Davison Community Schools
ADVISORY CURRICULUM COUNCIL
Phase II, December 8, 2014

AP Statistics

Course Essential Questions: How can the discipline of statistics be used to search for patterns and relationships in real-life?

Unit 1: Exploring Data: describing patterns and departures from patterns - Exploring Data

Essential Questions:	Essential Understanding:
• What is statistics and why should we study it? • How is data analysis used to explore categorical data? • What are different ways to display the distribution of a variable and how do you determine which is the best to use? • How do we use and interpret statistical summaries to describe quantitative data? • How do we identify where a value is located in a distribution? • What role does the Normal curve play into approximating distributions of data? • What are the important characteristics that describe the association between two variables in a scatterplot?	• Statistics is the science of learning from data. • Statistics gives us a language for talking about uncertainty that is understood by statistically literate people everywhere. • The most important information about any statistical study is how the data were produced. • A carefully chosen graph is often more instructive than a bunch of numbers. • Percentiles are used to locate a value in a distribution. • Correlation is a numerical summary that provides the strength of a linear association. • A least-square regression line models the relationship between variables in a scatter plot.

Curriculum Standards- DOK noted where applicable with Standards

I. Exploring Data
- Constructing and interpreting graphical displays of distributions of univariate data
- Summarizing distributions of univariate data
- Comparing distributions of univariate data
- Exploring bivariate
- Exploring categorical data

III. Anticipating Patterns
- The Normal distribution

LEARNING TARGETS

Knowledge/Content I Know …	Skills/Processes I Can …

● Individuals are the objects described by a set of data. ● A variable. is any characteristic of an individual. ● A categorical variable places an individual into one of several groups or categories. ● A quantitative variable takes a numerical values for which it makes sense to find an average. ● That the distribution of a variable tells us what values the variable takes and how often it takes these values. ● That a frequency table displays the counts of a variable. ● The components of a bar graph. ● The components of a histogram. ● The components of a pie chart. ● The components of a stemplot. ● The marginal distribution of one of the categorical variables in a two-way table of counts is the distribution of values of that variable among all individuals described by the table. ● A conditional distribution of a variables describes the value of that variable among individuals who have a specific value of another variable. ● That there is an association between two variables if knowing the value of one variable helps predict the value of the other. ● That when you examine the distribution of a quantitative variable you should look for the overall pattern and for striking departures. ● You can describe the overall pattern can be described as shape, center, and spread. ● An important kind of departure is an outlier, an individual value that falls outside the overall pattern. ● The measures of center. (Mean, Median, Mode) ● The different shapes of distributions. (left skewed, right skewed, uniform, symmetric) ● The different measures of spread. (range, standard deviation, variation) ● When to use which measure of center to represent a data set. ● The components of a boxplot ● The process for calculating the quartiles and interquartiles. ● What makes up the 5-number summary. ● That a percentile represents the value with p percent of the observations less than it.	● Identify the individuals and variables in a set of data **(intro)** ● Classify variables as categorical or quantitative**(intro)** ● Display categorical data with a bar graph. **1.1** ● Decide whether it would be appropriate to make a pie chart **1.1** ● Identify what makes some graphs of categorical data deceptive **1.1** ● Calculate and display the marginal distributions of a categorical variable from a two-way table. **1.1** ● Calculate and display the conditional distribution of a categorical variable for a particular value of the other categorical variable in a two-way table **1.1** ● Describe the association between two categorical variables by comparing appropriate conditional distributions **1.1** ● Make and interpret dotplots and stemplots of quantitative data. **1.2** ● Describe the overall pattern of a distribution and identify any major departures from the pattern **1.2** ● Identify the shape of a distribution from a graph as roughly symmetric or skewed **1.2** ● Make and interpret histograms or quantitative data **1.2** ● Compare distributions of quantitative data using dotplots, stemplots, or histograms **1.2** ● Calculate measures of center **1.3** ● Calculate and interpret measures of spread **1.3** ● Choose the most appropriate measure of center and spread in a given setting **1.3** ● Identify outliers using the 1.5 X IQR rule **1.3** ● Make and interpret boxplots or quantitative data **1.3** ● Use appropriate graphs and numerical summaries to compare distributions of quantitative variables. **1.3** ● Find and Interpret the percentile of an individual value within a distribution of data. **2.1** ● Estimate percentiles and individual values using a cumulative relative frequency graph **2.1** ● Find and interpret the standardized score of an individual value within a distribution of data **2.1** ● Describe the effect of adding, subtracting, multiplying by, or dividing by a constant on the shape, center, and spread of a distribution of data. **2.1** ● Estimate the relative locations of the median and mean on a density curve **2.2**

• The components of a cumulative relative frequency graph. • How the mean, median, and mode are located depending on the different shapes of the density curves. • That the z-score tells us how many standard deviations from the mean an observation falls, and in what direction. • The process for computing the z-score. • How adding, subtracting, multiplying by, or dividing by a constant will affect the shape, center, and spread of a distribution of data. • What a density curve is and characteristics of it. • A Normal distribution is described by a Normal Density curve. • Why Normal distributions are important in Statistics. • The 68-95-99.7 rule • A response variable measures an outcome of a study. • An explanatory variable may help explain or predict changes in a response variable. • A scatterplot shows the relationship between two quantitative variables measured on the same individuals. • The basic properties of correlation *r*, including how the correlation is influenced by outliers. • A regression line is a line that describes how a response variable changes as an explanatory variables changes. • Extrapolation is the use of a regression line for prediction far outside the interval of values of the explanatory variable x used to obtain the regression line. • A residual is the difference between an observed value of the response variable and the value predicted by the regression line. • The least-squares regression line of the response variable on the explanatory variable is the line that makes the sum of the squared residuals as small as possible. • The components for a residual plots and the process to assess whether a linear model is appropriate. • The dangers of extrapolation. • How the slope, y intercept, standard deviation of the residuals, and r squared are influenced by outliers. • Why association does not imply causation between an explanatory and response variable.	• Use the 68-95-99.7 rule to estimate areas in a Normal distribution **2.2** • Find the proportion of z-values in a specified interval **2.2** • Find a z-score from a percentile in the standard Normal distribution **2.2** • Determine whether a distribution of data is approximately Normal from graphical and numerical evidence **2.2** • Find the areas in any normal distribution using Table A or technology **2.2** • Find values from areas in any normal distribution using Table A or technology. **2.2** • Identify explanatory and response variables in situations where one variable helps to explain or influences the other. **3.1** • Make a scatterplot to display the relationship between two quantitative variables **3.1** • Describe the direction, form, and strength of a relationship displayed in a scatterplot and recognize outliers in a scatterplot. **3.1** • Interpret the correlation **3.1** • Use technology to calculate correlation **3.1** • Explain why association does not imply causation **3.1** • Interpret the slope and y intercept of a least-squares regression line **3.2** • Use the least-squares regression line to predict y for a given x. **3.2** • Explain the dangers of extrapolation. **3.2** • Calculate and interpret residuals **3.2** • Explain the concept of least squares **3.2** • Determine the equation of a least-squares regression line using technology. **3.2** • Construct and interpret residual plots to assess whether a linear model is appropriate **3.2** • Interpret the standard deviation of the residuals and r square. **3.2** • Use the residuals and r square to assess how well the least-squares regression line models the relationship between two variables. **3.2** • Describe how the slope, y intercept, standard deviation of the residuals, and r squared are influenced by outliers. **3.2** • Find the slope and y intercept of the least squares regression line from the means and standard deviations of x and y and their correlation. **3.2**

Phase III Textbook/Materials

Phase IV Summative Assessment Evidence

Common Summative Unit Assessments: (*identifies Performance Task)	**Agreed Upon Interim Summative Assessments:** (*identifies Performance Task)

Phase V Learning Plan

Unit 2: Sampling and experimentation: Planning and conducting a study

Essential Questions:	Essential Understanding:
• What is the difference between a population and a sample? • What are unbiased methods of sampling and describe the appropriateness of each? • What characteristics make up a well-designed experiment? • How does the data produced in a study determine the different conclusions that can be made?	• Because a census is almost always infeasible, it is important for a researcher to understand what multiple methods are and when to use them for gathering data. • Well-designed experiments should contain a comparison, random assignment, control, and replication. • Depending on the data that is collected, one could make inferences about populations or conclude cause and effect.

Curriculum Standards- DOK noted where applicable with Standards

II. Sampling and Experimentation

- Planning and conducting surveys and experiments using appropriate methods of data collection
- Generalizability of results and types of conclusions that can be drawn from observational studies, experiments, and surveys

LEARNING TARGETS

Knowledge/Content **I Know …**	**Skills/Processes** **I Can …**
• The population in a statistical study is the entire group of individuals we want information about. • A census collects data from every individual in the population • A sample is a subset of individuals in the population from which we actually collect data. • Ways to sample badly: Convenience and voluntary response samples. • Ways to sample well: Random sampling, Simple Random Sampling. • The steps in choosing a simple random sample with technology. • Other random sampling methods and the advantages and disadvantages of each. (ex: cluster, simple random, stratified, systematic) • The design of a statistical study shows bias if it would consistently underestimate or consistently overestimate the value you want to know.	• Identify the population and sample in a statistical study **4.1** • Identify voluntary response samples and convenience samples. **4.1** • Explain how sampling methods can lead to bias **4.1** • Describe how to obtain a random sample using slips of paper, technology, or a table of random digits. **4.1** • Distinguish a simple random sample from a stratified random sample or cluster sample. **4.1** • Give the advantages and disadvantages of each sampling method. **4.1** • Explain how undercoverage, non response, question wording, and other aspects of a sample survey can lead to bias. **4.1** • Distinguish between an observational study and an experiment. **4.2**

- Undercoverage occurs when some members of the population cannot be chosen in a sample.
- Nonresponse occurs when an individual chosen for the sample can't be contacted or refuses to participate.
- An Observational study observes individuals and measures variables of interest but does not attempt to influence the responses.
- An experiment deliberately imposes some treatment on individuals to measure their responses.
- Confounding occurs when two variables are associated in such a way that their effects on a response variable cannot be distinguished from each other.
- How confounding effects drawing conclusions.
- A specific condition applied to the individuals in an experiment is called a treatment.
- How to experiment badly.
- How to experiment well.
- In an experiment, random assignment means that experimental units are assigned to treatments using a chance process.
- Know the difference between explanatory and response variables.
- The Principles of an experimental design. (Comparison, Random assignment, Control, Replication)
- In statistics, replication means "use enough subjects."
- In a completely randomized design, the experimental units are assigned to the treatments completely by chance.
- What the purpose of a control group in an experiment is.
- The placebo effect is the response to a dummy treatment and its purpose in experiments.
- In a double blind experiment, neither the subjects nor those who interact with them and measure the response variable know which treatment a subject received.
- An observed effect so large that it would rarely occur by chance is called statistically significant.
- A block is a group of experimental units that are known before the experiment to be similar in some way that is expected to affect the response to the treatments.

- Explain the concept of confounding and how it limits the ability to make cause-and-effect conclusions. **4.2**
- Identify the experimental units, explanatory and response variables, and treatments in an experiment. **4.2**
- Explain the purpose of comparison, random assignment, control, and replication in an experiment. **4.2**
- Describe a completely randomized design for an experiment. **4.2**
- Describe the placebo effect and the purpose of blinding in an experiment **4.2**
- Interpret the meaning of statistically significant in the context of an experiment. **4.2**
- Explain the purpose of blocking in an experiment. **4.3**
- Describe a randomized block design or a matched pairs design for an experiment **4.3**
- Describe the scope of inference that is appropriate in a statistical study. **4.3**

• In a randomized block design, the random assignment of experimental units to treatments is carried out separately within each block. • The process for creating a randomized block design. • A matched pairs design is a common form of blocking for comparing just two treatments. • The basic data ethics. (institutional review, informed consent, and confidentiality)	

Phase III Textbook/Materials

Phase IV Summative Assessment Evidence

Common Summative Unit Assessments: (*identifies Performance Task)	Agreed Upon Interim Summative Assessments: (*identifies Performance Task)

Phase V Learning Plan

Unit 3: Anticipating patterns: exploring random phenomena using probability and simulation.

Essential Questions:	Essential Understanding:
• How does probability play into the role of bridging designing a study and drawing conclusions from them. • How can the way that probability distributions are displayed aid in their interpretation? • What is the difference between discrete and continuous variables? • What are the two most common types of discrete random variables and what does each count? • What is the difference between a statistic and a parameter? • What is the difference between the distribution of the population, the distribution of the sample, and the sampling distribution of a sample statistic? • How does the size of the sample affect the variability and shape of a distribution?	• The probability of an outcome in a chance process describes the proportion of times the outcome would occur in the long run. • Binomial random variables count the number of successes in a fixed number of trials, while geometric random variables count the number of trials needed to achieve a single success. • A parameter is a number that describes some characteristics of a population. A statistic estimates the value of a parameter using a sample from the population. • Statistics vary. • To adequately describe a sampling distribution, one needs to address shape, center, and spread.

Curriculum Standards- DOK noted where applicable with Standards
III. Anticipating Patterns • Probability • Combining independent random variables • The Normal distribution • Sampling distributions
LEARNING TARGETS

Knowledge/Content I Know …	Skills/Processes I Can …
• The probability of any outcome of a chance process is a number between 0 and 1. • Probability describes the proportion of times the outcome would occur in a very long series of repetitions. • The steps for performing a simulation. • The sample space of a chance process is the set of all possible outcomes. • A probability model is a description of some chance process that consists of two parts: a	• Interpret probability as a long-run relative frequency. **5.1** • Use simulation to model chance behavior. **5.1** • Describe a probability model for a chance process **5.2** • Use the complement rule. **5.2** • Use the addition rule for mutually exclusive events. **5.2**

sample space and a probability for each outcome. ● An event is any collection of outcomes from some chance process. ● Two events are mutually exclusive if they have no outcomes in common and so can never occur together. ● The complement rule. ● The addition rule for mutually exclusive events. ● The general addition rule for calculating probabilities for two events. ● The probability that one event happens given that another event is already known to have happened is called a conditional probability. ● The formulas for computing conditional probabilities. ● The general multiplication rule for probabilities. ● Two events are independent if the occurrence of one event does not change the probability that the other event will happen. ● The multiplication rule for independent events. ● A random variable takes numerical values that describe the outcomes of some chance process. ● The probability distribution of a random variable gives its possible values and their probabilities. ● The process for computing the expected value of a discrete random variable. ● How to use a probability distribution to compute probabilities of discrete and continuous random variables. ● Discrete random variables arise from counting something. ● A continuous random variables arise from measuring something, takes all values in an interval of numbers and the probability distribution of it is described by a density curve. ● A binomial setting arises when we perform several independent trials of the same chance process and record the number of times that a particular outcome occurs. ● The conditions for using a binomial distribution known as BINS but also including the 10% and large count condition. ● The binomial probability formula. ● The steps for finding binomial probabilities. ● The key strokes for computing a binomial probability on the calculator.	● Use a two-way table or Venn diagram to model a chance process and calculate probabilities involving two events. **5.2** ● Use the general addition rule to calculate probabilities. **5.2** ● Calculate and interpret conditional probabilities **5.3** ● Use the general multiplication rule to calculate probabilities **5.3** ● Use tree diagrams to model a chance process and calculate probabilities involving two or more events. **5.3** ● Determine if two events are independent. **5.3** ● When appropriate, use the multiplication rule for independent events to compute probabilities. **5.3** ● Compute probabilities using the probability distribution of a discrete random variable. **6.1** ● Calculate and interpret the mean (expected value) of a discrete random variable. **6.1** ● Calculate and interpret the standard deviation of a discrete random variable. **6.1** ● Compute probabilities using the probability distribution of certain continuous random variables. **6.1** ● Describe the effects of transforming a random variable by adding and subtracting a constant and multiplying or dividing by a constant. **6.2** ● Find the mean and standard deviation of the sum or difference of independent random variables. **6.2** ● Find probabilities involving the sum or difference of independent Normal random variables. **6.2** ● Determine whether the conditions for using a binomial random variable are met. **6.3** ● Computer and interpret probabilities involving binomial distributions. **6.3** ● Calculate the mean and standard deviation of a binomial random variable. **6.3** ● Interpret the mean and standard deviation in context. **6.3** ● Find probabilities involving geometric random variables **6.3** ● When appropriate, use Normal approximation to the binomial distribution to calculate probabilities. **6.3** ● Distinguish between a parameter and a statistic. ● Use the sampling distribution of a statistic to evaluate a claim about a parameter. **7.1**

• The formulas for computing the mean and standard deviation for a binomial situation. • A geometric setting arises when we perform independent trials of the same chance process and record the number of trials it takes to get the first success. • The conditions for using a geometric probability model. • The formula for computing geometric probability. • When the normal approximation can be used to represent a binomial situation. • A parameter is a number that describes some characteristic of the population. • A statistic is a number that describes some characteristic of a sample. • The sampling distribution of a statistic is the distribution of values taken by the statistic in all possible samples of the same size from the same population. • The relationship between sample size and the variability of a statistic. • The formulas for finding the mean and standard deviation of the sampling distribution of a sample proportion. • The formulas for finding the mean and standard deviation of the sampling distribution of a sample mean. • The Central Limit Theorem says that when a sample is large, the sampling distribution of the sample mean is approximately Normal. • Normal/Large sample condition for sample means. • How the shape of the sampling distribution of the sample mean is affected by the shape of the population distribution and the sample size.	• Distinguish among the distribution of a population, the distribution of a sample and the sampling distribution of a statistic. **7.1** • Determine whether or not a statistic is an unbiased estimator of a population parameter. **7.1** • Describe the relationship between sample size and the variability of a statistic. **7.1** • Find the mean and standard deviation of the sampling distribution of a sample proportion. **7.2** • Determine if the sampling distribution of the sample proportion is approximately normal. **7.2** • If appropriate, use a Normal distribution to calculate probabilities involving the sample proportion. **7.2** • Find the mean and standard deviation of the sampling distribution of a sample mean. **7.3** • Explain how the shape of the sampling distribution of the sample mean is affected by the shape of the population distribution and the sample size. **7.3** • If appropriate, use a Normal distribution to calculate probabilities involving the sample mean. **7.3**

Phase III Textbook/Materials
Phase IV Summative Assessment Evidence

Common Summative Unit Assessments: (*identifies Performance Task)	**Agreed Upon Interim Summative Assessments:** (*identifies Performance Task)

Phase V Learning Plan

Unit 4: Statistical Inference: estimating population parameters and testing hypotheses.

Essential Questions:	**Essential Understanding:**
• What conditions do you need to check when calculating a confidence interval for a population proportion? • How do you determine the appropriate sample size when planning a study? • What are the steps of conducting a significance test? • How are confidence intervals and significance tests used to compare two proportions or two means?	• A point estimate is the single best guess for the value of a population parameter. • A confidence interval is an interval of plausible values for a parameter. • The Four step process is ideally suited for confidence interval questions and should be used each time one calculates and interprets a confidence interval. • The difference between two sample proportions varies due to the chance involved in random selection is essential for drawing conclusions based on an observed difference in proportions in a study. • The chi-square statistic is a way to measure the difference between an observed distribution and a hypothesized distribution of categorical data.

Curriculum Standards- DOK noted where applicable with Standards

IV. Statistical Inference

- Estimating population parameters and testing hypotheses
- Tests of significance

LEARNING TARGETS

Knowledge/Content **I Know …**	**Skills/Processes** **I Can …**
• A point estimator is a statistic that provides an estimate of a population parameter. • A confidence interval gives an interval of plausible values for a parameter. • How the sample size and confidence level affect the length of a confidence interval. • How practical issues can affect the interpretation of a confidence interval. • The conditions for constructing a confidence interval for a population mean. • The four step process for creating a confidence interval. (State, Plan, Do, Conclude) • Conditions for constructing a confidence interval about a proportion.	• Determine the point estimate and margin of error from a confidence interval. **8.1** • Interpret a confidence interval in context. **8.1** • Interpret a confidence level in context. **8.1** • Describe how the sample size and confidence level affect the length of a confidence interval. **8.1** • Explain how practical issues can affect the interpretation of a confidence interval. **8.1** • State and check the Random, 10%, and Large Counts conditions for constructing a confidence interval for a population proportion. **8.2** • Determine critical values for calculating a Confidence interval for a population proportion. **8.2**

• When the standard deviation of a statistic is estimated from data, the result is called the standard error of the statistic. • The formula for computing a one-sample z-interval for a population proportion. • The formula for determining the necessary sample size that will yield a specific confidence interval. • How the t distributions are different from the standard Normal distribution. • Why it is necessary to use a t distribution when calculating a confidence interval for a population mean. • The formula for computing the standard error of the sample mean. • The process for choosing sample size for a desired margin of error when estimating the population mean. • The claim we weigh evidence against in a statistical test is called the null hypothesis. • The claim about the population that we are trying to find evidence for is the alternative hypothesis. • The difference between Type 1 and Type 2 errors and how they affect our inference about a hypothesis. • The conditions for performing a significance test about a population parameter. • What factors affect the power of a test. • The relationship among the probability of a Type 1 error, the probability of a Type 2 error, • The conditions for performing a significant test about a population mean. • The relationship among the probability of a Type 1 error, the probability of a Type 2 error, and the power of a test. • The conditions for doing inference about sample proportions. • The conditions for doing inference about the difference of sample means. • When it is appropriate to use two-sample t procedures versus paired t procedures. • A Chi-Square statistic is a measure of how far the observed counts are from the expected counts. • The conditions for performing a chi-square test for goodness of fit. • The process for computing the chi-square statistic, degrees of freedom, and P-value for a chi-square test for goodness of fit. • The conditions for performing inference about the slope of the population regression line.	• Construct and interpret a confidence interval for a population proportion. **8.2** • Determine the sample size required to obtain a C% confidence interval for a population proportion with a specified margin of error. **8.2** • State and check the conditions for constructing a confidence interval for a population mean. **8.3** • Explain how the t distributions are different from the standard Normal distribution. **8.3** • Explain why it is necessary to use a t distribution when calculating a confidence interval for a population mean. **8.3** • Determine critical values for calculating a Confidence interval for a population mean using a table or technology. **8.3** • Construct and interpret a confidence interval for a population mean. **8.3** • Determine the sample size required to obtain a C% confidence interval for a population mean with a specified margin of error. **8.3** • State the null and alternative hypotheses for a significance test about a population parameter. **9.1** • Interpret a P-value in context. **9.1** • Determine whether the results of a study are statistically significant and make an appropriate conclusion using a significant level. **9.1** • Interpret a Type 1 and Type 2 error in context and give a consequence of each. **9.1** • State and check the conditions for performing a significance test about a population parameter. **9.2** • Perform a significance test about a population proportion. **9.2** • Interpret the power of a test and describe what factors affect the power of a test. **9.2** • Describe the relationship among the probability of a Type 1 error, the probability of a Type 2 error, and the power of a test. **9.2** • State and check the conditions for performing a significant test about a population mean. **9.3** • Perform a significance test about a population mean. **9.3** • Use a confidence interval to draw a conclusion for a two-sided test about a population parameter. **9.3** • Perform a significance test about a mean difference using paired data. **9.3** • Determine the shape, center, and spread of the sampling distribution of the difference of sample proportions. **10.1**

• The formula for finding the sampling distribution of a slope. • The formula for calculating the t interval for the slope. • The t test for the slope.	• Determine whether the conditions are met for doing inference about sample proportions. **10.1** • Construct and interpret a confidence interval to compare two proportions. **10.1** • Perform a significance test to compare two proportions. **10.1** • Describe the shape, center, and spread of the sampling distribution of the difference of sample means. **10.2** • Determine whether the conditions are met for doing inference about the difference of sample means. **10.2** • Construct and interpret a confidence interval to compare two means. **10.2** • Perform a significance test to compare two means. **10.2** • Determine when it is appropriate to use two-sample t procedures versus paired t procedures. **10.2** • State appropriate hypotheses and compute expected counts for a chi-square test for goodness of fit. **11.1** • Calculate the chi-square statistic, degrees of freedom, and P-value for a chi-square test for goodness of fit. **11.1** • Perform a chi-square test for goodness of fit. **11.1** • Conduct a follow-up analysis when the results of a chi-square test are statistically significant. **11.1** • Check the conditions for performing inference about the slope of the population regression line. **12.1** • Interpret the values of a, b, s, SEb, and r square in context. **12.1** • Determine the values of a, b, s, SEb, and r square from a computer output. **12.1** • Construct and interpret a confidence interval for the slope of the population regression line. **12.1** • Perform a significance test about the slope of the population regression line. **12.1**

Phase III Textbook/Materials
Phase IV Summative Assessment Evidence

Common Summative Unit Assessments: (*identifies Performance Task)	**Agreed Upon Interim Summative Assessments:** (*identifies Performance Task)

Phase V Learning Plan