PISCATAWAY TOWNSHIP SCHOOLS

Dr. Frank Ranelli
Superintendent of Schools

Dr. William Baskerville
Assistant Superintendent

AP Statistics

Content Area:	**Mathematics**
Grade Span:	**9–12**
Revised by:	**Jackie Puleio**
Title	**Supervisor of Mathematics, 7–12**
Approval Date:	

Piscataway Township Schools
1515 Stelton Road
Piscataway, NJ 08854–1332
732 572–2289, ext. 2561
Fax 732 572–1540
www.piscatawayschools.org

COURSE OVERVIEW

Description

Advanced Placement Statistics introduces students to the major concepts and tools for collecting, analyzing, and drawing conclusions from data. There are four themes in the AP Statistics course: exploring data, sampling and experimentation, anticipating patterns, and statistical inference. Students use technology, investigations, problem solving, and writing as they build conceptual understanding, reinforce procedural knowledge, apply problem solving strategies, reason and justify their conclusions, and develop a positive disposition towards learning and doing mathematics.

The purpose of AP Statistics is to demonstrate the relevance of the course content to a range of relevant everyday issues that can be addressed with the help of statistical analysis techniques. This course is meant to expand the understanding of data collection and the role of statistics in making inferences from data. Applications from realistic contexts such as business and economics, the social and physical sciences, healthcare, education, engineering and leisure activities will be examined throughout the text. Students will use realistic data that they have either collected or found using simulations or laboratory investigations and then analyze it for class assignments/projects. Technology, via graphing calculator and spreadsheet software, is used daily and will enhance the depth of particular topics and lessen the drudgery of hand calculations. Major topics include descriptive statistics, probability and probability distributions, hypothesis testing, correlation and regression analyses, and contingency tables as outlined by the College Board of Advanced Placement.

Goals

In addition to the content standards, skills, and concepts set forth by Collegeboard, this course also seeks to meet the Standards for Mathematical Practice set forth in the New Jersey Student Learning Standards. These practices include generally applied best practices for learning mathematics, such as understanding the nature of proof and having a productive disposition towards the subject, and are not tied to a particular set of content.

The eight Standards for Mathematical Practice are outlined below:

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precision.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

Scope and Sequence

Unit	Topic	Length
Unit 1	Displaying and Describing Data	20 days
Unit 2	Correlation and Regression	15 days
Unit 3	Gathering Data	10 days
Unit 4	Randomness and Probability	15 days
Unit 5	Inferences About Proportions	20 days

Unit 6	Inferences About Means	10 days
Unit 7	Chi–Square and Inferences About Regression	5 days
Unit 8	Constructing, Implementing, and Evaluating the Success of a Risk–Based Business Model	8 days

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Each unit will include lab activities and a culminating project that will require students to gather and analyze collected data and organize a final product that represents their findings. These items include formal analyses that outline their statistical procedures, data from their findings, and a presentation of their analyses. This projects can be represented formally in a lab report or more creatively in a poster, movie, or presentation.
**see each unit for specific "lab" activities*

After the Advanced Placement Statistics Exam, students will work in groups on the Final Exam Project: The AP Statistics Carnival. This project requires students to reference knowledge and skills from each unit covered in the curriculum to create and successfully execute a "profitable" carnival game. Students then have the opportunity to construct,advertise, and run the carnival game in front of a live audience. While the carnival is being run, students collect real–time data to be analyzed to determine the success of their game. These analyses include surveys and several forms of hypothesis tests. All of the information gathered during the carnival is then organized into a final presentation to evaluate the success of their enterprise.

UNIT 1: Displaying and Describing Data

Summary and Rationale

This unit focuses on developing the foundational knowledge and vocabulary necessary for the remainder of the course. Specific focuses will include analyzing both categorical (one– and two– variable) and quantitative (one–variable) data, creating data displays, and an introduction to the Normal Model. This unit lays the groundwork for many of the skills required for more in–depth analysis in the statistical concepts that will be explored in this curriculum. There will be a strong emphasis on articulating a clear descriptive analysis of data sets as writing skills and communication are imperative skills required to be successful on the AP exam.

Recommended Pacing

20 days

Instructional Focus

Unit Enduring Understandings

- Information can be interpreted in multiple ways; supporting conclusions with evidence is essential.
- Data is meaningless without context.
- Graphical representations are a tool that can be used to represent key components of a data set.
- Many real–world scenarios can be modeled using the Normal Distribution.
- Technology can be used to assist in the interpretation/analysis of data.
- Utilizing a data display is essential in an efficient and effective analysis.
- Utilizing an appropriate numerical summary can persuade public perception of an analysis.
- Data displays and numerical summaries are interrelated and both are essential to a statistical analysis.
- A superior analysis of data sets is composed of numerical associations and comparative display(s).
- Methods of grouping data can affect the patterns and relationships that become apparent.
- Outliers may emerge depending on the method of grouping data and choice of displayed distributions.
- Shifting and rescaling affect measures of center, while measures of spread are only affected by rescaling data.

Unit Essential Questions

- What information can be used to describe a univariate quantitative data set?
- How can association between two categorical variables be identified?
- How can the Normal Model be used to help us describe a data set?
- How can data be presented in a way that illustrates its story?
- How can appropriate display(s) and numerical summaries be chosen to highlight comparisons of multiple data sets?
- How do measures of center and spread vary when working with data?

Objectives

Students will know:

- The difference between categorical and quantitative data
- The various measures of center (mean, median, etc.) and spread (standard deviation, variance, IQR, etc.)
- The difference between continuous and discrete data
- Various shapes of quantitative data (symmetric vs. skewed, uniform, unimodal vs. bimodal vs. multimodal)
- How data sets and measures of center & spread are affected by transformations of the data set
- The key features of the Normal Model

Students will be able to:

- Create and interpret appropriate data displays for a given set of data (for one– and two– variable categorical and one–variable quantitative data)
    - This should be done with and without the assistance of technology
- Describe a set of categorical/quantitative data set in context
- Use a two–way table/contingency table and segmented bar chart to determine if an association exists between two categorical variables
- Calculate probabilities from a two–way/contingency table
- Analyze a quantitative data set using appropriate descriptors of shape, center, and spread
- Use technology to identify the mean, standard deviation, variance, 5 number summary, etc. of a data set
- Determine if a data point is an outlier/unusual
- Create written comparisons of quantitative data sets (including comparisons of specific positions in the data, ie. percentiles)
- Find the z–score of a given data point and use the z–score to find the actual data point
- Standardize data that follows a Normal Model in order to be able to calculate probabilities (using normalcdf on the TI–84+) and using probabilities/percentiles to calculate the data point (invNorm)

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:

Lab: Gathering, Displaying & Analyzing Univariate Data

Students obtain their own data, create appropriate data displays (both with and without the assistance of technology, and interpret the information in context.

Lab: Contingency Tables (Probabilities & Association)

Students gather information from a sample of bivariate categorical data (e.g. Recording the frequency of different colored M&Ms in bags of various types of M&Ms) to investigate the existence of an association.

Lab: Effects of Transforming Univariate Data

Starting with one set of quantitative data, students explore the effects of adding/subtracting and multiplying/dividing each number in the data set by a constant.

Lab: Exploring the Normal Model

Students create problems that can be addressed using the Normal Model.

UNIT 2: Correlation and Regression

Summary and Rationale

This unit introduces and investigates bivariate quantitative data. Scatter plots, regression analysis, and residuals will be the central focus. Students will continue to develop writing skills and put mathematical relationships within the regression in context. They will also be utilizing several formulas related to the calculation and interpretation of a least square regression line to help formulate reliable predictions based on explanatory variables. Students will continue to manipulate data and explore the effects of transformations on variables to meet the conditions required to implement linear regression analysis.

Recommended Pacing

15 days

Instructional Focus

Unit Enduring Understandings

- Information can be interpreted in multiple ways; supporting conclusions with evidence is essential.
- Data is meaningless without context.
- Graphical representations are a tool that can be used to represent key components of a data set.
- Technology can be used to assist in the interpretation/analysis of data.
- Data can be used to make predictions so long as the conditions for the analysis are met.
- A strong correlation between two variables does not indicate causation.
- The relationship between two variables, in which the association is linear, is summarized using a linear equation.
- The slope of the regression line is greatly impacted by outliers.
- The magnitude of correlation is invalid for a nonlinear relationship.
- The correlation coefficient and the regression slope are related.
- The significance of a point can be determined influential by understanding its leverage and residual effect in comparison to the data set.

Unit Essential Questions

- When is it reliable to make predictions based on an explanatory variable?
- How can we determine the strength of an association between two quantitative variables?
- How can we determine if a least square regression line (LSRL) adequately “fits” the data set?
- How can the relationship between two variables be identified, quantified, and described?
- How can outliers and influential points be identified to yield a more accurate regression model?
- What is the relationship between residuals and the coefficient of determination (R^2)?
- How can residuals be used to analyze the appropriateness of a linear model?
- How can two variables with a nonlinear association be analyzed and relationships quantified using a linear regression model?

Objectives

Students will know:

- The definitions of terms associated with regression analysis (correlation, coefficient of determination, etc.)
- The formulas associated with regression analysis (slope, y–intercept, etc.)
- The definition of a residual
- Data can be transformed to become more linear
- The difference between an outlier and an influential point
- Extrapolation far from the mean of the x variable can lead to unreliable predictions
- An R^2 (coefficient of variation) of 100% does not indicate causation between the two variables

- If omitting a point from the data can result in a very different slope or overall regression model, then the point is considered influential
- The purpose of data re–expression allows the researcher to utilize a regression line for nonlinear data

Students will be able to:

- Identify the roles of the variables as explanatory and response
- Construct a scatter plot by hand and through the use of a scatterplot
- Describe the strength, form, and direction of an association between two quantitative variables in context
- Calculate the correlation coefficient (r) and the coefficient of determination (R^2) of a bivariate data set
- Write the equation of a least squares regression line by calculating the slope and y–intercept
- Interpret the correlation coefficient of determination, slope and y–intercept in context for linear and nonlinear regression transformations
- Calculate the residual of a given data point and interpret its meaning in context
- Construct/interpret a residual plot
- Pair a residual plot with its scatter plot/LSRL
- Recognize and interpret data that has been transformed (using logs, exponents, square roots, etc.)
- Transform data with the calculator or spreadsheet (if time allows)
- Investigate, display, re–express, and analyze residuals in quantitative bivariate data to build linear models to formulate accurate/reliable predictions

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Lab: Gathering Data to Create and Analyze a Least Squares Regression Line
Students participate in an activity that results in a bivariate quantitative data set in order to create a scatter plot, analyze the correlation, and develop regression model.
Lab: What Does this Residual Mean?
Students will develop residual plots from scatter plots and vice versa.
Lab: Dealing with Non–Linear Data
Class will run a simulation that results in nonlinear data. Students test which transformation of the variable (one or both) results in the most linear data.

UNIT 3: Gathering Data

Summary and Rationale

This unit focuses on the importance of the methods used to collect data as well as the important differences between a study and an experiment. Students will also explore the different types of bias that can occur when collecting data. The main goal of this unit is to have students realize that the sampling techniques/design of the experiment used to collect the data is just as important as the results of the data collection as this will determine whether results can be generalized or if causation can be discussed.

Recommended Pacing

10 days

Instructional Focus

Unit Enduring Understandings

- The manner in which data is gathered affects what can be concluded about a population.
- Evidence of causation can be established by the results of a repeated experiment (when properly designed); replication validates the results of a study..
- Sample statistics are used to estimate population parameters if and only if a representative sample is drawn.
- An element of randomization is essential in collecting data in both experiments and observational studies.

Unit Essential Questions

- What are the factors of a well–designed experiment?
- Why is it important to use random selection when gathering data?
- When can we generalize the results of a study/experiment to the population as a whole?
- Why is a control an essential component of experimental design?
- How can experimental design be used to reduce confounding variables?
- How can you determine the difference between an observational study and an experiment?

Objectives

Students will know:

- The various methods of sampling (simple random, convenience, stratified, cluster, systematic, multi–stage)
- The various types of bias as well as other factors that can affect the outcome of a study/experiment (lurking/confounding variables)
- The characteristics of a well–designed experiment
- The difference between a census and a sample survey, between an observational study and an experiment
- The basic principles of experimental design (control, randomize, replicate, block)
- The components of an experimental design are factors, treatments, and response variable
- A control of experimental design can be implemented using a placebo or multiple treatment groups as a means of comparison
- The Placebo Effect can only occur in treatments of humans, hence a blinding through the use of a placebo is not always necessary in experimental design
- In experiments, the element of randomized treatments is essential in well-designed study

Students will be able to:

- Collect data using the various sampling methods
- Design simulations by defining the component, assignment, trial, and random variable and run using random numbers or another method of randomization

- Discuss the results of a simulation and how it extends to the real world
- Recognize when a simulation is more useful in modeling random behavior than the act of collecting data
- Identify types of bias/other issues that could affect the results of a particular study/experiment
- Design a completely randomized experiment (including a control group, blocking, random assignment, replication, etc.)
- Report the results of an experimental using component, random assignment, defining a trial, and identifying the response variable
- Discuss the benefits/downfalls of each method of sampling
- Analyze the results of a study or experiment in context
- Recognize random outcomes in the real world
- Recognize when to use an observational study/experiment
- Identify observational studies and classify them as retrospective or prospective
- Use graphical displays to compare responses for different treatment groups

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Lab: Bias & Error in Sample Surveys
Students create and distribute their own surveys and discuss the bias and error that possibly occurred.
Lab: Designing Simulations
Groups are presented with varying scenarios for which they are required to design and run simulations.
Lab: Conducting the Same Study with Different Sampling Methods
Students are presented with data from 100 plots of land and use various sampling methods to estimate the true yield of crops from the field. They will investigate which method(s) provided the best estimate.

UNIT 4: Randomness and Probability

Summary and Rationale

This unit explores the concepts of randomness, probability, and binomial modeling. Students will be creating various representations to facilitate their comprehension and computations, such as tree diagrams, Venn diagrams, and frequency tables. Students will tinker with real–life scenarios to help them understand the nuances between theoretical and empirical probabilities, disjoint and independent events, and marginal, joint, and conditional frequencies; they will use this knowledge to construct a schema for random variables and the binomial and normal models that emerge from it.

Recommended Pacing

15 days

Instructional Focus

Unit Enduring Understandings

- The probability of any singular or series of events must be between 0 (impossible) and 1 (certain)
- In a conditional probability scenario, new given information has the potential to adjust the values of a numerator or denominator (or both) compared to an original non–conditional scenario.
- The numeric value of a random variable is based on the outcome of a random event.
- The uses of different probability models may vary in order to accurately evaluate risk for a given scenario.
- Probability is a way of predicting outcomes, but does not assure outcomes.
- Random phenomena are unpredictable in the short term but show patterns in the long run that enable predictions to be made.
- Risk and fairness are interrelated.
- Binomial probability models and approximating the probability using the Normal Model sets the foundation for many tools within statistical inference.

Unit Essential Questions

- What do the numerator and denominator of a probability's fraction each represent?
- How can a Venn diagram be used to explore the concepts of union, intersection, and disjoint?
- How can tree diagrams and frequency tables be used to compute conditional probability?
- What are the similarities and differences between a discrete and continuous random variable?
- What conditions must be met to establish that a model is binomial, and then what further must be shown to conclude that the binomial model approximates the Normal Distribution?
- How can the conditions of a situation impact the calculation of its probability?
- How can we make predictions in the efficiency and success of a game using probability?

Objectives

Students will know:

- The definition of probability and the domain of possible values
- The characteristics of and relationships between multiple probability events (fundamental counting principle, independence, union, intersection, disjoint, complement, conditional)
- The concept of a random variable, along with its two subgroups: discrete and continuous
- There are two conditions to classify a Bernoulli Trial: (1) there are/can be reduced to two possible outcomes two outcomes, (2) the probability of success is unchanging/trials are independent
- The relevant assumptions and conditions needed to establish a binomial or geometric probability, and to conclude normality

- A geometric probability problem is a special case of the binomial, in that the number of successes is exactly 1 and occurs only after a set number of failures
- Operations between independent random variables, *X* and *Y*, affect the mean and standard deviation in the following manner:
    - $E(X \pm Y) = E(X) \pm E(Y)$
    - $Var(X \pm Y) = Var(X) + Var(Y)$

Students will be able to:

- Compare empirical probabilities and theoretical probabilities through the lens of The Law of Large Numbers
- Assess whether or not events are independent, disjoint, both, or neither
- Draw diagrams and analyze tables that represent frequencies, randomness, and probabilities
- Identify random variables in applications and construct the relevant probability model(s) for that random variable
- Compute the expected value and standard deviation of a random variable, as well as how to combine independent random variables' expected values or standard deviations
- Calculate probabilities using the Geometric, Binomial, and Normal Models using technology
- Understand when binomial probability can be approximated using the Normal Model
- Interpret the mean and standard deviation of the Binomial, Geometric, and Normal Models

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Lab: Exploring the Law of Large Numbers

Students will use two dice to gather empirical data regarding the probabilities of rolling each possible sum. They will compare their individual results to results when combining their results with a partner and to the whole class', etc. to demonstrate how the empirical probability better approximates the theoretical probability as the number of trials increases.

Lab: Calculating & Interpreting Expected Value

Using a spinner with specific "prizes" associated with different sections of the spinner, students calculate expected winnings/loss and the effects of adding variables/multiplying by a constant.

Lab: Binomial vs. Geometric

Students compare, calculate, and examine the properties of binomial and geometric probability distributions.

Unit 5: Inferences About Proportions

Summary and Rationale

This unit extensively explores the normal model by introducing sampling distributions, z–tests, confidence intervals, and p–values. Students will learn about the strategy involved in estimating statistical constants and declaring significant results, as they design their own hypothesis tests and come to understand the consequences that researchers face if they make various errors in judgment. This unit's focus is primarily on sample proportions from data, which sets the stage for the next unit to mirror these concepts by concentrating on sample means from data.

Recommended Pacing

20 days

Instructional Focus

Unit Enduring Understandings

- Regardless of the shape of a population distribution, the sampling distribution for any data set approaches a Normal Distribution as the sample size increases; this major tenet of statistics is called the Central Limit Theorem.
- If one can conclude that the counts from different groups are from the same population (i.e. contain the same parameters), one can "pool the data" together to compute and analyze more precisely.
- Significance tests can determine the relationship between different sample statistics to make inferences about the population(s).
- Confidence intervals of sample statistics measure the magnitude of the relationships between populations.
- Tests of significance help to draw conclusions about a population in an efficient and cost–effective manner.
- Significance level and confidence level are interrelated.
- Confidence Intervals are a balance between precision and the certainty of a statement about a population's parameter.
- The power of a hypothesis test is dependent on sample size, effect size, and the probability of Type I and Type II errors.
- The possibility of Type I and Type II Errors exist within every hypothesis test; it is the responsibility of the statistician to both minimize the impacts through the elements of a hypothesis test (significance level, sample size, etc) and consciously evaluate the effects of each error in their findings.

Unit Essential Questions

- How does a sampling distribution differ from a sample distribution?
- What does it mean to have a 95% level of confidence?
- When is it appropriate to reject the null hypothesis, and why is "failing to reject" the null hypothesis different than "accepting" the null hypothesis?
- How does the decision to construct a one–tailed or two–tailed hypothesis test affect the ability to reject the null hypothesis?
- What is the purpose of statistical inference?
- How can tests of significance and confidence intervals be used to make decisions?
- How are power, confidence, level of significance, and error interrelated?

Objectives

Students will know:

- The relevant assumptions and conditions needed to model a sampling distribution for sample proportions, and to conclude normality

- The notation and relationship between null and alternative hypotheses
- How to accurately describe the null and alternative hypothesis and the importance of using the correct vocabulary
- Type I error is rejecting the null hypothesis when it is true; Type II error is failing to reject the null hypothesis when it is false
- A p–value is the probability of encountering the collected data given that the null hypothesis is true
- When it is appropriate to construct a one–tailed or two–tailed test
- The connections between test error, alpha (significance), beta, power, effect size, and sample size

Students will be able to:

- Apply their previous knowledge of the approximation of probabilities in the Normal Model to sampling distributions
- Identify, explain, and compute standard errors, margins of error, critical values (z*), and confidence intervals
- Create and run an appropriate one–proportion z–test for a specific scenario
- Discern whether a one–proportion or two–proportion z–interval (or z–test) is an appropriate methodology

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Lab: Central Limit Theorem
Students will calculate the means from sample data for increasing sample sizes and examine the changes that occur to the sampling distributions of the class' data to demonstrate the Central Limit Theorem.
Lab: Constructing a Confidence Interval Based on Sample Data
Students will flip a coin multiple times, record the proportion of heads, and use this sample proportion to construct a confidence interval to estimate the true proportion of heads. Students will be able to see a visual representation of each individual's confidence interval to see the percentage of confidence intervals that contain the true proportion.
Lab: What is a p–value?
Students will explore the concept of a p–value (probability of obtaining specified sample statistic if the null hypothesis is true).
Lab: Collecting & Using Sample Data to Run a Hypothesis Test
Students will collect a simple random sample from a set of data and use it to conduct a one–proportion z–test.

Unit 6: Inferences About Means

Summary and Rationale

This unit takes the previous unit's underlying principles of hypothesis testing, statistical significance, and modeling data and explores them through the lens of sample means (rather than sample proportions), and with the additional caveat of potentially having incomplete information. Students will explore the reality of needing to use certain assumptions and conditions to justify a more conservative and safer approach to statistical analysis using concepts from previous units . The t–distribution will be extensively covered, and students will recognize the parallels between last unit's z–tests with this unit's t–tests.

Recommended Pacing

10 days

Instructional Focus

Unit Enduring Understandings

- According to the Central Limit Theorem, the mean of a sampling distribution of means is an unbiased estimator of the population mean, and the standard deviation (standard error) of a sampling distribution of means is the population standard deviation divided by the square root of the sample size.
- When analyzing means (as opposed to proportions), the population standard deviation is often unknown, and so the sample standard deviation is computed as a necessary substitute to estimate the parameter.
- As the sample size of a t–distribution increases, the shape of the t–distribution approaches that of the Normal Distribution.

Unit Essential Questions

- What are the similarities and differences between a t–test and z–test?
- How can one verify whether or not the "nearly normal" condition is met, and why is this an important characteristic?
- How does the margin of error, confidence interval, and overall precision of a t–interval compare to that of a z–interval, and why?
- What is the difference between situations that require the use of a t–distribution and the z–distribution?
- What is the similarities between the assumptions for tests that are modeled via a t–distribution and z–distribution?
- How are the t–distribution and the z–distribution related?
- Why are confidence intervals an essential component when making decisions based on hypothesis testing?
- Why is pooling mandatory for a two–proportion test but optional (and less likely) for a two–sample t–test for means?

Objectives

Students will know:

- When to use a t–test (or t–interval) instead of a z–test (or z–interval)
- The relevant assumptions and conditions needed to use a t–distribution model
- The distinction between a two–sample test (or interval) and a paired–differences test (or interval)

Students will be able to:

- Compute the degrees of freedom for a t–test or t–interval
- Run a variety of statistical tests for (or between) means, and analyze the results
- Apply their knowledge of testing hypotheses, pooling data, and recognizing one–sample vs. two–sample scenarios from the previous unit to this unit

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Lab: One– and Two Sample t–tests

Students use means and data from two samples to conduct both one and two sample t–tests. (e.g. Students design and construct two paper airplanes, run trials to determine the mean hang time of each plane, run one–Sample t–tests comparing each plane separately to a known population parameter and a two–sample t–test comparing the average hang times to each other.

Lab: When can we Pair Data?

Using data from the class, students will conduct a paired t–test discussing when two samples result in data that can be paired.

Unit 7: Chi–Square and Inferences About Regression

Summary and Rationale

This unit combines a categorical analysis of chi–squared (χ^2) modeling with a quantitative extension of further regression inference. Students will take their knowledge of hypothesis testing, distributions, correlation, and other statistical topics practiced earlier in the year and use them for novel, complex computations and evaluations. This unit also utilizes supplementary software output to provide students with other representations of data that they must learn to interpret.

Recommended Pacing

5 days

Instructional Focus

Unit Enduring Understandings

- Chi–squared tests and distributions are modeled based on the counts from categorical data and degrees of freedom.
- Inferences for regression include a hypothesis test that there exists an association between variables and a confidence interval for the slope of a regression line.

Unit Essential Questions

- How are the three primary chi–squared tests in this unit similar and different?
- How do observed data and expected data play a role in computing and assessing a chi–squared test statistic?
- What are the connections between the regression analysis performed in Unit 2: Linear Regression and the additional regression inferences explored in this unit?

Objectives

Students will know:

- The notation, degrees of freedom, hypothesis tests, and characteristics of each chi–squared test
- The relevant assumptions and conditions needed to use a chi–squared model or to make inferences about the coefficients of a regression equation
- A value in a data set that has the greatest influence on the chi-square value (χ^2) is essential to examine and reference in an analysis
- Inferences for regression extend from and combine earlier topics, such as correlation, regression, sampling distributions, confidence intervals, p–values, and t–distributions

Students will be able to:

- Describe how sample size affects the shape of chi–squared distributions
- Evaluate whether or not a collection of data fits a given trend, is proportionally consistent when measuring across a single variable, or reveals a significant association between two variables
- Create tables to organize chi–squared models
- Use software outputs to model and interpret inferences for regression

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)

Online supplemental resources (e.g., Albert.io)

Exploratory Activities:
Lab: Chi–Square – The Three Tests

Students play Rock, Paper, Scissors against a computer and gather data as they play to answer questions such as:

- Is there a difference in skill level when playing against the computer in “novice” or the “expert” mode?
- Did the computer play rock, paper, or scissors the same in both novice and expert mode?
- Did human influence prove to be more/less successful over randomization when playing against the computer?

They can then use this data to run one of each of the three types of chi–square tests.

Unit 8: Constructing, Implementing, and Evaluating the Success of a Risk–Based Business Model

Summary and Rationale

AP Statistics students are tasked with designing, constructing, and implementing a unique carnival game enterprise to explore the balance of success and entrepreneurial risks for their new "business". Students explore ideas of marketing, data, and cost/profit analysis to ultimately determine the success of their design. Students are inspired by their instructors to develop this game in the post–AP Exam timeframe and will use the data collected from the carnival, typically run in a middle school, to report on the success of their game for their final exam presentation. The themes and objectives of the carnival games are only limited by the students' imaginations and the elimination of players' skills from their design. Students draw on inspiration from a variety of disciplines and interests, ranging from sports to animals to the arts and sciences. This allows for AP Statistics students to benefit from the authenticity of this course curriculum, while inspiring the youth of the district to pursue AP Statistics in their future mathematical careers.

Recommended Pacing

8 days

Instructional Focus

Unit Enduring Understandings

- In order to design a successful enterprise, all variables must be taken into account and analyzed.

Unit Essential Questions

- How can the success of a risk–based business model be predicted?
- How can the success of a risk–based business model be measured?
- How can surveying and experimentation help to make changes to a business model to increase success?

Objectives

Students will know:

- Expected value is essential in estimating profit and success of a business model
- Surveys and experiments are useful in predicting human behavior
- Hypothesis testing is a method of measuring the success of a design and a business model in the long run

Students will be able to:

- Design, construct, implement, and evaluate the success of a risk–based business model
- Apply concepts learned through the entire curriculum in a real life context
- Design a completely risk based business model
- Build their carnival game and have the student body play games to gather raw data
- Collect data on the players of the game and analyze their profit margins. Success is analyzed in comparison to other games at the event
- Analyze a budget based off of risk and profit
- Analyze whether designed predictions for the risk based model were accurate
- Examine whether the variation from a theoretical business model are due to the faults of risk–based design or due to sample variation

Resources

Core Text: Bock, Velleman, and De Veux's "Stats: Modeling the World (3rd edition)".

Suggested Resources:
Graphing Calculator (TI–83, TI–84)
Online supplemental resources (e.g., Albert.io)

Exploratory Activities:

- Game Blueprint
- Risk–Based Business Model Design
- Experimental Design and Analysis
- Advertisement for Carnival Day
- Identifying Claims to Measure the Success of the Business