TOWNSHIP OF UNION PUBLIC SCHOOLS

Statistics
Adopted June 17, 2015
Updated December 18, 2018

Mission Statement

The mission of the Township of Union Public Schools is to build on the foundations of honesty, excellence, integrity, strong family, and community partnerships. We promote a supportive learning environment where every student is challenged, inspired, empowered, and respected as diverse learners. Through cultivation of students' intellectual curiosity, skills and knowledge, our students can achieve academically and socially, and contribute as responsible and productive citizens of our global community.

Philosophy Statement

The Township of Union Public School District, as a societal agency, reflects democratic ideals and concepts through its educational practices. It is the belief of the Board of Education that a primary function of the Township of Union Public School System is to formulate a learning climate conducive to the needs of all students in general, providing therein for individual differences. The school operates as a partner with the home and community.

Course Description

This mathematics course will prepare students for college level probability and statistics courses. These college courses are often requirements for many college majors. The TI-84+ graphing calculator is used extensively throughout the course. Upon completion of this course, students will be able to describe events using statistics; organize and summarize data; determine probability of compounded events; analyze and draw appropriate inferences from data; and use listing, counting, and algorithmic methods to solve real world problems.

Units

Unit 1: **Exploring and Understanding Data**

- Exploratory analysis of data makes use of graphical and numerical techniques to study patterns and departures from patterns.

Unit 2: **Exploring Relationships between Variables**

- In examining distributions of one and two categorical and quantitative data, students should be able to detect important characteristics, such as shape, location, variability, and unusual values. From careful observations of patterns in data, students can generate conjectures about relationships among variables.

Unit 3: **Gathering Data**

- Both the type of analysis that is appropriate and the nature of conclusions that can be drawn from that analysis depend in a critical way on how the data was collected. Emphasis will be placed on sampling methods to include: simple random, stratified, systematic, and cluster sampling. Data must be collected according to a well-developed plan if valid information is to be obtained. Experiments must be designed with respect to blocking, blinding, the placebo effect, and methods of randomization

Unit 4: **Randomness and Probability**

- Recognize independence, the role it plays in calculating conditional probabilities and use them to interpret data. The rules of probability allow for the calculation of the probability of compound events in a uniform probability model. Calculate expected value and use it to solve problems and evaluate the outcomes of decisions.

Unit 5: **From the data at hand to the world at large**

- Making inferences from data can be thought of as the process of selecting a reasonable model and generalizing sample data to draw justifiable conclusions about the entire population.

Unit 6: **Accessing associations between variables**

- Determine if an association exists between two categorical variables by conducting the appropriate chi-square significance test and determine if an association exists between two quantitative variables by making inferences about the slope linear regression equation.

Recommended Textbooks

Student Text:

Statistics: Intro Stats 4th Edition (DeVeaux Velleman Bock) - Pearson

Recommended Student Calculator:
TI 84 Plus Silver Edition Color

Course Proficiencies

EACH STUDENT WILL BE ABLE TO:

- Read and understand terminology in published statistical reports and journals
- Create and interpret graphs from data and calculate descriptive statistics
- Calculate and interpret correlation between two variables
- Determine the influence of one variable on another.
- Know the characteristics of and the difference between observational studies and experiments
- Know the different ways of selecting samples from a population
- Know how to design a study and be aware of the biases that may exist
- Calculate the probabilities of real world phenomena
- Students will understand and use sampling distributions
- Calculate and interpret confidence intervals in real world situations
- Perform significance tests and draw valid conclusions from their findings
- Perform chi-squared test on categorical data

Pacing Guide Statistics
Semester I

Unit 1

Exploring and Understanding Data	**Sections**	**# of days**
Stats Start Here	1-3	5
Displaying and Describing Categorical Data	1-2	5
Displaying and Summarizing Quantitative Data	1-7	10
Understanding and Comparing Distributions	1-3	8
Standard Deviation and the Normal Model	1-5	10

Unit 2

Exploring Relationships between Variables		
Scatterplots, Association, and Correlation	1-3	5
Linear Regression	1-7	12
Regression Wisdom	1-5	8

Unit 3

Gathering Data		
Understanding Randomness	1-2	5
Sample Surveys	1-7	11
Experiments and Observational Studies	1-6	11
	Total	90

Pacing Guide – Statistics
Semester II

Unit 4

Randomness and Probability	**Sections**	**# of days**
From Randomness to Probability	1-3	5
Probability Rules	1-5	7
Random Variables and Probability Models	1-5,7	8

Unit 5

From the data at hand to the world at large

Sampling distribution models	1-4	6
Confidence intervals for proportions	1-4	8
Testing hypotheses about proportions	1-5	10
Inferences about means	1-5	10
More about Tests and Intervals	1-5	9

Unit 6

Accessing associations between variables

Comparing Groups	1-7	9
Paired samples and Blocks	1-4	6
Comparing counts	1-4	6
Inferences for Regression	1-4	6
	Total	90

Unit 1: Exploring and Understanding Data

Essential Questions	NJSLS/Instructional Outcomes	Activities	Assessments
1. How can we construct and interpret graphical displays of distributions of univariate data? (stemplot, histogram) 2. How can we summarizing distributions of univariate data? 3. How can we compare distributions of univariate data (stem plots and parallel boxplots)? 4. How can we explore bivariate data? 5. How can we explore categorical data? 6. How can we use standard deviation to measure differences from an expected value?	1. Identify center, spread, clusters, outliers and other unusual features of univariate data by reading graphical representations. **S-ID2** 2. Calculate center, spread and position of univariate data **S-ID3** 3. Compare and contrast features of different univariate distributions.**S-ID3**	1. Instructor and student use of interactive computer software applet to demonstrate resistance of mean and median by outliers in a data set. 2. Use of correlation by eye computer applet designed to aid students in recognizing strength of correlation. 3. Use Microsoft Excel computer program to complete least squares regression project **(See Appendix A)** 4. Organize univariate data into logical graphical representations using the graphing calculator that can be used to make conclusions about univariate distributions. 5. Enter data into graphing calculator and run one variable statistics to find the mean, median, mode, range, standard deviation, percentiles, and z-scores for univariate data. 6. Estimate population percentages using normal distribution. Use calculators, spreadsheets, and tables to estimate areas under the normal curve **(Interdisciplinary Connection)** 7. Data comparing populations across states and nations will be used to motivate lessons and generate discussions on students' global perspectives. **(Global Perspectives)**	Use given data of test grades from the previous statistics test to answer questions 1-3 58, 89, 67, 99, 74, 91, 84, 86, 70, 73, 97, 61, 52, 88, 55, 12, 78, 60 1. Calculate mean, median, mode, range, and standard deviation of data. 2. Construct a stem leaf plot of the data. 3. Describe the distribution of the data. Calculate boundaries for and make mention of any outliers.

Suggested Differentiation for Unit 1

- **Tier 1 Learners:**
  - Have guided notes filled out at different levels according to ability.
  - Give assignments that contain tasks of varying difficulty. Each task should focus on essential learning that all students should master, but the tasks will vary in difficulty.
  - Group students by similar interest when working on application problems.
  - Use mini lessons to reteach to those having difficulty.
  - Group students so that each group contains all level learners. The tier 3 learners can serve as peer helpers.
  - Assign a basic homework assignment. Require students to spend a set amount of time to work (showing effort) on the assignment rather than completing the entire assignment.
  - Allow students to choose a method for completing a project: video, PowerPoint, paper, or presentation.
- **Tier 2 Learners:**
  - Utilize foldables creating tangible products to help students digest information while incorporating several of the multiple intelligences.
- **Tier 3 Learners:**
  - Have problems posted around the room. Have students loop to specific questions based on difficulty.

Curriculum Development Resources

- Course textbook
- Graphing calculator
- Kuta software
- Quia software

Unit 2: Exploring Relationships between Variables

Essential Questions	NJSLS/Instructional Outcomes	Activities	Assessments
1. How can we use scatterplots to determine association between quantitative variables? 2. How can we create a linear prediction model? 3. How good of a predictor is our linear regression model? 4. What are the reasons that correlation does not imply causation? 5. In what ways do we need to exercise caution when making predictions?	1. Analyze patterns in scatterplots, recognize correlation and linearity. 2. Find least-squares regression line and verify its validity by checking residual plots, outliers, and influential points**. S-ID6** 3. Construct and interpret representations of categorical data. **S-ID5**	1. Check the validity of least-squares regression by analyzing patterns in residual plots 2. Fit a function to the data, use linear, quadratic and exponential functions fitted to data to solve problems in the context of the data. 3. Clarify that r-squared is a measure of how much of the variability in the response variable can be explained by variability in the explanatory variable. 4. Explore examples that clarify why correlation does not infer causation. **Interdisciplinary Connections:** Most of the data given to students in this unit will be from Real-world sources relating to sociology, science, or economics.	1. Given a data list of SAT scores and hours of preparation, create a linear regression model and make predictions for students that: don't study, study for 20 hours, and study for 100 hours. 2. A regression equation that uses height to predict weight is y=110+0.5x. How much would you expect a person that is 70 inches tall to weigh?

Suggested Differentiation for Unit 2

- **Tier 1 Learners:**
  - Have guided notes filled out at different levels according to ability.
  - Give assignments that contain tasks of varying difficulty. Each task should focus on essential learning that all students should master, but the tasks will vary in difficulty.
  - Group students by similar interest when working on application problems.
  - Use mini lessons to reteach to those having difficulty.
  - Group students so that each group contains all level learners. The tier 3 learners can serve as peer helpers.
  - Assign a basic homework assignment. Require students to spend a set amount of time to work (showing effort) on the assignment rather than completing the entire assignment.
  - Allow students to choose a method for completing a project: video, PowerPoint, paper, or presentation.
- **Tier 2 Learners:**
  - Utilize foldables creating tangible products to help students digest information while incorporating several of the multiple intelligences.
- **Tier 3 Learners:**
  - Have problems posted around the room. Have students loop to specific questions based on difficulty.

Curriculum Development Resources

- Course textbook
- Graphing calculator
- Kuta software
- Quia software

Unit 3: Gathering Data

Essential Questions	NJSLS/Instructional Outcomes	Activities	Assessments
1. What is randomness and how can it be used to simulate reality? 2. What are the methods of data collection? 3. How can we conduct random sampling using stratified, cluster, systematic, and simple random sampling? 4. How can we plan and conduct surveys to avoid bias? 5. How can we plan and conduct experiments? 6. When is it appropriate to use each of the following experimental techniques will be discussed: blocking, randomization, blinding, double blinding, replication, and the placebo effect?	1. Understand methods of data collection. **S-IC3** 2. Identify characteristics of a well-designed, well-conducted survey. **S-IC3** 3. Randomly select from a population in order to have a sample from which valid conclusions can be drawn.**S-IC4**	1. Discussion of different types of bias including sampling bias, response bias, nonresponse bias, observer bias. 2. Plan a well-designed survey. 3. Draw an appropriate sample from the population. Real studies from around the country, and around the world, will be studied. Students will be required to consider the method of sampling used and how that relates to the culture and people sampled. Cultural bias will also be discussed. **Global Perspectives**	1. A survey is to be conducted in your high school. There is to be a total of 40 students in the sample. Describe how you would choose the participants if: a. There are to be the same number of freshman, sophomores, juniors, and seniors in the sample. b. There are to be the same number of males and females in the sample. c. There are no restrictions on the choice of the participants. 2. Identify the source of the bias and specify the direction of the bias. A flour company wants to know what fraction of Minneapolis households bake tier

			own bread. An SRS of 500 residential addresses is drawn and interviewers are sent to these addresses. The interviewers are employed during regular working hours on weekdays and they interview only during those hours.

Suggested Differentiation for Unit 3

- **Tier 1 Learners:**
  - Have guided notes filled out at different levels according to ability.
  - Give assignments that contain tasks of varying difficulty. Each task should focus on essential learning that all students should master, but the tasks will vary in difficulty.
  - Group students by similar interest when working on application problems.
  - Use mini lessons to reteach to those having difficulty.
  - Group students so that each group contains all level learners. The tier 3 learners can serve as peer helpers.
  - Assign a basic homework assignment. Require students to spend a set amount of time to work (showing effort) on the assignment rather than completing the entire assignment.
  - Allow students to choose a method for completing a project: video, PowerPoint, paper, or presentation.
- **Tier 2 Learners:**
  - Utilize foldables creating tangible products to help students digest information while incorporating several of the multiple intelligences.

- **Tier 3 Learners:**
  - Have problems posted around the room. Have students loop to specific questions based on difficulty.

Curriculum Development Resources

- Course textbook
- Graphing calculator
- Kuta software
- Quia software

Unit 4: From Randomness to Probability

Essential Questions	NJSLS/Instructional Outcomes	Activities	Assessments
1. How can we use the Law of Large Numbers to understand probability? 2. How can we use the Addition rule, multiplication rule, conditional probability, and independence to find the probability of an event occurring? 3. How can we combine independent random variables 4. How can we use the normal distribution to understand probability? 5. How does the Central limit Theorem play a role in Sampling distributions?	1. Interpret probability including long-run relative frequency interpretation. **S-CP2** 2. Find probabilities based on distributions of discrete random variables. **S-MD2** 3. Simulate the random behavior of events based on probability distributions. **S-MD1** 4. Calculate expected value and standard deviation of a random variable and linear transformations of a random variable. **S-ID6** 5. Differentiate between independent and dependent events.**S-CP2** 6. Identify properties of a normal distribution and use normal distribution tables **S-MD3** 7. Reach conclusions about data using the following sampling distributions: a. Sampling distribution of a sample proportion b. Sampling distribution of a sample mean c. Students use computer applet for Central Limit	1. Calculate probability based on both independent and dependent events 2. Find probability based on binomial and geometric random variables. Use graphing calculator functions binompdf and binomcdf to calculate binomial probabilities. 3. Find the mean and standard deviation for sums and differences of independent random variables. 4. Find probabilities based on normally distributed random variables. 5. Use sampling distributions reach conclusions about sample data. 6. Complete Sample Survey Project **(Project description in Appendix B)**	1. A set of 2000 measurements has a symmetric, mound-shaped distribution. The mean Is 5.3 and the standard deviation is 0.7. Determine an interval that contains approximately 1360 data values. 2. In a group of 100 scouts who took the physical exam for summer camp, 20% had type A blood. Six percent had both blond hair and type A blood. Find the probability that one scout selected at random will have blond hair, given that the blood test reveals type A. 3. Rogers High will play Memorial High in baseball six times during the upcoming season. Assume the teams are of equal ability; that is, p = .5. Within the context of a binomial experiment, what is the probability that:

	Theorem to draw different sample sizes from a population and draw conclusions of the sampling distribution. d. Difference between two independent sample proportions e. Difference between two independent sample means f. Simulate sample distributions g. t-distribution h. chi-square distribution **S-MD-7**	**Interdisciplinary Connection)**	Rogers will win 4 games and lose 2 Rogers will win *at least* 4 games?

Suggested Differentiation for Unit 4

- **Tier 1 Learners:**
  - Have guided notes filled out at different levels according to ability.
  - Give assignments that contain tasks of varying difficulty. Each task should focus on essential learning that all students should master, but the tasks will vary in difficulty.
  - Group students by similar interest when working on application problems.
  - Use mini lessons to reteach to those having difficulty.
  - Group students so that each group contains all level learners. The tier 3 learners can serve as peer helpers.
  - Assign a basic homework assignment. Require students to spend a set amount of time to work (showing effort) on the assignment rather than completing the entire assignment.

- Allow students to choose a method for completing a project: video, PowerPoint, paper, or presentation.

- **Tier 2 Learners:**
  - Utilize foldables creating tangible products to help students digest information while incorporating several of the multiple intelligences.
- **Tier 3 Learners:**
  - Have problems posted around the room. Have students loop to specific questions based on difficulty.

Curriculum Development Resources

- Course textbook
- Graphing calculator
- Kuta software
- Quia software

Unit 5: From the Data at Hand to the World at Large

Essential Questions	NJSLS/Instructional Outcomes	Activities	Assessments
1. What is the Central limit theorem? 2. How can we generalize results and types of conclusions that can be drawn from observational studies, experiments, and surveys? 3. What is the relationship between point estimators and confidence intervals? 4. How can we conduct Tests of Significance?	1) Estimation, Calculate and interpret confidence intervals for the following: a. a proportion b. difference between two proportions c. a mean d. difference between two means (unpaired and paired)**S-MD5** 2). Perform and interpret tests of significance for the following: a) a proportion b) difference between two proportions c) a mean d) difference between two means (unpaired and paired) e) slope of a least squares regression line f) chi-square tests for goodness of fit and independence **S-MD5**	1. Make conclusions about a population based on confidence intervals. 2. Use graphing calculator to create any confidence interval from data or from summary statistics. 3. Complete final Project **(Apendix C) (Interdisciplinary Connection)** 4. Make conclusions about a population based on tests of significance. a. Set up correct null and alternative hypotheses based on the context of the problem. b. Verify assumptions for the test chosen are satisfied c. Use graphing calculator to calculate test	1. Inference and Justifying Conclusions Problems i. A random sample of 64 students were asked to rate the school cafeteria food on a scale from 1 to 30. The sample mean was 22 and the standard deviation was 2.5. Determine the limits of a: 68% confidence interval 80% confidence interval 95% confidence interval 2. Suppose your world history teacher had given a particular exam for several years and has determined that the scores are normally distributed and that the population mean score on the exam is 84 and the population standard deviation is 6. Her present class of 36

		statistic and P-value of the significance test. d. Draw conclusions of the test based on the test statistic and P-value in the context of the problem.	students obtains a mean score of 86. Should she retain the hypothesis that the class is representative of the population as defined by previous classes? Test at the 0.05 level with a two-tailed test to see if the null hypothesis is valid. 3. A random sample is taken from two different high schools. In high school A, 14 of 120 students test positive for drug use. In high school B, 19 of 133 test positive. Test at the 0.05 level to see if there is a significant difference between positive drug test rates at the two schools.

Suggested Differentiation for Unit 5

- **Tier 1 Learners:**
  - Have guided notes filled out at different levels according to ability.
  - Give assignments that contain tasks of varying difficulty. Each task should focus on essential learning that all students should master, but the tasks will vary in difficulty.

- Group students by similar interest when working on application problems.
- Use mini lessons to reteach to those having difficulty.
- Group students so that each group contains all level learners. The tier 3 learners can serve as peer helpers.
- Assign a basic homework assignment. Require students to spend a set amount of time to work (showing effort) on the assignment rather than completing the entire assignment.
- Allow students to choose a method for completing a project: video, PowerPoint, paper, or presentation.

- **Tier 2 Learners:**
  - Utilize foldables creating tangible products to help students digest information while incorporating several of the multiple intelligences.
- **Tier 3 Learners:**
  - Have problems posted around the room. Have students loop to specific questions based on difficulty.

Curriculum Development Resources

- Course textbook
- Graphing calculator
- Kuta software
- Quia software

Unit 6: Accessing associations between variables

Essential Questions	NJSLS/Instructional Outcomes	Activities	Assessments
1. How can we compare two proportions? 2. How can we compare two means? 3. How can we test for differences before and after an event? 4. How can we test for independence of categorical variables? 5. How can we test for independence of quantitative variables?	1) prediction interval for predicted value) 2) slope of a least squares regression line 3) chi-square tests for goodness of fit and independence **S-MD5**	1. Use graphing calculator to create any confidence interval from data or from summary statistics. 2. Make conclusions about a population based on tests of significance. 3. Set up correct null and alternative hypotheses based on the context of the problem. 4. Verify assumptions for the test chosen are satisfied 5. Use graphing calculator to calculate test statistic and P-value of the significance test. 6. Draw conclusions of the test based on the test statistic and P-value in the context of the problem. 7. Complete Final Project **(Project description in Appendix C)**	1. Given SAT scores of 35 randomly selected students before and after a review course, Perform a paired difference t-test to test whether or not there is a significant increase in test scores after completing the review course? 2. We hypothesize that the distribution of appointments at a dialysis center is uniformly distributed across the four seasons of the year. If there are 220 in the spring, 245 in the summer, 215 in the fall and 255 in the winter, can you reject the null hypothesis that there is a uniform distribution? 3. Given the number of hours studying and test score for 50 students, test the null hypothesis that there is no association between studying and test score at the 0.05 significance level.

Suggested Differentiation for Unit 6

- **Tier 1 Learners:**
  - Have guided notes filled out at different levels according to ability.
  - Give assignments that contain tasks of varying difficulty. Each task should focus on essential learning that all students should master, but the tasks will vary in difficulty.
  - Group students by similar interest when working on application problems.
  - Use mini lessons to reteach to those having difficulty.
  - Group students so that each group contains all level learners. The tier 3 learners can serve as peer helpers.
  - Assign a basic homework assignment. Require students to spend a set amount of time to work (showing effort) on the assignment rather than completing the entire assignment.
  - Allow students to choose a method for completing a project: video, PowerPoint, paper, or presentation.
- **Tier 2 Learners:**
  - Utilize foldables creating tangible products to help students digest information while incorporating several of the multiple intelligences.
- **Tier 3 Learners:**
  - Have problems posted around the room. Have students loop to specific questions based on difficulty.

Curriculum Development Resources

- Course textbook
- Graphing calculator
- Kuta software
- Quia software

Additional Suggested Modifications for Units

Below is an additional list of modifications and accommodations opportunities. This includes, but is not limited to,:

1. English Language Learners.
    a. Read written instructions.
    b. Model and provide examples
    c. Extended time on assessments when needed.
    d. Establish a non-verbal cue to redirect student when not on task.
    e. Students may use a bilingual dictionary.

English Language Development Standard 3: Language of Mathematics: English language learners communicate information, ideas and concepts necessary for academic success in the content area of mathematics.

2. Special Education/504 Students.
    a. Extended time on assessments when needed.
    b. Preferred seating to be determined by student and teacher.
    c. Provide modified assessments when necessary.
    d. Student may complete assessments in alternate setting when requested.
    e. Establish a non-verbal cue to redirect student when not on task.
    f. Maintain strong teacher / parent communication.
    g. Conversion chart

New Jersey Student Learning Standards - Technology

8.1 Educational Technology: All students will use digital tools to access, manage, evaluate, and synthesize information in order to solve problems individually and collaborate and to create and communicate knowledge.

A. Technology Operations and Concepts: Students demonstrate a sound understanding of technology concepts, systems and operations

B. Creativity and Innovation: Students demonstrate creative thinking, construct knowledge and develop innovative products and process using technology.

C. Communication and Collaboration: Students use digital media and environments to communicate and work collaboratively, including at a distance, to support individual learning and contribute to the learning.

E: Research and Information Fluency: Students apply digital tools to gather, evaluate, and use information.

F: Critical thinking, problem solving, and decision making: Students use critical thinking skills to plan and conduct research, manage projects, solve problems, and make informed decisions using appropriate digital tools and resources.

Students will be instructed on how to use TI-84 Plus graphing calculators, Microsoft Excel, and Fathom to generate graphs, compute statistics, and analyze data. Such technology will be required on homework, projects, and other assessments.

*See Activities/Appendix for further Technology Integration.

Career Readiness Practices

- CRP1. Act as a responsible and contributing citizen and employee.
- CRP2. Apply appropriate academic and technical skills.
- CRP4. Communicate clearly and effectively and with reason.
- CRP6. Demonstrate creativity and innovation.
- CRP8. Utilize critical thinking to make sense of problems and persevere in solving them.
- CRP11. Use technology to enhance productivity.

NJSLS 9.2 - Career Awareness, Exploration, and Preparation

9.2.8.B.3 Evaluate communication, collaboration, and leadership skills that can be developed through school, home, work, and extracurricular activities for use in a career.

Supplemental Teacher Resources	Supplemental Student Resources
American Statistical Association **https://www.amstat.org/ASA/Education/K-12-Statistics-Education-Resources.aspx** **Statistics Education Web** **https://www.amstat.org/asa/education/stew/home.aspx** **The World of Statistics** **http://www.worldofstatistics.org/** **Stats.org** **http://senseaboutscienceusa.org/stats/**	**Khan Academy** **Desmos**

Appendix A: Statistics Displaying Data Project

Due: _____

Objective

To use descriptive statistics to organize, summarize, and graphically represent data.

Project Requirements

1. Collect quantitative data on a variable of interest to you. List the values of the observations. ($n \geq 30$) Sample topics are listed on the back of this page.
2. Determine if the variable is discrete/continuous. Explain why.
3. Make a frequency table from the data.

☆ 4. a. Construct a dotplot from the data. Describe the main features of the dotplot.
b. Construct a stem and leaf plot from the data. Describe the main features of the stem and leaf plot.
c. Construct a histogram from the data. Describe the main features of the histogram.

☆ 5. Determine the shape of the distribution of the data.
6. Calculate the mean, median and mode of the data.
7. Determine whether the mean or median is a better estimate of the center for your data? Explain why.
8. Calculate the range, variance and standard deviation of the data.
9. Does the empirical rule apply to your data? Explain why or why not?
10. Calculate the 5-Number Summary for the data.
11. State and interpret the quartiles for the data.
12. Calculate the interquartile range for the data.
13. Are there any outliers in your data? Show your calculations.
14. Construct a boxplot for the data.
15. Calculate the z-score for the last observation you collected based on your data. Show your calculation. Interpret this z-score.
16. Give a brief 2-minute presentation showing the results of your descriptive statistic analysis.
17. Complete all starred requirements using a computer (preferably Excel).

Appendix B: Statistics Baseball Project

Objective: To describe the association between quantitative variables and to predict values of the response variable based on the values of the explanatory variable. In baseball terms, we will find which variable is most correlated with winning and predict the number of wins based on this variable. All starred steps must be done on the computer using EXCEL.

Requirements

☆1. Collect the following baseball data from the _____ baseball season:

Hitting Stats: *R, HR, SB, BA,* _____

Pitching Stats: ***Wins***, *R, SO,* _____

Select one more Batting and one More Pitching statistic that you want to explore

Directions

a. Go to http://www.baseball-reference.com/leagues/MLB/

b. click on the link for that year which brings you to Team and League Standard Batting

c. Copy and paste batting and pitching statistics into EXCEL. (be sure to select everything from "Team and League Standard Batting" down to the bottom of the page.

d. Put Batting on Sheet 1 and rename Batting, Put Pitching on Sheet 2 and rename Pitching

e. Highlight then Copy/Paste pitching wins column into batting Stats page.

f. Clean up the data and delete all unnecessary columns. (be careful)

g. Print out batting data landscape on one page

h. Print out pitching data landscape on a separate page.

☆2. Use EXCEL to construct 8 scatterplots that plots each quantitative variable as the x variable against *Wins* as the y variable. (Note: since every win is credited to a pitcher on the team, pitcher wins is the same as team wins. Rename Sheet 3 Scatterplots and put all scatterplots there.

INSERT>CHART>SCATTERPLOT>SELECT COLUMNS>LABEL AXES>ADD AS OBJECT

Add the regression equation (trend line) and r^2 value by right clicking on a point on the scatterplot and clicking "add trendline"; then on the options tab add the equation and r^2 value.

Print scatterplots all on the same page (create as objects in one sheet instead of giving each scatterplot its own sheet.) On a separate page, discuss any associations visible in each scatterplot.

☆2. Calculate and interpret the correlation for each of these eight variables with *Wins*. Comment on the validity of the correlation.

3. Interpret r^2 for each of the eight variables with *Wins*.

4. Are there any influential points in any of your scatterplots? If so, identify them by *Team* name.

5. Based on the results of steps #2-4, identify which of the eight variables seems to be the best predictor of *Wins*? Why?

6. Use the regression line of the variable you selected in step #5 to predict the number of *Wins* the New York Yankees ***should*** have had, based on this variable.

7. Calculate and interpret the residual for the New York Yankees based on the same regression line.

☆8. Use EXCEL to construct a residual plot for the same regression you chose in step #5. Does this residual plot verify that the least squares regression line fits the data well? Explain. (Show the predicted values, residuals and residual plot all on the same page.)

Display the data collected on your *individual* topic question from the survey.

1. Construct a table to display the data.
2. Make an appropriate graph based on your data.
3. If the data is categorical, calculate the relative frequency of each category.
   If the data is quantitative, calculate the mean, median, mode, standard deviation.
4. Summarize the conclusions that you found based on the graphs and calculations. Explain the results of the sample.

Results Presentation

Students must present their findings to the class in a clear and concise manner with emphasis on the use of proper statistical vocabulary and terminology.

Appendix C: Statistics Final Project

Overview
Each group of 2-3 students will prepare a poster.
The project will include an analysis of data from an existing dataset or a dataset generated by the students. The end product is a poster (on 22 by 28 inch poster board) that will be presented to a group of fellow students at a poster session scheduled for **Wednesday, June 8**. During the poster presentations you will also be responsible for reviewing the poster presentations of other groups in your class. Posters will be graded after the presentations are completed.

Proposal
A proposal for the project will be required from each group. The proposal must be a single page that gives a brief statement of your topic, the issue or question you are addressing, the source of the data (be specific), how the data was (or will be) collected, the sample size, and the graphical and statistical analysis techniques that will be used.
The proposal will be due by **Wednesday June 1**.

Poster Presentation Reviews
At the completion of each poster presentation, the other members of the class will independently review the poster and presentation by completing the attached review form (blank review forms will be provided).

Analysis of Data
The project will consist of 1) locating or collecting data in your field of interest that will require both graphical displays and statistical analyses, 2) preparing a poster that presents the aforementioned graphical display(s) and results of statistical tests, 3) verbally presenting the poster to fellow students at a poster presentation session, and 4) reviewing the poster presentations of other class members.

Step 1 – Selection of the Dataset or Design of the Study

Identify a field of potential interest, such as medicine, biology, economics, sports, etc. The internet has a vast source of datasets on almost any subject. Just search for your favorite topic using an internet search engine. Be sure to select a dataset that has a clear description of its context (e.g. background, study design, specification of variables, etc.). Sports are a good source for datasets (e.g. try www.baseball-reference.com). If you have already collected your own dataset or plan on collecting your own data, you will need to be concerned about aspects of study design. What are the primary study objectives? What are the primary variables? How was the sample size determined? How were the individuals selected? How are you avoiding bias? What statistical techniques will be used? If you plan on collecting data from other individuals (e.g. surveying your fellow students), please see the instructor to discuss issues of consent and confidentiality before you start.

Step 2 – Conduct your Study and Analyze your Data

The best approach to data analysis is to have a clear plan that focuses on your primary objectives.
Your plan should include methods for understanding your data [graphs and tables] and the analysis [the analysis method(s), key graph(s) and table(s) reporting the analysis results].

Step 3 – Preparation of your Poster

The presentation should fit onto a 22 by 28 inch poster board, use a font size of 28 or higher throughout (for readability), and should consist of the following elements:

1) A brief descriptive title of the research topic.

2) If you used data collected by someone else, you should give a complete formal statement of the reference for the data. If you generated the dataset yourself, you should give a single sentence summary description of the dataset accompanied by the statement "collected and analyzed by (your name)". If someone else assisted you in the collection of the data, then their name should also be given.

3) An abstract giving the following:
   a. Background (10-40 words)
   b. Methods (10-40 words)
   c. Results (20-100 words)
   d. Conclusions (10-40 words)

4) A statement of your study design, indicating the specific source of the data, the sample size, and how the data was collected (50-100 words).

5) A table describing the key characteristics of your dataset (e.g. frequency table or contingency table).

6) A graphical display(s) useful for understanding the results of your study. The most common type of graphical display is the histogram or bar chart.

7) A paragraph describing the main features of the graphical display(s) (50-100 words).

8) The results of the main statistical tests you conducted. In general, the most common types of statistical analyses are t-tests, chi-square tests, and regression. All parts of the significance test or confidence interval must be shown.

9) A paragraph describing the conclusions of the statistical test results (50-100 words).

10) A single paragraph describing the strengths (20-100 words) and weaknesses (20-100 words) of the graphical display(s) and statistical test results (e.g. weaknesses may involve a small number of subjects, suboptimal study design, unmet assumptions, or variables that you would have liked to control for, but were not available in your data).

Important notes

1. Focus on presenting only 1-2 analysis points clearly. Don't try to present lots of complex material. One of the goals of this project is for you to develop the skill to be able to present the results of an analysis in a clear, succinct manner that will maximize the likelihood that it will convince others of your main conclusions. Do not fill up your poster with unnecessarily complex detailed graphs, analyses and text.

2. Word limits and font sizes have been designed for this poster project to encourage the development of an effective poster. Follow them.

3. If your current text and material doesn't fit onto the poster, your first thought should be to condense your text to just the essential elements. The skill of condensing text to fit within the word limits of professional journals is a skill nearly all will eventually need in their career. It is

recommended that the student write at least twice as much text as needed for the poster and then condense it down to reach the word limits. Text initially written to a target word limit often looks disjoint and sloppy. Also, please do not have any material hang over the edges of the poster.

4. Discussing strengths and weaknesses. Referring to the poster paragraph describing the strengths and weaknesses of the graphical display(s) and statistical test results (section 10), please carefully note the following points. Study weaknesses (e.g. suboptimal study design) should be identified using specific comments (rather than general comments) and a statement should be included that indicates how the problem could be solved in a practical manner. For example, a true experiment may not be possible, but it might be feasible to identify suitable individuals that could be used as controls. If there is a problem with a lack of testing of assumptions, the specific unreasonable assumptions should be identified, the method of testing or assessing the assumption should be identified, and a statement added giving your balanced judgment regarding the extent of potential impact the violation of this assumption would have on study results. If you can't show causality, describe why not? What causal relationship would you like to demonstrate if you could?

Poster Project Review Comments

Name of students who prepared the poster _____
_____
_____

Please use the following rating scale for the 5 questions below
(5 – Outstanding, best in the class; 4 - Very clear and well-done; 3 – Very good, about average for the class; 2 – Good, but could be improved; 1 – Some weaknesses, needs to be improved)

_____ Abstract
_____ Graphical display(s) and descriptive paragraph of the concept/conclusions
_____ Statistical tests (tables) and descriptive paragraph of the concept/conclusions
_____ Description of the strengths and weaknesses of the graphical display(s) and statistical test results
_____ Verbal presentation

Please give at least 20 words describing the strength of the poster and/or presentation
_____
_____
_____
_____
_____

Please give at least 20 words of constructive criticism regarding the poster and/or presentation
_____
_____

_____
_____
_____

_____ Name of student who prepared this review

Appendix D: New Jersey Student Learning Standards Statistics

Common Core State Standards S-ID, S-IC, S-CP and S-MD are used throughout curriculum. Students will develop their mathematical skills as well as their problem solving strategies and their ability to interpret information and data. They will become efficient and creative problem solvers and will acquire an understanding of mathematical concepts. Students will be able to solve problems numerically, graphically, and analytically. They will use technology to reinforce concepts and also as an efficient problem solving tool.

Appendix E: New Jersey Scoring Rubric

Scoring Guide for Mathematics Open-Ended (OE) Questions
(Generic Rubric)

3-Point Response

The response shows complete understanding of the problem's essential mathematical concepts. The student executes procedures completely and gives relevant responses to all parts of the task. The response contains few minor errors, if any. The response contains a clear, effective explanation detailing how the problem was solved so that the reader does not need to infer how and why decisions were made.

2-Point Response

The response shows nearly complete understanding of the problem's essential mathematical concepts. The student executes nearly all procedures and gives relevant responses to most parts of the task. The response may have minor errors. The explanation detailing how the problem was solved may not be clear, causing the reader to make some inferences.

1-Point Response

The response shows limited understanding of the problem's essential mathematical concepts. The response and procedures may be incomplete and/or may contain major errors. An incomplete explanation of how the problem was solved may contribute to questions as to how and why decisions were made.

0-Point Response

The response shows insufficient understanding of the problem's essential mathematical concepts. The procedures, if any, contain major errors. There may be no explanation of the solution, or the reader may not be able to understand the explanation. The reader may not be able to understand how and why decisions were made.

The generic rubric above is used as a guide to develop specific scoring guides or rubrics for each of the open-ended (OE) questions that appear on the New Jersey fourth-grade (ESPA), eighth-grade (GEPA), and eleventh-grade (HSPA) proficiency assessments in Mathematics. The generic rubric helps ensure that students are scored in the same way for the same demonstration of knowledge and skills regardless of the test question.