A1 M2

UNIT PLAN # 2 Descriptive Statistics	
Title/Theme	Descriptive Statistics
Grade/Subject	CCSS Algebra 1
Length of Unit/Timeframe	5 weeks

Unit Summary

In this module, students reconnect with and deepen their understanding of statistics and probability concepts first introduced in Grades 6, 7, and 8. There is variability in data, and this variability often makes learning from data challenging. Students develop a set of tools for understanding and interpreting variability in data and begin to make more informed decisions from data. Students work with data distributions of various shapes, centers, and spreads. Measures of center and measures of spread are developed as ways of describing distributions.

A1 M2

Essential Questions:
How do we summarize, represent, and interpret data on a single count or measurement variable?
How do we summarize, represent, and interpret data on two categorical and quantitative variables?
What kind of data can be represented by linear models?

A1 M2

Standards – Content

S-ID.A.1 Represent data with plots on the real number line (dot plots, histograms, and box plots).★

S-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.★

S-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).★

Summarize, represent, and interpret data on two categorical and quantitative variables.

S-ID.B.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.★

S-ID.B.6 Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.★

a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. *Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models.*

b. Informally assess the fit of a function by plotting and analyzing residuals.

c. Fit a linear function for a scatter plot that suggests a linear association.

Interpret linear models.

S-ID.C.7 Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.★

S-ID.C.8 Compute (using technology) and interpret the correlation coefficient of a linear fit.★

S-ID.C.9 Distinguish between correlation and causation.★

Standards for Mathematical Practice

MP1 Make sense of problems and persevere in solving them.

MP2 Reason abstractly and quantitatively.

MP3 Construct viable arguments and critique the reasoning of others.

MP4 Model with Mathematics

MP5 Use appropriate tools strategically.

MP6 Attend to precision.

A1 M2

Enduring Understandings: Variables can be categorical or quantitative Data distributions will vary depending on the data It is important to consider outliers when analyzing data	**Skills (Students will be able to):** Collect and sort data Use appropriate tools to analyze the data
Formative and/or Alternate Assessments ❑ Section quizzes ❑ Making drawings and Illustrations of concepts. ❑ Predicting results ❑ Writing for mathematics - Finding information through the use of research. ❑ In class presentations ❑ Daily Student Reflection ❑ Evaluating arguments	❑ Oral Recitation ❑ Formulating feedback ❑ Group Reports ❑ Assessment through guided questioning ❑ Summarizing data ❑ Arguing logically ❑ Critiquing explanations ❑ Pretest (Written) ❑ Post Test (Written)
Summative Assessment(s): ❑ Unit test	

UNIT READINGS AND VOCABULARY

- Appendix A CCSS Traditional Pathway Algebra 1.
- http://greatminds.net/maps/images/math_documents/G6-M1_Teacher_Edition.pdf

ESSENTIAL UNIT VOCABULARY

Skewed Data Distribution, Outlier, Sample Standard Deviation, Interquartile Range, statistical association, Conditional Relative Frequency, Residual Plot, Residual, Correlation Coefficient

INSTRUCTIONAL ACTIVITIES

❑ Constructing	❑ Language
❑ Reading	❑ Vocabulary
❑ Writing	❑ Viewing
❑ Speaking/Listening	❑ Critical Thinking
❑ Interpreting	❑ Peer Critique
❑ Research	❑ Illustrating
❑ Evaluating	❑ Outline
❑ Accountable Talk/Socratic Forum	❑ Graphic Organizers
	❑ Performance

Instructional Resources

❑ Internet	❑ www.jmap.org
❑ Math Articles	❑ www.Regentsprep.org
❑ Math Journals	❑ https://www.khanacademy.org/www.mrwana
❑	❑

Differentiated Instruction

A1 M2

- **Look at data to see where individual students are strong and weak**
- **Look at item analysis to adjust foci**
- **Design tasks that allow multiple entry points**
- **Use problem solving strategy to help students get started**
- **How can we help students understand scientific language?**
- **Use of diagrams and other learning aids**
- **Understand what students are thinking**
- **Problem solving and discussion time among students**
- **Different approaches and views of the same problem**

ESL Modifications:

Flashcards for Vocabulary enrichment
Directed Reading
Predictive Reading

Special Education Modifications:

IEP: Extended time for quiz/test
Special Needs Activities and Modified Tests
Questions should be read aloud
Simple demonstration for students to see concrete results

Sequential Progression by Topic

Topic A	**Topic B**
Topic A Shapes and Centers of Distributions (S-ID.A.1, S-ID.A.2, S-ID.A.3, Distributions and Their Shapes, Describing the Center of a Distribution, & Estimating Centers and Interpreting the Mean as a Balance Point .	Topic B: Describing Variability and Comparing Distributions (S-ID.A.1, S-ID.A.2, S-ID.A.3, Summarizing Deviations from the Mean, Measuring Variability for Symmetrical Distributions, Interpreting the Standard Deviation, Measuring Variability for Skewed Distributions (Interquartile Range, & Comparing Distributions.

Mid-Module Assessment and Rubric .....
Topics A through B (assessment 2 days, return and remediation or further applications 4 days)

Topic C	**Topic D**
Categorical Data on Two Variables (S-ID.B.5, S-ID.C.9, Summarizing Bivariate Categorical Data , Summarizing Bivariate Categorical Data with Relative Frequencies, Conditional Relative Frequencies and Association.	Numerical Data on Two Variables (S-ID.B.6, S-ID.C.7, S-ID.C.8, S-ID.C.9), Relationships Between Two Numerical Variables, Modeling Relationships with a Line, Interpreting Residuals from a Line , More on Modeling Relationships with a , Analyzing Residuals, Interpreting Correlation

End-of-Module Assessment and Rubric .....
Topics A through D (assessment 2 days, return and remediation or further applications 4 days)

A1 M2

Curriculum Map High School Courses

	Grade 9 -- Algebra I	Grade 10 -- Geometry	Grade 11 -- Algebra II	Grade 12 -- Precalculus	
20 days	M1: Relationships Between Quantities and Reasoning with Equations and Their Graphs (40 days)	M1: Congruence, Proof, and Constructions (45 days)	M1: Polynomial, Rational, and Radical Relationships (45 days)	M1: Complex Numbers and Transformations (40 days)	20 days
20 days					20 days
20 days	M2: Descriptive Statistics (25 days)	M2: Similarity, Proof, and Trigonometry (45 days)	M2: Trigonometric Functions (20 days)	M2: Vectors and Matrices (40 days)	20 days
20 days	M3: Linear and Exponential Functions		M3: Functions (45 days)		20 days
20 days	State Examinations (35 days)	State Examinations	State Examinations	State Examinations	20 days
20 days	M4: Polynomial and Quadratic Expressions, Equations and Functions (30 days)	M3: Extending to Three Dimensions (10 days)	M4: Inferences and Conclusions from Data (40 days)	M3: Rational and Exponential Functions (25 days)	20 days
20 days		M4: Connecting Algebra and Geometry through Coordinates (25 days)		M4: Trigonometry (15 days)	20 days
20 days	M5: A Synthesis of Modeling with Equations and Functions (20 days)	M5: Circles with and Without Coordinates (25 days)		M5: Probability and Statistics (30 days)	20 days
20 days	Review and Examinations	Review and Examinations	Review and Examinations	Review and Examinations	20 days

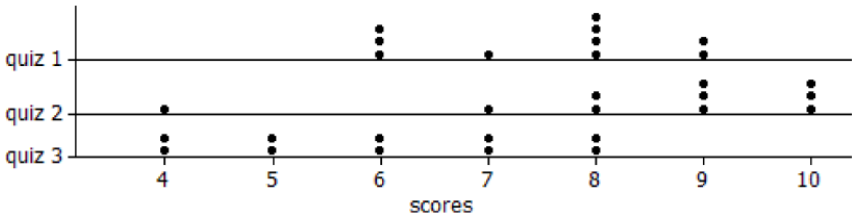
Key: Number and Quantity and Modeling; Geometry and Modeling; Algebra and Modeling; Statistics and Probability and Modeling; Functions and Modeling

A1 M2 Mid Module Assessment P1

A1 M2

Name _____ Date _____

1. The scores of three quizzes are shown in the following data plot for a class of 10 students. Each quiz has a maximum possible score of 10. Possible dot plots of the data are shown below.



a. On which quiz did students tend to score the lowest? Justify your choice.

b. Without performing any calculations, which quiz tended to have the most variability in the students' scores? Justify your choice based on the graphs.

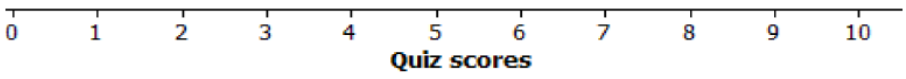
A1 M2

A1 M2 Mid Module Assessment P2

c. If you were to calculate a measure of variability for Quiz 2, would you recommend using the interquartile range or the standard deviation? Explain your choice.

d. For Quiz 3, move one dot to a new location so that the modified data set will have a larger standard deviation than before you moved the dot. Be clear which point you decide to move, where you decide to move it, and explain why.

e. On the axis below, arrange 10 dots, representing integer quiz scores between 0 and 10 so that the standard deviation is the largest possible value that it may have. You may use the same quiz score values more than once.



A1 M2

A1 M2 Mid Module Assessment P3

Use the following definitions to answer questions (f)–(h).

- The *midrange* of a data set is defined to be the average of the minimum and maximum values: $\frac{\text{min} + \text{max}}{2}$.
- The *midhinge* of a data set is defined to be the average of the first quartile (Q_1) and the third quartile (Q_3): $\frac{Q_1+Q_3}{2}$.

f. Is the midrange a measure of center or a measure of spread? Explain.

g. Is the midhinge a measure of center or a measure of spread? Explain.

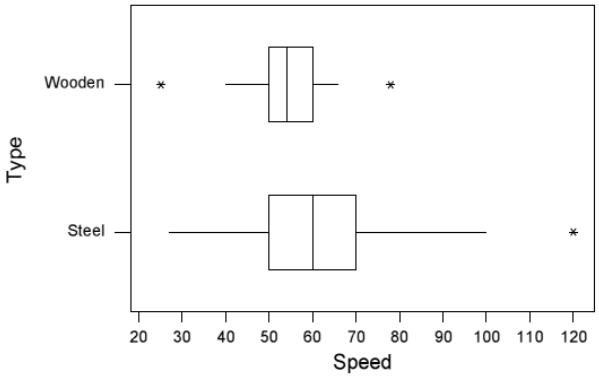
h. Suppose the lowest score for Quiz 2 was changed from 4 to 2, and the midrange and midhinge are recomputed. Which will change more?

A. Midrange
B. Midhinge
C. They will change the same amount.
D. Cannot be determined

A1 M2

A1 M2 Mid Module Assessment P4

2. The box plots below display the distributions of maximum speed for 145 roller coasters in the United States, separated by whether they are wooden coasters or steel coasters.



Based on the box plots, answer the following questions or indicate that you do not have enough information.

a. Which type of coaster has more observations?

A. Wooden
B. Steel
C. About the same
D. Cannot be determined

Explain your choice:

b. Which type of coaster has a higher percentage of coasters that go faster than 60 mph?

A. Wooden
B. Steel
C. About the same
D. Cannot be determined

Explain your choice:

A1 M2

A1 M2 Mid Module Assessment P5

c. Which type of coaster has a higher percentage of coasters that go faster than 50 mph?

A. Wooden
B. Steel
C. About the same
D. Cannot be determined

Explain your choice:

d. Which type of coaster has a higher percentage of coasters that go faster than 48 mph?

A. Wooden
B. Steel
C. About the same
D. Cannot be determined

Explain your choice:

e. Write 2–3 sentences comparing the two types of coasters with respect to which type of coaster normally goes faster.

A1 M2

A1 M2 End Module Assessment P1

Name _____ Date _____

1. A recent social survey asked 654 men and 813 women to indicate how many "close friends" they have to talk about important issues in their lives. Below are frequency tables of the responses.

Number of Close Friends	0	1	2	3	4	5	6	Total
Males	196	135	108	100	42	40	33	654
Females	201	146	155	132	86	56	37	813

a. The shape of the distribution of the number of close friends for the males is best characterized as

A. Skewed to the higher values (right or positively skewed).
B. Skewed to the lower values (left or negatively skewed).
C. Symmetric.

b. Calculate the median number of class friends for the females. Show your work.

c. Do you expect the mean number of close friends for the females to be larger or smaller than the median you found in part (b), or do you expect them to be the same? Explain your choice.

d. Do you expect the mean number of close friends for the males to be larger or smaller than the mean number of close friends for the females, or do you expect them to be the same? Explain your choice.

A1 M2

A1 M2 End Module Assessment P2

2. The physician's health study examined whether physicians who took aspirin were less likely to have heart attacks than those who took a placebo (fake) treatment. The table below shows their findings.

	Placebo	Aspirin	Total
Heart attack	189	104	293
No heart attack	10,845	10,933	21,778
Total	11,034	11,037	22,071

Based on the data in the table, what conclusions can be drawn about the association between taking aspirin and whether or not a heart attack occurred? Justify your conclusion using the given data.

A1 M2

A1 M2 End Module Assessment P3

3. Suppose 500 high school students are asked the following two questions:

   - What is the highest degree you plan to obtain? (check one)
     ☐ High school degree ☐ College (Bachelor's degree)
     ☐ Graduate school (e.g., Master's degree or higher)

   - How many credit cards do you currently own? (check one)
     ☐ None ☐ One ☐ More than one

   Consider the data shown in the following frequency table.

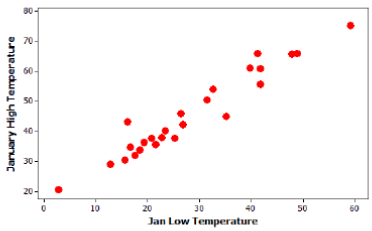
	No Credit Cards	One Credit Card	More than One Credit Card	Total
High school	?		6	59
College	120	240	40	394
Graduate school				47
Total		297		500

Fill in the missing value in the cell in the table that is marked with a "?" so that the data would be consistent with no association between education aspiration and current number of credit cards for these students. Explain how you determined this value.

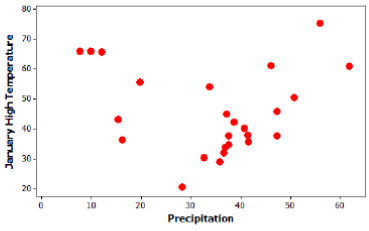
A1 M2

A1 M2 End Module Assessment P4

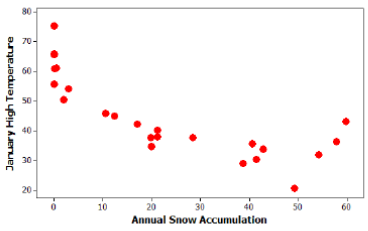
4. Weather data were recorded for a sample of 25 American cities in one year. Variables measured included January high temperature (in degrees Fahrenheit), January low temperature (in degrees Fahrenheit), annual precipitation (in inches), and annual snow accumulation. The relationships for three pairs of variables are shown in the graphs below (Jan. Low Temperature—Graph A; Precipitation—Graph B; Annual Snow Accumulation—Graph C).



Graph A



Graph B



Graph C

a. Which pair of variables will have a correlation coefficient closest to 0?

A. Jan. high temperature and Jan. low temperature
B. Jan. high temperature and precipitation
C. Jan. high temperature and snow accumulation

Explain your choice:

b. Which of the above scatterplots would be best described as a strong nonlinear relationship? Explain your choice.

A1 M2

A1 M2 End Module Assessment P5

A1 M2

c. Suppose we fit a least squares regression line to Graph A. Circle one word choice for each blank that best completes this sentence, based on the equation:

If I compare a city with a January low temperature of 30°F to a city with a higher January low temperature, then the ___(1)___ January high temperature of the second city will _____(2)_____ be _____(3)___.

(1) actual, predicted

(2) probably, definitely

(3) smaller, larger, the same, equally likely to be higher or lower

d. For the city with a January low temperature of 30°F, what do you predict for the annual snow accumulation? Explain how you are estimating this based on the three graphs above.

A1 M2

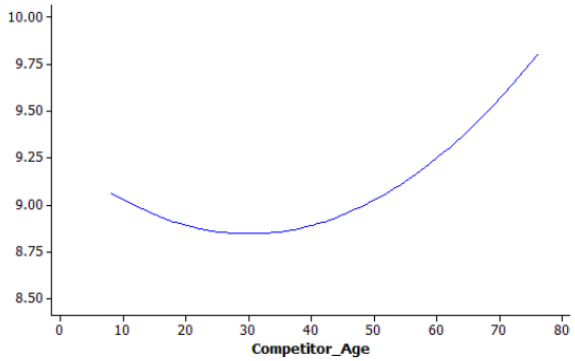
5. Suppose times (in minutes) to run one mile were recorded for a sample of 100 runners, ages 16–66 years, and the following least squares regression line was found.

   Predicted time in minutes to run one mile $= 5.35 + 0.25 \times (age)$

   a. Provide an interpretation in context for this slope coefficient.

   b. Explain what it would mean in the context of this study for a runner to have a negative residual.

   c. Suppose, instead, that someone suggests using the following curve to predict time to run one mile. Explain what this model implies about the relationship between running time and age, and why that relationship might make sense in this context.



A1 M2

d. Based on the results for these 100 runners, explain how you could decide whether the first model or the second model provides a better fit to the data.

e. The sum of the residuals is always equal to zero for the least squares regression line. Which of the following must also always be equal to zero?

A. The mean of the residuals
B. The median of the residuals
C. Both the mean and the median of the residuals
D. Neither the mean nor the median of the residuals

A1 M2

Rubric A1 M2

A Progression Toward Mastery				
Assessment Task Item	**STEP 1** Missing or incorrect answer and little evidence of reasoning or application of mathematics to solve the problem.	**STEP 2** Missing or incorrect answer but evidence of some reasoning or application of mathematics to solve the problem.	**STEP 3** A correct answer with some evidence of reasoning or application of mathematics to solve the problem or an incorrect answer with substantial evidence of solid reasoning or application of mathematics to solve the problem.	**STEP 4** A correct answer supported by substantial evidence of solid reasoning or application of mathematics to solve the problem.