***Statistical Understanding Through Competitive Decisions and Results or Should We, in the Words of Nike "Just Do It"[1]?***

by Lew Davidson, PhD, 2013 CTI Fellow
Mallard Creek High School

This curriculum unit is recommended for:
AP Statistics courses/typically taught in High Schools Grades 11-12

**Keywords:** AP Statistics, Sports Analytics, Probability, Moneyball, Sabermetrics, Brackets, Won-Loss Standings, False Positives, Testing Errors, Decision Theory

**Teaching Standards:** See **Appendix "Implementing Common Core Standards"** for teaching standards addressed in this unit.

**Synopsis:** This Curriculum Unit, intended for experienced AP Statistics teachers use at the start of a course, proposes a framework for teaching both to enhance student engagement and to enable a deeper understanding of statistical reasoning. It provides sports-based classroom activities to address specific knowledge requirements and to inspire student generated collaborative group projects based on sports or more general competitive situations. The intuitively inspiring merging of Sports and Math is evolved into the Framework of Competitive Decision Making and Math to address many diverse student interests.

Provided student activities address 1) personalization of AP Statistics incorporating key trends, 21st century skills, needed statistical insight, and the objectives in the CMS Strategic Plan 2018; 2) testing for the alleged use of performance enhancing drugs in Andrus Veerpalu's Gold Medal 2006 Olympics performance; 3) the current high school football playoff brackets for the NC State Championships to explore associations between a team's ability to dominate Won-Loss standings and the size of its student population; and 4) what play to use on a fourth down.

*I plan to teach this unit during the coming year to 62 students in AP Statistics.*

***Statistical Understanding Through Competitive Decisions and Results or Should We, in the Words of Nike "Just Do It"?***

*Lew Davidson*

## Prelude

Nike popularized the phrase of "Just Do It" and in fact for much of what happens today in both sports and other competitive encounters, that is exactly what is happening. People "Just do it" based on conventional approaches. Can we do better? Bill James[2], an avid Boston Red Socks fan, is viewed by many as the catalyst for not accepting conventional practice on its face, but rather probing for the data that objectively reveal the most probable associations between decisions and results (Aha Statistics!). James coined the phrase "Sabermetrics", but this concept became more widely know through a book and movie called "Moneyball". The impact of what is basically data driven decision making is spreading widely and now is known as "Sports Analytics" one of the hottest areas in today's sports world. Professor David Romer observed in a benchmark work[3] that sports can serve as a laboratory for economic analysis. I would add also for many decisions in life. So let's use Statistics to investigate areas requiring competitive decisions that are of interest to each student and see if we can take "Just do it" to a new plane where it is based on established results and not only conventional wisdom.

## Introduction

### Rationale

The announcement of a CTI 2013 seminar with a focus on the use of sports to enhance math education created great excitement. The potential of such a strategy is immediately apparent. Coupling this with my quest of continually enhancing the teaching and impact of AP Statistics, I applied to join the seminar, with a proposal that envisioned a unit for an AP Statistics course with specific guidance from the outstanding work of Tabor and Franklin[4]. This proposal concept has been a beacon throughout the development of this unit.

### Acknowledgments

This Curriculum Unit was developed as part of the tremendous 2013 CTI Math and Sports Seminar experience. With much appreciation, it is a pleasure to share that this unit has profited immensely from interactions with our Seminar Leader, colleague 2013 Seminar Fellows, and content experts who kindly shared their time and insight at seminar meetings. The author is very fortunate and deeply appreciated to have had the opportunity to interact with such dynamic, knowledgeable and helpful professionals.

- Our Seminar Leader is Professor Tim Chartier, Davidson College
- 2013 Seminar Fellows are:
    - Lisa Ashworth, Barringer Academic Center
    - Cassandra Black, Kindergarten, Ashley Park Pre-K-8
    - Stephanie Coggins, Bailey Middle
    - Madalina Corneanu, Harding University High School
    - Monica Echols, Francis Bradley Middle School (our seminar coordinator)
    - Cynthia Hicks, Kindergarten, Steele Creek Elementary School
    - Susan Jones, Kindergarten, Steele Creek Elementary School
    - Mindy Passe, Barringer Academic Center
    - Michael Pillsbury, now CMS Mathematics Curriculum Coordinator
    - Latoya Scott, Kindergarten, Ashley Park Pre-K-8
- 2013 Seminar Guest Speakers included
    - John Brenkus, ESPN Sports Science Host
    - Bob McKillop, Davidson Wildcats, Basketball Coach
    - Jason Rosenfeld, Director of Basketball Analytics, Charlotte Bobcats
    - Henry Segerman, Assistant Professor, Oklahoma State (3-D Printing)
    - Donour Sizemore, Michael Waltrip Racing
    - Glen Whitney, Co-Executive Director, National Museum of Mathematics

**Content Objectives**

The objectives for this unit are twofold.

1. The first objective is to proposes a framework for teaching AP Statistics that is anticipated to enhance student engagement and enable a deeper understanding of how to use statistical reasoning to better student's personal and professional lives.
2. The second objective is to provide sports based activities both to address specific AP Statistics knowledge requirements and to inspire student generated collaborative group projects based on sports or more general competitive situations.

The context for these objectives is to address the identified trends in AP Statistics and inspire an ever increasing number of students to mastery a course equivalent to a first year college level Statistics course. In particular this Curriculum Unit is intended for experienced AP Statistics teachers to use at the start of the course. As such, this unit can set the course tone and provide a framework for individual student personalized course themes. Factors considered in the development of these objectives included:

1. Adherence to the College Board AP Course, which include many Common Core State Standards
2. Impact on Student Learning for the entire course

3. Likelihood of enthusiastic and proactive student engagement (Student engagement, especially in 90 minute class blocks, is a focus item)
4. Likelihood of an increase in student engagement with a "Sports" oriented theme consistent with our Seminar's fundamental belief of the synergism between Sports and Mathematics.

The College Board's AP Statistics Course description provides the specific standards about which this unit seeks to establish understanding using a Sports context[5]. The standards that are specifically addressed in this unit are:

II Sampling and Experimentation: Planning and conducting a study (10%–15% *[of the course content])*

Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. This plan includes clarifying the question and deciding upon a method of data collection and analysis.

A. Overview of methods of data collection

1. Observational study

B. Planning and conducting surveys

1. Characteristics of a well-designed and well-conducted survey
2. Populations, samples and random selection
3. Sources of bias in sampling and surveys
4. Sampling methods, including simple random sampling, stratified random sampling and cluster sampling

D. Generalizability of results and types of conclusions that can be drawn from observational studies, experiments and surveys

III Anticipating Patterns: Exploring random phenomena using probability and simulation. Probability is the tool used for anticipating what the distribution of data should look like under a given model.

A Probability

Objectives Overview

This unit is structured in two parts since it shares not only specific application of sporting, or more generally competitive, scenarios to the teaching of Statistics at the AP [Advanced Placement, i.e. college] level; but also in what is believed to be a new focus, using competitive scenarios as a framework for teaching AP Statistics in a manner personalized for each student. In the first part this unit discusses considerations for the Framework and then the second part provides specific applications to inspire and guide students in developing projects of special interest to them.

In "Part 1: The Framework", the intuitively inspiring merging of Sports and Math is evolved into the wider reaching Framework of Competitive Decision Making and Math. This evolution resulted from a finding that not all students are inspired by sports. It is believed that this enlarged framework will significantly enable student engagement in a manner that addresses the many diverse interests of our students. Thus the framework deals with the challenge of inspiring not only student sport enthusiasts, but also student non-sport enthusiasts. It also examines key trends in AP Statistics; 21st century skills and insight that will be required of those who master the course content; and the associated objective in the CMS Strategic Plan 2018. Part 1 includes a student activity designed to facilitate student personalization of the use of statistics in a field of interest to the student.

In "Part 2: The Applications" three specific scenarios are investigated both to inspire by demonstrating the power of these concepts and to serve as examples for the personalized projects expected from students.

1. The first student activity is based on the ongoing work of ensuring that athletic achievements are obtained without the use of performance enhancing drugs.[6] Using the ongoing controversy over allegations that Andrus Veerpalu had use such drugs in his Gold Medal 2006 Olympics Cross-country Skiing win as a Case Study, students can directly witness the issues involved in trying to reach a conclusion based on a test. Students in collaborative teams are then expected to apply, with prior teacher agreement, these or similar concepts to a competitive activity of their choice.

2. The second student activity is based on the current high school football playoffs for the NC State Championship. There is much excitement, analysis and discussion concerning the playoff brackets. During that process the following question was posed: Is a team's ability to dominate the Won-Loss standings associated with the size of the student population at the team's school? As a Case Study the 2013 Western Region brackets are examined to answer this question. This case study enables students to see how data that is readily available on the internet may be analyzed to provide answers to interesting questions. Students in collaborative teams are then expected to apply, with prior teacher agreement, these or similar concepts to a competitive activity of their choice.

3. The third student activity explores a critical football decision as to what play to use on a fourth down. Conventionally coaches call for a punt on a fourth down ("Just Do It")!. However, there is strong evidence that in many situations a team would be better trying for a first down and not punting. This Case Study of the application of Sports Analytics[7] to football is the subject of much ongoing analysis and contention in today's media. Students in collaborative teams are then

expected to apply, with prior teacher agreement, these or similar concepts to a competitive activity of their choice.

**The Teaching Strategy - Part 1: The Framework**

What is AP Statistics?

AP Statistics is an algebra based study of Statistics that is designed for equivalency to typical first-year college statistics courses. As such the College Board, who oversees the course description and implementation, cite Algebra 2 (now Math III in CMS) as the only prerequisite. AP Statistics affords the wonderful opportunity for students to gain their introduction to Probability and Statistics in a small classroom setting and assimilate the important concepts of Statistical Thinking in the very interactive setting of a high school classroom.

AP Statistics is designed to develop mastery of all the content found in a typical First Year Undergraduate Introductory Statistics course. The degree of achievement of this objective is validated by means of a comprehensive and rigorous AP Exam at the end of course encompassing both Multiple Choice and Free Response Questions. Thus, students must both develop a strong understanding of the objectives articulated in the course standards and refine/enhance their critical thinking and communication cognitive skills.

*Trends in AP Statistics*

Trend 1: Increasing participation and achievement:
There are important current trends in CMS for AP Statistics with the CMS Strategic Plan 2018 establishing key focus areas. In one area: Focus Area III: Access to Rigor, CMS has set two specific goals for Advanced Placement courses:

a. Increase the "Percentage of graduates successfully completing at least one AP/IB course" from 44% in 2012-2013 to 60% in 2017-2018 (IB stands for courses meeting the International Baccalaureate standards)
b. Increase the "Percentage of AP Exams (3 or higher)" from a TBD 2013-2014 baseline to a TBD 2017-2018 goal (AP Exams are graded on a 5 point scale with scores of 3 to 5 resulting in college credit for the courses at many colleges).

By looking at these goals it is clear that an ever increasing number of AP students in CMS will be what is currently known as "non-traditional AP students". Unfortunately such students often do not possess a love of learning and are not driven by an internal goal to master material. To address their needs, ever improving strategies for engagement need to be implemented. Given the large estimated percentage of students who are inspired by sports of one form or another, it seems clear that sports based activities have a high probability of enhancing student engagement.

Trend 2: Increasing relevance of AP Statistics course content to lifetime needs and decision making.

Professor Jessica Utts, the new AP Statistics Chief Reader Designate in her keynote address[8] at the 2013 AP Statistics Examination Reading challenged AP Statistics teachers as she discussed the question: "What do Future Senators, Scientists, Social Workers, and Sales Clerks Need to Learn from Your Statistics Class?" She identified the current "Top 10 Important Topics" for AP Statistic teacher focus within the content of the College Board prescribed course description:

1. Observational studies, confounding, causation
2. The problem of multiple testing
3. Sample size and statistical significance
4. Why many studies fail to replicate
5. Does decreasing risk actually increase risk?
6. Personalized risk
7. Poor intuition about probability/expected value
8. The prevalence of coincidences
9. Surveys and polls – good and not so good
10. Average versus normal

*Increasing Engagement*

Most AP Statistics course activities use primarily data from a wide variety of actual real life scenarios. Mathematics has such a broad application to sports that many available topics are inherently interesting to a large number of students, but typically not all students in the class. Such interest can enhance engagement and thus a goal is to seek activities that engage a large proportion of the class.

It seems intuitively apparent that if in fact students could personalize the area of their activity to their own interests, engagement might be optimized. Sports, or more broadly competitive activities provide such a framework. Moreover, the broad application of most, if not all, statistical techniques to sports provides the opportunity for students to select a personal theme for the course. Thus, students can experience the excitement and satisfaction of building a portfolio of knowledge and skills to enhance their enjoyment, and for many their performance, of a sport of particular interest.

I began reflecting on this topic and discovered:

- Challenge #1: the Common Core State Standards (CCSS) do not include the AP Statistics focus on "Gathering Data", e.g. Observational Studies and as such this topic is not covered in CCSS oriented resources. Therefore a sports oriented treatment of Gathering Data became a clear matter of priority for this unit.

- Challenge #2 Sports are appealing, but – ***DISCOVERY*** – a significant number of students have no interest in Sports, in fact --- Sports are a "Turn-Off" for them

These are pivotal Challenges! So first let's deal with the individual issues inherent in understanding Sports from a broad perspective and ask "What is a Sport"?

1. What is the nature of competition?
2. What are the important aspects of a) winning and b) success?

What is a "Sport"? Off to the Dictionary we go to find: A Sport is "an athletic activity requiring skill or physical prowess and often of a competitive nature, as racing, baseball, tennis, golf, bowling, wrestling, boxing, hunting, fishing, etc."[9] A key word in this definition is "athletic", which primarily means "physically active and strong".[10] Strong in part is defined as "…mentally powerful or vigorous … or especially able, competent, or powerful in a specific field or respect: She's very strong in mathematics."[11] The Dictionary's examples imply physical sports, as I suspect most of us think of when we hear "Sports". Conceptually however, the concept of "Sports" is inherently much broader.

So whereas I primarily had focused on the physical aspects of sports and athletics, when I reflected on our Seminar, I soon realized that this unit could have a much broader perspective and thus potentially enhance the degree of engagement to hopefully most, if not all students. As the coach of our school's Mathlete team, the concept of mentally based sports is quite familiar. One can think also of the games of poker (regularly broadcasted on ESPN), Chess, Spelling Bees, and Quiz Bowls as primarily competitions of a mental nature. Under "skills' we can include many interest competitions such as those in art, debate, and marketing (DECA).

A possible categorization might be:

1. Responsive Interaction - e.g. chess games or volleyball
2. Simultaneous with little or no opposing Interaction - e.g. a swim meet, IRON CHEF or a "sing off" as duals (a mode used on the VOICE)
3. No direct opposition - e.g. individual time trials, a talent show, a Sudoku game, competing for a college scholarship, math competitions (Inner Competition)

The notion "of a competitive nature" is a major consideration! This unit's focus on "Sports" is rooted in the broad context of this notion. This immediately leads to a categorization of activities based on the nature of the competitiveness:

1. Whether the activity is engaged in as a team or as an individual
2. Absolute (e.g. breaking a world record) vs. relative (e.g. more goals than the other team)
3. Mode:
    a. Simultaneous with direct opposing interaction - e.g. wrestling or football actions[12]. Note by the term action we are referring to the relatively short period after an orderly resumption of intense competitive interactions
    b. Simultaneous with little or no opposing interaction - e.g. a swim meet or a "sing off" as duets (e.g. a mode used on the TV reality competition show called the VOICE),
    c. Responsive with alternating one sided interaction - e.g. chess games or volleyball
    d. No direct opposition during the activity - e.g. individual time trials, a talent show, a Sudoku game, competing for a college scholarship.

This line of thought brings one to what I will call the "Ultimate Sport": Achieving our own potential! Isn't this an absolute competition? What are the competitive forces for this competition? Can we broadly define an achievement in this context? A fantastic Coach and English Teacher, John Wooden, concluded that success is more important than winning. He developed his definition of success as "*Success is peace of mind which is a direct result of self-satisfaction in knowing you did your best to become the best you are capable of becoming.*"[13]

**Part 2 Classroom Activities**

*Activity #1:*

The first activity in this unit will follow the above described Teaching Strategy and provide a convenient means for students to personalize AP Statistics. Students will receive as an essay prompt the previous thoughts for their reflection on their personalized choice. The project deliverable will be an essay wherein they discuss how and why they picked their “Sport” or Competitive Activity as a theme for personalizing AP Statistics (especially for projects and open ended activities). Student choices will be used as drivers for the themes of our in class activities. The essay will have two parts:

1. A "Sports" topic that will serve as a theme for each student's individual assignments throughout our course. Students will be encouraged to work collaboratively with others who share similar interests.

2. Their personal plans for achieving success, both in our course and in their lives, either as defined by Coach Wooden or as otherwise defined.

With reference to "Content Objectives" above, the two major aspects of gathering data that must be understood are Observational Studies and Experiments. Students should determine projects to explore these areas. The specific content of the project will be the students' choice in accordance with the following hierarchy: [please note that this hierarchy is only for the purpose of inspiring and determining topics that are consistent with our competitive theme. Topics from any aspect of this hierarchy are equally capable of gaining credit]. Project content must receive prior approval from the teacher.

1. A conventional sports related topic
2. A competitive based topic
3. An equivalent topic based on consultation with the instructor. (As part of such a consultation, the student shall explain specifically why neither #1 or #2 above is appealing to the student.)

It became clear that because of its potential power, this activity would be very effective in the first unit in an AP Statistics course. Thus, this unit will be on "Gathering Data" to provide a rigorous study of Observational Studies, including Surveys; Experiments; and Simulation. I discussed this concept with Daren Starnes and Josh Tabor, authors of key Statistics resources. Of particular focus was the fact that many AP Statistics teachers and books start with Analyzing Data, not Gathering Data. Daren pointed out that experienced teachers are comfortable starting with Gathering Data because they already have the insight that is needed, whereas Analyzing Data draws upon prior math skills and provides a more comfortable start for new teachers. Thus, the course placement of this activity is a teacher's option.

"Gathering Data" is a single chapter of our textbook, but it covers material that accounts for 10%–15% of the AP Exam content. It requires Critical Thinking and Effective Communications at the highest level. It is the first part of any statistical investigation. Simply put the quality of a statistical investigation can be no better than the quality of the input data

*Rubrics for Activities*

Each year the American Statistical Association (ASA) conducts student poster and project competitions with relevant rubrics. For this unit student activities should be assessed as per the concepts demonstrated in the ASA rubric.[14]

*Activity #2*

This Activity introduce students to the necessity for understanding the perspectives of a statistical question. It addresses three of the Important Topics articulated by Professor Utts as described above

1. Observational studies, confounding, causation
3. Sample size and statistical significance
7. Poor intuition about probability/expected value

Societies inherently seek fair competitions. To ensure fairness, various requirements (often called rules or laws) are established and means, i.e. tests, are established to ensure adherence to the requirements. A challenge is to have foolproof tests. Often uncertainties exist to the degree that those who do not support the test results can validly dispute findings. Most are familiar with the story of Lance Armstrong and his many wins of the grueling Tour du France cycling competition. Because that story is so well known, I carefully considered using it for this activity. However, that story is now a fully settle situation due to Lance's confession of guilt and I resolved to find a current active scenario.

The Olympics very seriously focus on the requirement of not using performance enhancing drugs to win competitions. Let us consider the Cross Country Skiing - Men's 15Km Competition at the Turin 2006 Winter Olympic Games.

1. (Setting the stage in an engaging manner) Students view Andrus Veerpalu Gold Medal 2006 Olympics performance[15]

2. (Introducing the dilemma with a April 2011 news report) "Andrus Veerpalu tests positive for HGH"[16]

3. (Build excitement) Enter the “Defense Team” claiming "Innocence" and filing an Appeal[17] "a team of Estonian Ph.D.s disclosed the findings of a 1.5-year study which they said categorically shows that the World Anti-Doping Association's test is unreliable."

4. Enter the World Anti-Doping Agency [WADA][18] "The principle of WADA's work is largely built on trust and it would be unfortunate if this trust were damaged," Port said, adding that, according to WADA scientists, the chance of Veerpalu being wrongly deemed guilty is one in 10,000.

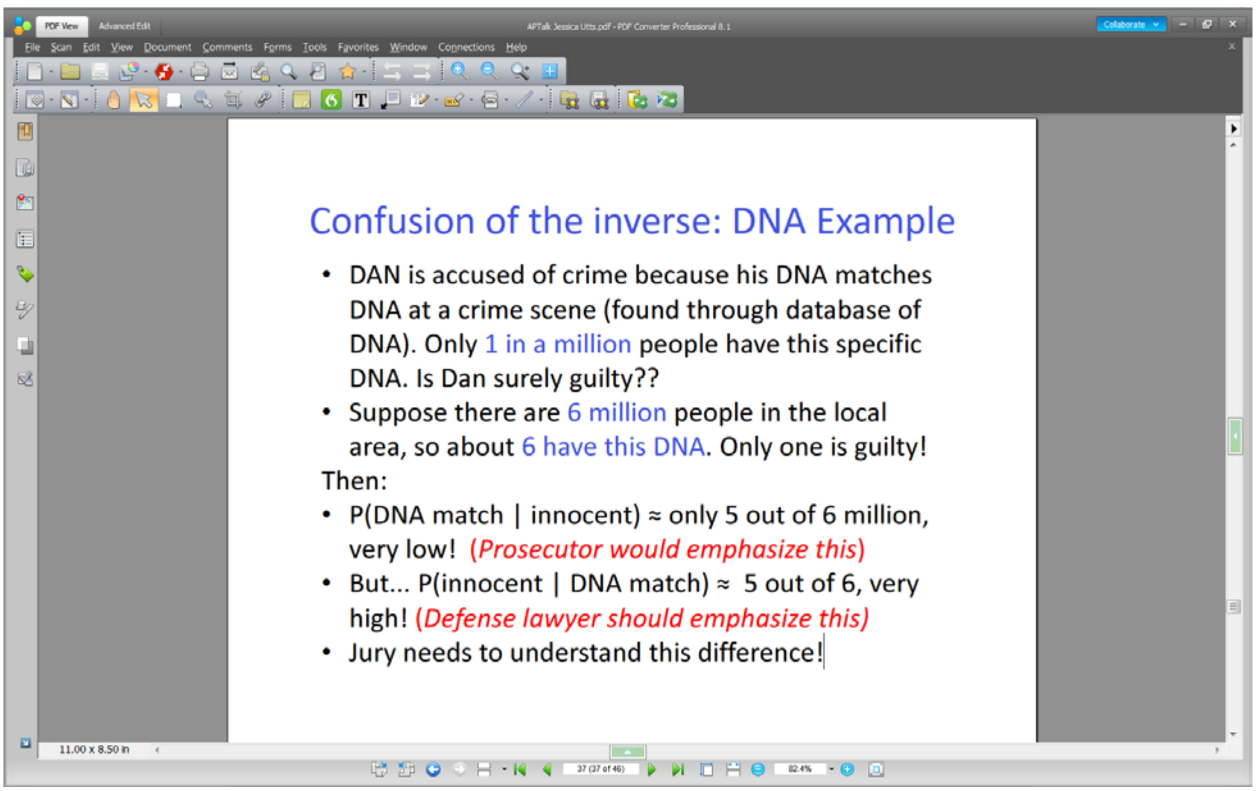
5. Bring the issue to another current day context: Enter the NFL[19]

To better understand this issue, we turn to an example from Professor Utts' presentation to illustrate a Top Ten Topic:

"7. Poor intuition about probability/expected value"

Consider the common use of DNA testing she showed in the following slide[20]. The probability of an innocent person having a specific DNA found at a crime scene is very low (only one chance in a million). So the prosecution would emphasize that the finding of DNA that matches that of a specific individual is a significant consideration that must be considered.

However, the defense should emphasize that the above probability is really not the question that needs to be considered in determining guilt. Rather, we need to consider what was actually found and where. One can then determine the probability that the DNA was not from the defendant. For the large city example discussed by Professor Utts, that Probability might easily be as high as 5 chances out of 6. Thus a finding of guilt is not conclusively established by the finding of DNA that matches a defendant's DNA. This is certainly counter to the general understanding in today's society and the view that is popularized in today's media.




<u>Activity #2</u> (Student Group Collaborative Project Activity):
After reviewing the above information, students should determine the following:

1. For the competitive environment of the student's choice, determine if there are any requirements? (If so list and describe up to 5 such requirements, if not, say that)
2. For a representative requirement, what is the test(s) to ensure compliance for that requirement (limit your analysis to a maximum of 5 tests)
3. In a manner similar to the above examples, perform a Web Quest to search for any examples of disagreement with the test results and describe your findings.
4. If a student finds no areas of issue in the field of their choice, they may obtain data from an example with medical tests. The analysis described in slides 30 to 37 of Professor Utts' presentation may be used as a guide[21]

*Activity #3:*

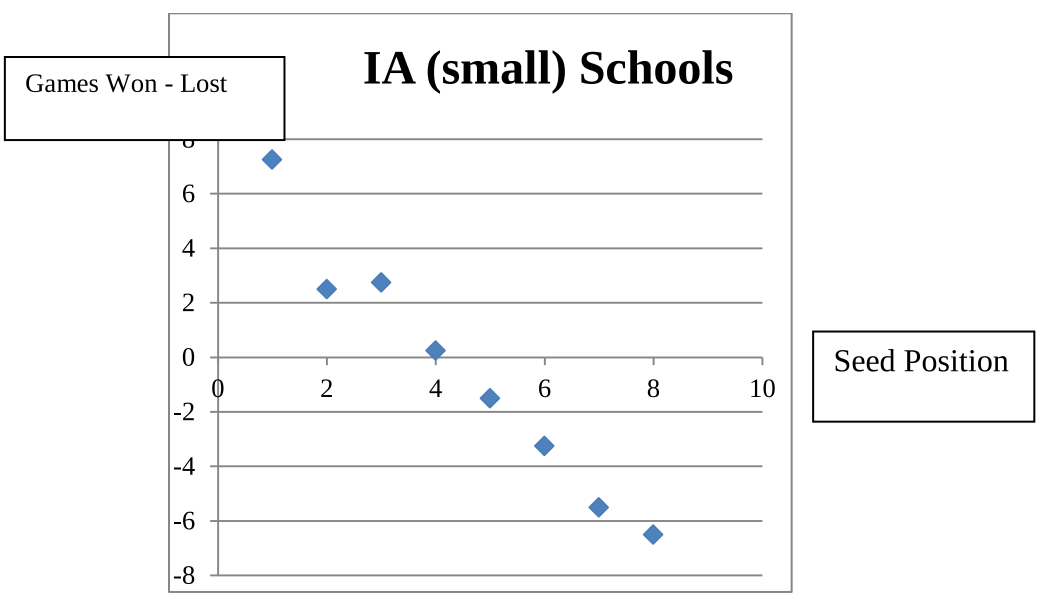
The NCHSSA[22] administers high school athletics in North Carolina. Schools are classified in terms of their student enrollment and compete in conferences consisting of geographically similar schools and similar student enrollments.[23] "The high schools in the state are organized into four classifications by the size of their student populations. These classifications are designated 1A, 2A, 3A, & 4A classifications. 4A is made up of the largest schools. The schools are split so that approximately 25% of the schools are in each classification. The classifications are reordered every four years based on updated student population numbers.

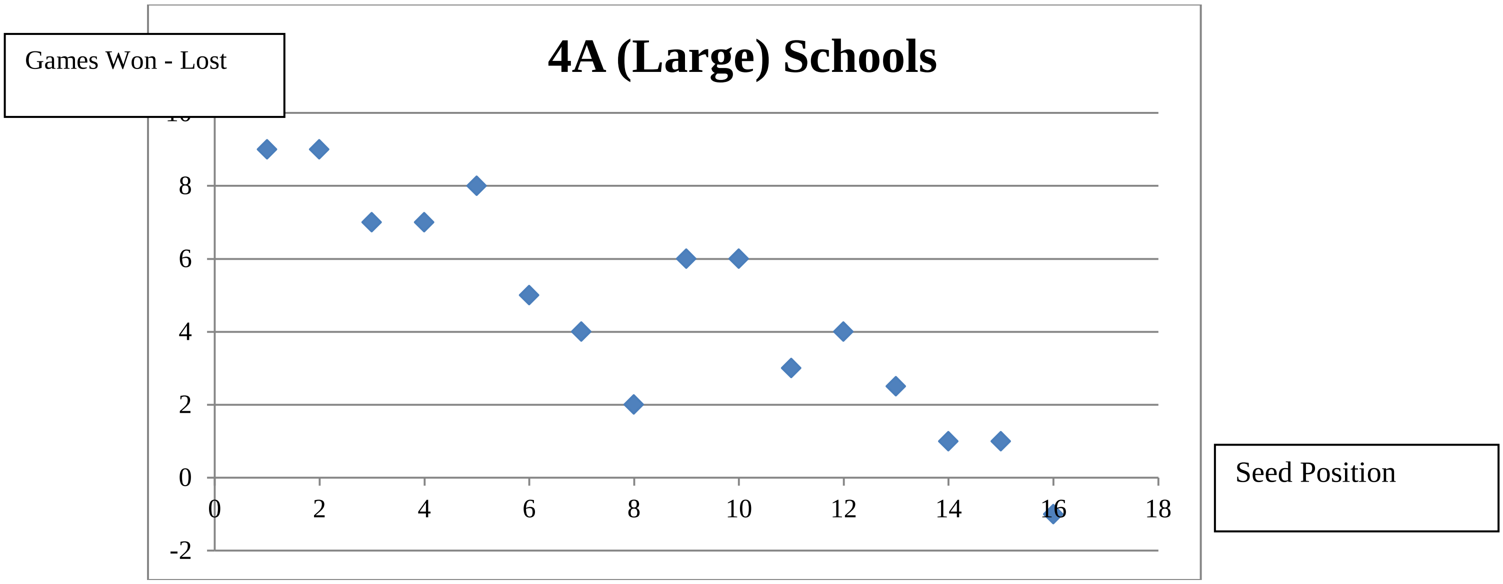
Team sports have a separate state championship competition and title for each of the classifications. The only exceptions are cases in which the 1A & 2A classifications are combined, or in football, where each classification is separated into a single "A" and double "A" (AA) classification, with the double "A" classification being made up of larger schools than the single "A".

Here is a question of interest. Perhaps schools with larger enrollments can field teams with a high level of player capabilities at most if not all positions, while smaller schools may not enjoy the same degree of consistency in player capabilities across all positions. Football teams typically have large, talented and dedicated coaching staffs who presumably can identify weaknesses in an opponent and design plays to exploit such weaknesses. Thus one may postulate that a team with strength at all positions should achieve at a higher and more consistent rate. (However, there is an outlier scenario in that a team, of any size, may have a superstar athletic who could be a deciding factor in winning games.)

To explore this we will look at the difference in Wins and Losses for different classifications among teams of similar 2013 football playoff seeds from high schools of different sizes. High Schools in the 4AA classification typically have student populations that exceed 2,000 students, while schools in the 1A classification typically have less total enrollments of less than 800 students. The following table shows a metric which is the difference between the number of wins and the number of losses for the teams holding specific seeds in the 2013 NCHSAA classifications. Note E/W stands for Eastern/Western NC teams. Also the perceived strongest teams receive the lowest seeding positions. The geographic regions results are then combined to produce metrics averaged over the entire state (marked NC)

| | WINS minus LOSSES (mostly 11 game in the Fall 2013 season) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| http://www.nchsaa.org/championships.php?mode=viewchampPDF&champID=748 | | | | | | | | |
| Seed | 4AA(E) | 4AA(W) | 4AA(NC) | 1A(E1) | 1A(E2) | 1A(W1) | 1A(W2) | 1A(NC) |
| 1 | 9 | 9 | 9 | 10 | 9 | 5 | 5 | 7.25 |
| 2 | 7 | 11 | 9 | 3 | 3 | -3 | 7 | 2.5 |
| 3 | 5 | 9 | 7 | 5 | 0 | 1 | 5 | 2.75 |
| 4 | 9 | 5 | 7 | 3 | -3 | 0 | 1 | 0.25 |
| 5 | 7 | 9 | 8 | 1 | -3 | -5 | 1 | -1.5 |
| 6 | 5 | 5 | 5 | -1 | -3 | -6 | -3 | -3.25 |
| 7 | 5 | 3 | 4 | -5 | -7 | -7 | -3 | -5.5 |
| 8 | 5 | -1 | 2 | -5 | -9 | -7 | -5 | -6.5 |
| 9 | 5 | 7 | 6 | | | | | |
| 10 | 5 | 7 | 6 | | | | | |
| 11 | 1 | 5 | 3 | | | | | |
| 12 | 5 | 3 | 4 | | | | | |
| 13 | 2 | 3 | 2.5 | | | | | |
| 14 | 1 | 1 | 1 | | | | | |
| 15 | 1 | 1 | 1 | | | | | |
| 16 | -1 | -1 | -1 | | | | | |
| Mean | 4.4 | 4.8 | 4.6 | 1.4 | -1.6 | -2.8 | 1.0 | -0.5 |






The data presented above is timely data of interest to many high school students. It is expected to be engaging for students and can be used in bivariate correlation and regression studies. Both the graphs clearly show the expected association between the strength of a teams' Won/Lost record and their seed position. Moreover, the difference between school size classifications and the mean number of the metric of wins minus looses (4.6 versus -0.5) is consistent with a higher probability of large schools being able to dominate their peer teams than can be expected for smaller schools.

Activity # 3 Assignment (Student Group Collaborative Project Activity):
After considering the above data and analysis as an indication of the richness of the data and information that can be obtained to answer questions of interest, students should identify a statistical question of interest for their choice of competitive activity. To ensure relevance, students should review their plans with their instructor before performing their analysis and write ups. The student deliverable should include:

1. A statement of the question of interest in context
2. A discussion of the analysis plan and any needed assumptions and their validity
3. A display of the analysis, typically including a graph, and the findings
4. A statement of the conclusion in context

*Activity # 4 - Making Competitive Decisions*[24]*:*

Throughout the fall of each year thousands of football games occur across the United States at all levels of play from youth leagues through professional football, which is organized as the National Football League, the highest revenue producing sport in the United States. For quick reference a synopsis of the relevant rules of football is presented

in an Appendix of a paper by Professor David Romer from University of California at Berkley dated July 2005 and titled:
"DO FIRMS MAXIMIZE? EVIDENCE FROM PROFESSIONAL FOOTBALL".[25]))[26]

Each of our lives may be viewed in the context of an accumulation of decisions. Likewise businesses and economies may be viewed in a similar manner. In his benchmark paper, Professor David Romer, a noted macroeconomist, "examines a single, narrow decision—the choice on fourth down in the National Football League between kicking and trying for a first down—as a case study of the standard view that competition in the goods, capital, and labor markets leads firms to make maximizing choices." [27] He finds "….that teams' choices on fourth downs differ from the choices that would maximize their chances of winning in ways that are systematic and overwhelmingly statistically significant. Indeed, there are cases where teams consistently make choices that represent clear-cut and large departures from win-maximization." Moreover, Romer finds "…that over the entire field, teams' choices are far more conservative than the choices that would maximize their chances of winning."

So what is the conventional wisdom? (Interestingly, it is in overwhelming use today) Basically it is to punt on 4th down unless there is a reasonable chance of making a field goal. Romer points out that such logic considers only part of the consequences of this decision. One must also consider the impact of turning the ball over to the other team, who in turn may now have a higher probability of scoring from a more advantageous starting point.

His simplest case example of this phenomena is to consider a first down and goal opportunity from the opponent's 2-yard line. In greater than 90% of such situations a touchdown is scored, 6, 7, or 8 points are added to the team's score and a kick-off to the other team is the next play.

However, when the defense successfully executes a "goal line stand" for three downs, the offense gets to 4th and goal from still the 2-yard line. At this point, frustrated by the lack of ability to score and reasoning that the defense is "fired-up". i.e. has the momentum and will continue to be successful; most coaches call for the "sure thing" and kick a field goal. It typically ends up with the addition of 3 points to the team score and a kick-off to the other team.

Ok, so what is the problem with this conventional "sure thing" strategy? Using probability analysis Romer finds that a team facing fourth-and-goal within five yards of the end zone is better off, on average, trying for a touchdown. The reason for this is that if the team fails to score, the other team is now starting from their own 2 yard line, as opposed to on average starting at their 30 +/- yard line after a kick-off. So what is the impact? Using the expected value concept of Probability Theory, teams starting close to their own goal line rarely end up scoring points whereas they often suffer a 2 point safety against them or after 4 downs turn the ball over to the other team who are now in a very favorable position to score. The sum of all of these possible outcomes times their respective probabilities is called the expected points associated with a first down and 10 yards to go at a specific point on the field. It is vividly demonstrated using a concept called Expected Points AS well described by Brian Burke of Advanced NFL Stats at http://www.advancednflstats.com/2008/08/expected-points.html including a graph at the above link titled "Expected Point Value by Field Position".

ESPN analyst Greg Garber weighed in on this discussion [28] and reported that in the 1,100 instances where teams had more to gain in Romer's estimation by going for a touchdown, they did so only 108 times, or less than 10 percent.

Ok, so Romer's conclusions are definitely not the conventional practice, so what can happen in reality when Coaches listen to Statisticians? A vivid demonstration[29] shows the approach of Coach Kevin Kelley of Pulaski Academy in Little Rock, Arkansas. He follows Professor Romer's teachings to the point of "never punting" and always trying onside kicks. So what is the long term result of Coach Kelley playing the odds? Since 2003 Pulaski Academy has won 125 games and lost 23 (84% win rate!), won three state titles, and achieved an 11-1 season in 2013.

Activity #4 Assignment:

After considering the above data and analysis as an indication of the richness of the data and information that can be obtained to answer questions of interest, students should identify a common decision situation that often occurs for their choice of competitive activity. Preliminary research should be undertaken to ensure sufficient data is available or obtainable to discuss the merits of a particular type of decision. To ensure relevance and doability, students should review their plans with their teacher before performing their analysis and write ups. In some cases students may want to serve as a team statistician for a sport at their school and gather data that would be helpful for the coach.

The student deliverable should include:

1. A statement of the question of interest in context
2. A discussion of the analysis plan and any needed assumptions and their validity
3. A display of the data gather, analysis, (typically including a graph) and the findings
4. A statement of the conclusion in context

**Bibliography for Teachers**


1. Tabor, Josh and Franklin, Chris. *Statistical Reasoning in Sports*[30]. New York: 2012 This is a very informative and stimulating source of relevant ideas as Josh and Chris, leaders in the AP Statistics program, have authored in their words "a unique and powerful way to introduce the principles of statistical reasoning, …Completely covering the Common Core Standards for Probability and Statistics…"

2. W. H. Freeman & Company Book Companion Site for *Statistical Reasoning in Sports* *http://bcs.whfreeman.com/sris/#t_730892____. Course Proposal: Statistical Reasoning in Sports* Josh and Chris have provided a tremendous resource as suggestions for how a course might be proposed utilizing the concepts of their book. In particular, they show for each chapter of the book the Learning Objectives that are addressed and their alignment with the new Common Core State Standards (CCSS) for High School Statistics and Probability. Their book covers all of these (CCSS) standards.

3. Thill, Bill. *About teaching statistics, math, and other stuff,* http://roughlynormal.com , Current Blog. In this informative source of relevant ideas Bill, an AP reader, shares his thoughts and receives comments on these topics. It appears to be worth monitoring. It links to the class blog at http://ssshw.blogspot.com/

4. Thill, Bill. *"Stats, Sports and School:" A new adventure beyond AP Statistics. http://roughlynormal.com/2012/08/30/stats-sports-and-school-a-new-adventure-beyond-ap-statistics/*, Current Blog. This also is a source of relevant ideas as Bill shares his thoughts and receives comments on his efforts through his school's initiative to develop a new course. He states, "My course proposal will be entitled "Statistics, Sports, and School." My goals: Create and answer a substantive, interesting research question related to sports or sports medicine.

5. Professor Jessica Utts AP Statistics Professional Night Talk: What Do Future Senators, Scientists, Social Workers and Sales Clerks Need to Learn from Your Statistics Class? (pdf version) In this direction setting presentation, Professor Utts clearly sets the course for AP Statistics.

6. IAC Corporation. Dictionary.com, LLC. http:/dictionary.reference.com, Current Online resource, © 2013. In Statistics we seek precision in articulating imprecision in life. Thus, researching the precise meaning of words will be a critical part of our work. This is a critically important requirement in our communication; e.g. this is a major factor in the AP Statistics Exam where per the

rubric in a part of the response using the word "consumed" was graded a zero whereas the same sentence using the word "consumes" earned full credit!

*7.* http://www.ted.com/ This is a repository of very insightful and thought provoking talks by noted authorities

*8.* http://www.ted.com/talks/hans_rosling_reveals_new_insights_on_poverty.html In this talk Swedish Professor Hans Rosling uses his advanced data display technologies in a meaningful way to show the power of data analysis to bring understanding and insight to some of the most challenging problems in our world. It is an outstanding example of the impact of data analysis

9. ibidem. 4. The difference between winning and succeeding – John Wooden which inspires one to read: Wooden, John They Call Me Coach. http://www.ted.com/talks/john_wooden_on_the_difference_between_winning_and_success.html

**Reading List for Students**

*Rubrics*

1. *http://www.amstat.org/education/posterprojects/pdfs/ProjectJudgingRubric.pdf* This table provides clear areas of focus and requirements for projects.

*Activity # 1*

2. Tabor, Josh and Franklin, Chris. Statistical Reasoning in Sports[31]. New York: 2012 This is a very informative and stimulating source of relevant ideas regarding the relationship between sports and mathematics. It provides "a unique and powerful way to introduce the principles of statistical reasoning, …"

3. Professor Jessica Utts AP Statistics Professional Night Talk: What Do Future Senators, Scientists, Social Workers and Sales Clerks Need to Learn from Your Statistics Class? (pdf version) In this direction setting presentation, Professor Utts clearly sets the course for AP Statistics.

4. IAC Corporation. Dictionary.com, LLC. http:/dictionary.reference.com, Current Online resource, © 2013. In Statistics we seek precision in articulating imprecision in life. Thus, researching the precise meaning of words will be a critical part of our work. This is a critically important requirement in our communication; e.g. this is a major factor in the AP Statistics Exam where per the rubric in a part of the response using the word "consumed" was graded a zero

whereas the same sentence using the word "consumes" earned full credit!

5. Rubrics: http://www.amstat.org/education/posterprojects/pdfs/ProjectJudgingRubric.pdf

6. This table provides clear areas of focus and requirements for projects.

7. http://www.ted.com/ This is a repository of very insightful and thought provoking talks by noted authorities

8. http://www.ted.com/talks/hans_rosling_reveals_new_insights_on_poverty.html In this talk Swedish Professor Hans Rosling uses his advanced data display technologies in a meaningful way to show the power of data analysis to bring understanding and insight to some of the most challenging problems in our world. It is an outstanding example of the impact of data analysis

9. The difference between winning and succeeding – John Wooden which inspires one to read: Wooden, John They Call Me Coach. http://www.ted.com/talks/john_wooden_on_the_difference_between_winning_and_success.html

*Activity # 2*

10. http://www.wada-ama.org/en/ This website is the official site for the worldwide movement to combat use of performance enhancing drugs.

11. http://www.youtube.com/watch?v=57SdyA6VKf8 Highlights of the 2006 Winter Olympics event in

12. http://news.err.ee/sports/7b580679-cce8-4995-964e-7c99f4ad09a5 , http://news.err.ee/sports/c387d8a0-39a0-4621-96dd-db8a4d29db37 , and http://sports.espn.go.com/oly/crosscountry/news/story?id=6304402 these are news stories related to this drug use controversy.

*Activity # 3*

13. http://en.wikipedia.org/wiki/North_Carolina_High_School_Athletic_Association#3A.E2.80.934A This site provides a description of the NC High School Sports Classifications

14. http://www.nchsaa.org/championships.php?mode=selectchampsport#.UpttK8RDut8 this site provides links to the high school sports playoff brackets

*Activity #4:*

15. The Coach Who Never Punts http://www.youtube.com/watch?v=AGDaOJAYHfo
This video highlights a coach who uses statistical methods in coaching. His teams have a 84% Winning percentage for the last 10 years.

16. Do Firms Maximize? Evidence From Professional Football http://elsa.berkeley.edu/~dromer/papers/PAPER_NFL_JULY05_FORWEB_CORRECTED.pdf In this often cited benchmark work, Professor Romer provides a mathematical framework that determine in his view that in general the majority of football coaches are too conservative in their play calling and are missing opportunities.

17. http://static.espn.go.com/nfl/columns/garber_greg/1453717.html and http://www.huffingtonpost.com/2012/02/04/football-statistics-show-teams_n_1254352.html These analyses share well know coaches' their view of Professor Romer (#2 above) work

18. http://www.advancednflstats.com/2008/08/expected-points.html This report shares the very interesting NFL statistic of expected points given a starting yardage for a new set of downs

19. http://www.advancednflstats.com/2008/07/first-down-probability.html This reports shares the NFL probability of making a first down given the down and the yards to go.

**Materials For Classroom Use**

1. Computer Access: Students must be able to perform Web Quests. To ensure all students have access to the Internet, the class may go to a computer laboratory or bring a laptop or IPAD cart into the classroom

2. Statistical Data Analysis: Students must be able to analyze data in a statistically appropriate manner. A large amount of the data for examples illustrating the above standards may be found on the Internet. Many teachers are adopting technology based tools and using them as a primary platform for gathering, presenting, and analyzing data. At the best practices professional development session of the 2013 AP Statistics Exam Reading I learned about a free internet and mobile phone app that has tremendous functionality for data gathering and analysis. It is called STATKEY http://lock5stat.com/statkey/sampling_1_quant/sampling_1_quant.html At that site one can choose several sporting databases, such as Olympic Marathon times in minutes (Upper left link) and immediately visualize key aspects of the distribution. Statkey can serve as an effective platform for students to explore and gain an appreciation of the tools for data analysis.

**Appendix "Implementing Common Core Standards"**

North Carolina has adopted the Common Core Standards for Mathematics[32] The Probability and Statistics Domain of these standards consist or four clusters:

- Interpreting Categorical and Quantitative Data (S-ID)
  The ability to interpret data in a statistical manner is a fundamental 21st century skill. All three of this unit's classroom activities utilize this skill and provide rigorous applications of this skill
- Making Inferences and Justifying Conclusions (S-IC)
  The ability to make inferences and justifying conclusions in a sound statistical manner is a fundamental 21st century skill. The activities that include Case Studies on drug use testing and 4th down decisions are current, state-of-the-art applications of these techniques
- Conditional Probability and the Rules of Probability (S-CP)
- Using Probability to Make Decisions (S-MD)
  Understanding probabilistic analysis (S-CP) and using the resulting data to make decisions (S-MD) are fundamental 21st century skills. The activities that include Case Studies on drug use testing and 4th down decisions are current, state-of-the-art applications of these techniques

## Endnotes

[1] "Just Do It" is a Nike trademark. I discussed its use and guidelines as part of this title with the Nike World Office. I learned that it is up to an author in situations like mine to determine whether a proposed usage is acceptable. We discussed their intent and I have determined that this usage is acceptable.

[2] http://en.wikipedia.org/wiki/Bill_James#The_Bill_James_Baseball_Abstracts Many recognize Bill James as the "father" of what is now known as "Sports Analytics".

[3] *Do Firms Maximize? Evidence From Professional Football* http://elsa.berkeley.edu/~dromer/papers/PAPER_NFL_JULY05_FORWEB_CORRECTED.pdf In this often cited benchmark work, Professor Romer provides a mathematical framework that determine in his view that in general the majority of football coaches are too conservative in their play calling and are missing opportunities.

[4] *Tabor, Josh and Franklin, Chris. Statistical Reasoning in Sports . New York: 2012* This is a very informative and stimulating source of relevant ideas as Josh and Chris, leaders in the AP Statistics program, have authored in their words "a unique and powerful way to introduce the principles of statistical reasoning, …Completely covering the Common Core Standards for Probability and Statistics…"

[5] AP Statistics standards are described at page 8 in the College Board's Course description http://apcentral.collegeboard.com/apc/public/repository/ap-statistics-course-description.pdf

[6] The author is indebted to our seminar leader, Professor Chartier for pointing out the relevance of this topic of investigation to this unit.

[7] The author is indebted to Jason W. Rosenfeld, Director of Basketball Analytics, Charlotte Bobcats, and a speaker at both our CTI math seminar and the CTI 2013 Exploding Cannons for suggesting this reference and in response to our discussions on this unit.

[8] Professor Jessica Utts AP Statistics Professional Night Talk: What Do Future Senators, Scientists, Social Workers and Sales Clerks Need to Learn from Your Statistics Class? (pdf version)

[9] http://dictionary.reference.com/browse/athletic?s=t

[10] ibidem

[11] ibidem

[12] John Brenkus ESPN Sports Analyst pointed out in our seminar that in such sports when play begins or resumes in these sports, it is chaotic. That is why it is important to note that for this work we are only interested in the outcome of each instance of play, such as a down in football. There is a large amount of data available on the outcomes, which may be viewed as a random phenomena

[13] http://www.ted.com/talks/john_wooden_on_the_difference_between_winning_and_success.html

[14] http://www.amstat.org/education/posterprojects/pdfs/ProjectJudgingRubric.pdf

[15] http://www.youtube.com/watch?v=57SdyA6VKf8

[16] http://sports.espn.go.com/oly/crosscountry/news/story?id=6304402

[17] http://news.err.ee/sports/7b580679-cce8-4995-964e-7c99f4ad09a5

[18] http://www.wada-ama.org/en/

[19] http://news.err.ee/sports/c387d8a0-39a0-4621-96dd-db8a4d29db37 The NFL is reported to be carefully monitoring this challenge to fundamental World Anti-Doping Agency procedures from the aspect of potential spill over to NFL procedures.

[20] Professor Jessica Utts slides 37, AP Statistics Professional Night Talk: What Do Future Senators, Scientists, Social Workers and Sales Clerks Need to Learn from Your Statistics Class? (pdf version)

[21]Professor Jessica Utts slides 30 - 37, AP Statistics Professional Night Talk: What Do Future Senators, Scientists, Social Workers and Sales Clerks Need to Learn from Your Statistics Class? (pdf version)

[22] http://www.nchsaa.org

[23]

http://en.wikipedia.org/wiki/North_Carolina_High_School_Athletic_Association#3A.E2.80.934A

[24] This activity addresses Professor Utts' point #7: "Poor intuition about probability/expected value".

[25]

http://elsa.berkeley.edu/~dromer/papers/PAPER_NFL_JULY05_FORWEB_CORRECTED.pdf

[26] The author is indebted to Jason W. Rosenfeld, Director of Basketball Analytics, Charlotte Bobcats, and a speaker at both our CTI math seminar and the CTI 2013 Exploding Cannons for suggesting this reference in response to our discussions on this unit.

[27] Quotation from the Abstract of Professor Romer's paper explaining his view that highly competitive sports decisions such as 4th down play decisions may serve as "laboratories" in which to study the highly competitive decision making process in industry. (3. ibidem)

[28] http://static.espn.go.com/nfl/columns/garber_greg/1453717.html

[29] http://www.youtube.com/watch?v=AGDaOJAYHfo The author is indebted to Jason W. Rosenfeld, Director of Basketball Analytics, Charlotte Bobcats, and a speaker at both our CTI math seminar and the CTI 2013 Exploding Cannons for suggesting this reference in response to our discussions on this unit.

[30] http://highschool.bfwpub.com/catalog/Product/statisticalreasoninginsports-firstedition-tabor

[31] http://highschool.bfwpub.com/catalog/Product/statisticalreasoninginsports-firstedition-tabor

[32] http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf