen student understanding of the standard deviation as a measure of variability in a data distribution.” By having students focus on the conceptual understanding of the standard deviation as a measure of variability in four different ways, students develop a greater understanding of the meaning and application of the standard deviation both numerically and graphically. This provides students with an opportunity to look for and make use of structure while working with the standard deviation and variability MP.7. |
| **Target Standards** | S.ID.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.  S.ID.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** | MP.2, MP.3, MP.7 |



|  |  |
| --- | --- |
| **Section 6:** 1 day | **Variability for Skewed Distributions** |
| **Mathematical Goals** | Students will...   * Revisit the fact that median is better than mean as a measure of center for a skewed distribution S.ID.2 * Learn to create a 5-number summary, box plot, and compute the interquartile range using technology S.ID.2 * Interpet IQR as a description of variability better-suited to a skewed distribution S.ID.3 |
| **Narrative overview of section**  (and how the standards are achieved) | In this section students use technology to compute the IQR for two data sets. Students use this data to compare the data sets with specific emphasis on how to identify an outlier and the affects an outlier can have on a data set. |
| **Sample Activity 6.1** | [Identifying Outliers](https://www.illustrativemathematics.org/illustrations/1888), Illustrative Mathematics **WHAT:** In this task, students are given data about how far various students live from school. For the first set of data, students are given the Min/Max/Median/Mean/Q1/Q3 values and they must construct a box plot using these values. In a later part of the problem, one of the data points changes and students must calculate the new Min/Max/Median/Mean/Q1/Q3 values. Students can complete these calculations by hand or using technology MP.5. Students compare the calculated values with the original given values to analyze the statistical summary of the two data sets S.ID.2. The last parts of this task have students analyze and explain if certain data points are in fact outliers; students are asked to support their answers with appropriate work, providing an opportunity for students to construct viable arguments based on their mathematical calculations MP.3.  **WHY:** As stated by IM “The purpose of this task is to develop students understanding how extreme data points affect measures of center and how to use the mathematical definition to classify data points as “outliers” S-ID.3.” |
| **Target Standards** | S.ID.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.  S.ID.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** | MP.3, MP.5 |

|  |  |
| --- | --- |
| **Section 7:** 1 - 2 days | **Bringing it All Together** |
| **Mathematical Goals** | Students will...   * Starting with a bunch of raw data, represent it in different ways, and decide which way is most appropriate S.ID.1 * Select particular measures of center and spread appropriate to the shape of the distribution S.ID.3 * Compare and contrast two or more distributions by using appropriate measures to describe center, variability, and shape S.ID.2 |
| **Narrative overview of section**  (and how the standards are achieved) | In this section, students are presented with several data measures (mean, median, standard deviation, Q1, Q3, minimum & maximum values) for 5 sets of data. Students must decide how best to represent the data S.ID.1, interpret and understand differences in center and spread S.ID.2 and determine if there are any outliers S.ID.3 in order to accurately compare and contrast the given data sets. |
| **Sample Activity 7.1** | [Comparing Weather](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsSUZCMEhYQmxsN1E/edit), High Tech High, Jade White  **WHAT:** In this activity, students are given data regarding the average daily temperature in five U.S. cities. A claim is given based on the data provided and students must first read and interpret MP.1 the data before making models MP.4 and constructing their argument MP.3 as to whether or not they support the given claim.    **WHY:** This activity allows students to demonstrate their understanding of several statistics topics. In order to construct a sound argument, students must show strong evidence of understanding and proficiency of the statistics topics in the previous sections. Further, this problem allows students to tie together all of the topics learned in the previous 6 sections. |
| **Target Standards** | S.ID.1 Represent data with plots on the real number line (dot plots, histograms, and box plots).  S.ID.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.  S.ID.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** | MP.1, MP.3, MP.4 |

|  |  |
| --- | --- |
| **Section 8:** 2 days | **Two-way Frequency Tables** |
| **Mathematical Goals** | Students will...   * Interpret a table that divides data into different categories S.ID.5 * Understand that the choices made when organizing data can lead to different conclusions S.ID.5 |
| **Narrative overview of section**  (and how the standards are achieved) | Sometimes, it is illuminating to separate data into categories, especially when the category might influence the variable being measured. For example, in health studies, participants are often categorized as smokers vs non-smokers, or men vs women. This topic fits in this unit broadly because it’s another way to organize and interpret a list of data. |
| **Sample Activity 8.1** | [Musical Preferences](http://www.illustrativemathematics.org/illustrations/123), Illustrative Mathematics  **WHAT:** The 54 students in one of several middle school classrooms were asked two questions about musical preferences: “Do you like rock?” “Do you like rap?” The responses are summarized in a table. Students are asked several questions requiring them to understand and interpret the table, draw some conclusions, and explain them S.ID.5 MP.3. This problem introduces positive association and percentages based on the table.  **WHY:** The basic idea is for students to demonstrate that they know what it means for two variables to be associated: that if we knew someone were in one group (for example, they like rap), we now know more about their preferences for rock than if we knew nothing at all. This task also connects back to sample size and data collection from previous sections of this unit. |
| **Sample Activity 8.2** | [Lesson Plan](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsRjYwSXZVbmc4c3M/edit?usp=drive_web), [Teacher Note Outline & Example Problem](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsVW55RVdDeWtCbms/edit?usp=drive_web), [Example Practice Problems](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsemxhRFo4Q3ZhTjQ/edit?usp=drive_web), High Tech High, Jade White  **WHAT**: This activity allows students to understand two-way frequency tables based on data from their own class. It adds the ideas of joint, marginal and conditional frequencies S.ID.5.  **WHY**: This is largely an opportunity for procedural fluency practice. Working with data about themselves and their peers is motivating. |
| **Target Standards** | S.ID.5 Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data. |
| **Mathematical Practices** | MP.3 |

|  |  |
| --- | --- |
| **Section 9:** 1 day | **Summative Assessment** |
| **Unit Assessment Targets** | Assess students’ ability to   * calculate mean, median and mode S.ID.2 * create box-plots given data S.ID.1 * compare and contrast two frequency distributions S.ID.3 * articulate the differences and benefits of using mean and median for center S.ID.2 * read and interpret relative frequencies S.ID.5 |
| **Sample Activity** | [Unit Assessment](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsYVJqNUtVbktGUk0/edit) ([ANSWERS](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5CsS08wUEFDN0JiSFE/edit)\_frequency distributions missing from this copy)  [Assessment with Standards](https://docs.google.com/a/hightechhigh.org/file/d/0B784Ztw6k5Csb2x3bmx5alBGMWc/edit) BUT frequency distributions didn’t copy |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1.1 | 1.2 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 3.1 | 3.2 | 3.3 | 5.1 | 5.2 | 6.1 | 7.1 | 8.1 | 8.2 |
| S.ID.1 | **X** | **X** | **X** | **X** | **X** |  |  | **X** |  |  |  |  |  | **X** |  |  |
| S.ID.2 |  |  |  | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** |  |  |
| S.ID.3 |  |  |  |  | **X** | **X** | **X** | **X** | **X** | **X** | **X** |  | **X** | **X** |  |  |
| S.ID.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **X** | **X** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1.1 | 1.2 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 3.1 | 3.2 | 3.3 | 5.1 | 5.2 | 6.1 | 7.1 | 8.1 | 8.2 |
| MP.1 |  |  |  |  |  |  |  |  |  |  |  |  |  | **X** |  |  |
| MP.2 |  |  |  |  |  | **X** | **X** |  |  | **X** |  | **X** |  |  |  |  |
| MP.3 |  |  |  |  |  | **X** | **X** |  | **X** |  |  | **X** | **X** | **X** | **X** |  |
| MP.4 | **X** |  | **X** |  |  |  |  |  |  | **X** |  |  |  | **X** |  |  |
| MP.5 | **X** |  |  | **X** |  |  |  |  | **X** |  |  |  | **X** |  |  |  |
| MP.6 |  | **X** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| MP.7 |  |  |  |  |  |  |  |  |  |  |  | **X** |  |  |  |  |
| MP.8 |  |  |  |  |  | **X** | **X** |  |  |  |  |  |  |  |  |  |