YALE NATIONAL INITIATIVE
to strengthen teaching in public schools®

Curriculum Units by Fellows of the National Initiative
2015 Volume VI: Physiological Determinants of Global Health

Statistical Methods and Health in Chicago

Curriculum Unit 15.06.06, published September 2015
by Sarah J. Schneider

Content Objectives

Introduction and Rationale

Imagine asking your students the following sequence of questions in class: What does it mean to be healthy? What does it mean to be unhealthy? How much of your health is under your own control? How does your family affect your health? How does your environment affect your health? How does your socioeconomic status affect your health? What are the major health problems you may face in your lifetime? How can we predict and prevent these future health problems?

I imagine the answers in your classroom would be as varied as those I would find in my own; some would be very encouraging, some very disheartening, and a select few would be very comical. As educators of any discipline, we find it imperative that we further our students' understanding of and control over their own health through probing questions such as these and the subsequent lessons we teach to address these issues

As a mathematics educator, I also find it imperative that I inspire in my students an appreciation for mathematics and the power it gives us in understanding and communicating about the world around us. Too often students see math as an intimidating set of rules and processes that have little to no relevance in their daily lives. Even when you encounter the few individuals who are self-proclaimed "math people," they are not always able to see the connections math has to the world around them, let alone the impact a thorough understanding of mathematical processes can have on their own quality of life.

Through this unit, I intend to provide enriching opportunities for students to practice and apply core statistical methods while exploring health in their hometown of Chicago. It is my hope that this unit will allow students to see mathematics as an accessible, integral component to not only their academic studies, but also to understanding and communicating information in their daily lives. I hope that this unit will serve as a springboard for discussion as to how students can take better care of their own health and also be advocates for better health in their families and communities.

School Profile

Back of the Yards College Preparatory High School is located on the southwest side of Chicago and is entering its third year of serving both the Back of the Yards neighborhood students as well as students from

surrounding neighborhoods through our International Baccalaureate (IB) Program. All students engage in the Middle Years Program in 9th and 10th grade, a program that aims to push students to become more internationally minded, well-rounded students through specific interdisciplinary connections and rigorous curricular components. Students then have the opportunity to apply to the Diploma Program (DP), Career Program, or Subject Certificate courses for their 11th and 12th grade studies as part of the IB Program. Each of these programs works to continue to build IB learners who are inquirers, knowledgeable, thinkers, communicators, principled, open-minded, caring, risk-takers, balanced, and reflective[1] . Our student body is approximately 85% Hispanic, 7% Asian, 5% African American, and 3% White. Almost all students (99%) come from a low-income family. Approximately 11% of our students have disabilities and 24% of our students are English Language Learners[10] . There are definite health struggles within these specific communities that we will examine through this unit.

Course Profile

This curriculum unit is being developed as the first unit of the year for my two sections of 11th Grade DP Mathematics. The prerequisites for this course are high school Algebra I, Geometry, and Advanced Algebra with Trigonometry. All of the students in these two sections will be concurrently enrolled in 11th Grade DP Biology. I will have numerous opportunities to collaborate with the 11th Grade DP Biology instructor, and thus we will be able to build many interdisciplinary connections throughout this unit and subsequent units. The statistical methods we focus on in this unit will serve as a foundation both for this mathematics course as we build into more complex modeling scenarios, and for laboratory work in the 11th Grade DP Biology course. In addition, the 11th Grade Biology topic sequence will expose my math students to the biological underpinnings of health and disease in order to give them a deeper understanding of the science behind this initial unit. While this unit is developed for a rigorous, accelerated 11th grade mathematics course, components of this unit can be used at all levels of high school mathematics study, as well as in secondary science studies, particularly biology. The timeline for this unit as it stands is approximately five weeks of five 50-minute class periods (or roughly 20 instructional hours).

Background Information

Statistical Methods

Students will gain proficiency in mathematically summarizing data sets, and they will use these mathematical summaries to make logical predictions and informed decisions about health-related issues in their hometown. The statistical methods we will study in this unit will fall into one of three categories: measures of central tendency and dispersion, measuring frequency, and linear correlation[2]. The section below aims to provide an overview of the statistical methods necessary; for the scope of this class, students will not be expected to calculate certain measures by hand (i.e. standard deviation) but instead they will use technology. Thus many formulas have been omitted, though they will be discussed in class.

Measures of central tendency and dispersion are integral in understanding the average and the distance from the average within a data set. There are three main measures of central tendency: the mean (the sum of all data values divided by the number of data values), the median (the middle number of a data set when values are arranged in ascending order), and the mode (value occurring most frequently in a data set). Similarly, there are three measures of the spread of a data set: the range (the distance between maximum and minimum values of the data set), the interquartile range (the distance between the lower quartile and the

upper quartile), and the standard deviation (the average distance of data values from the mean – we will primarily use technology to calculate this measure). It is important to note that outliers are found by identifying data points that are 1.5 interquartile ranges (IQRs) away from the median. There are distinct advantages and disadvantages to each measure in summarizing a data set, as detailed in Table 1 below.

Table 1: Advantages and Disadvantages of Measures of Central Tendency and Dispersion

Measure	Advantage	Disadvantage
Mean **Range** **Standard Deviation**	Gives "big picture" of the data set by considering all data values	Can be inaccurately distorted as a result of outlier values
Median **Interquartile Range**	Unaffected by outlier values	Can omit some of the "big picture" of the data set by not representing all data values
Mode	Useful for qualitative data sets	Can have multiple values

Particularly when dealing with large data sets, as we often are in the realm of health, we create tables and diagrams to measure the frequency of occurrence of data values instead of listing all data values in a set. Frequency tables are useful in identifying the number of occurrences of each exact data value. The mean and standard deviation of a data set can also be calculated directly from a frequency table. In addition, we use frequency tables to group data into appropriate intervals. While we may lose some level of detail in such tables, they are incredibly useful in summarizing overall patterns and trends in data. We use the mid-interval value (the mean of the upper and lower interval boundaries) to calculate mean and standard deviation in these cases. In order to calculate median and quartiles of a frequency table, we construct cumulative frequency diagrams; cumulative frequency represents the running total of all frequencies up to a certain data value. Cumulative frequency diagrams are essentially graphs of the data set that detail the data values along the *x*-axis and cumulative frequencies along the *y*-axis. We can graphically estimate medians and quartiles from these curves. Box-and-whisker plots are also a useful diagram in representing the center and spread of a data set, and they are helpful in visualizing outliers. The final diagram we will use in representing frequency is a histogram. Histograms communicate the distribution of the data set and are similar to bar graphs. Histograms, however, are used uniquely for sets of continuous data, and the size of each bar must be representative of the frequency of each data value (bar graphs, on the other hand, are used for discrete or qualitative data sets and the size of each bar does not necessarily convey the frequency of the data value). Table 2 below provides examples of each of these frequency diagrams constructed from TI-84 Plus Emulator software (similar to what will be utilized in class) using my own sample set of data.

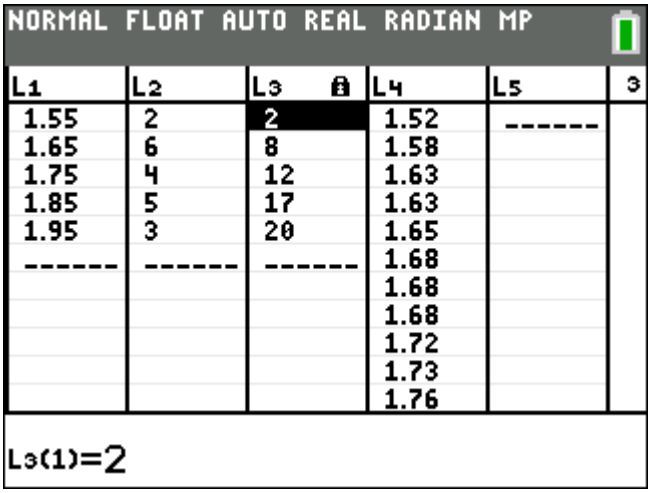
Table 2: Examples of Frequency Diagrams

Initial Data Set representing heights in meters of 20 students:

{1.52, 1.58, 1.63, 1.63, 1.65, 1.68, 1.68, 1.68, 1.72, 1.73,

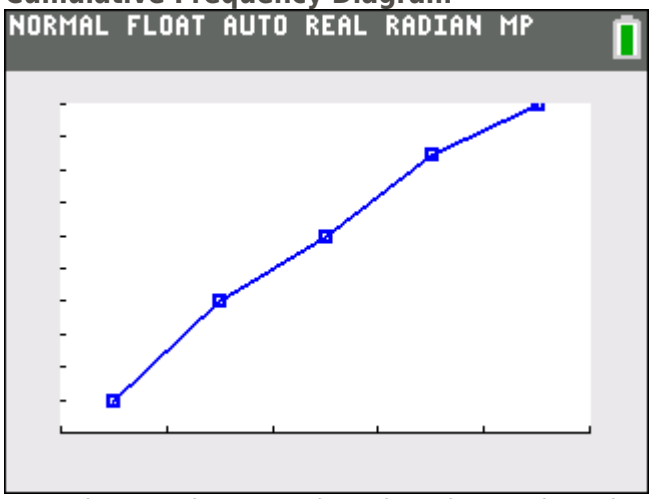
1.76, 1.76, 1.80, 1.81, 1.82, 1.85, 1.90, 1.93, 1.97, 1.99}

Frequency and Cumulative Frequency Table with 5 intervals of width 0.1



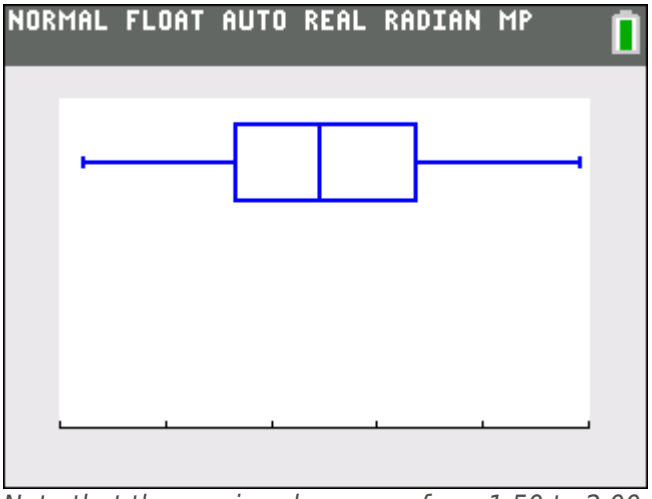
Note that the mid-interval value is used for each interval of width 0.1 in column L_1, the frequency of data values within each interval is listed in L_2, and the cumulative frequency for each interval is listed in L_3. L_4 is the full data set (continued beyond the screen).

Cumulative Frequency Diagram



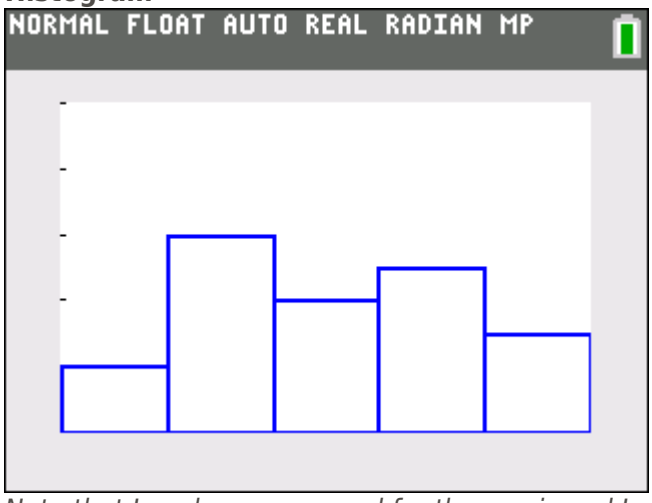
Note that L_1 values are plotted on the x-axis and L_3 values are plotted on the y-axis.

Box-and-Whiskers Plot



Note that the x-axis values span from 1.50 to 2.00 (the same scale that was used in the cumulative frequency diagram) and L_4 was utilized to create this plot.

Histogram

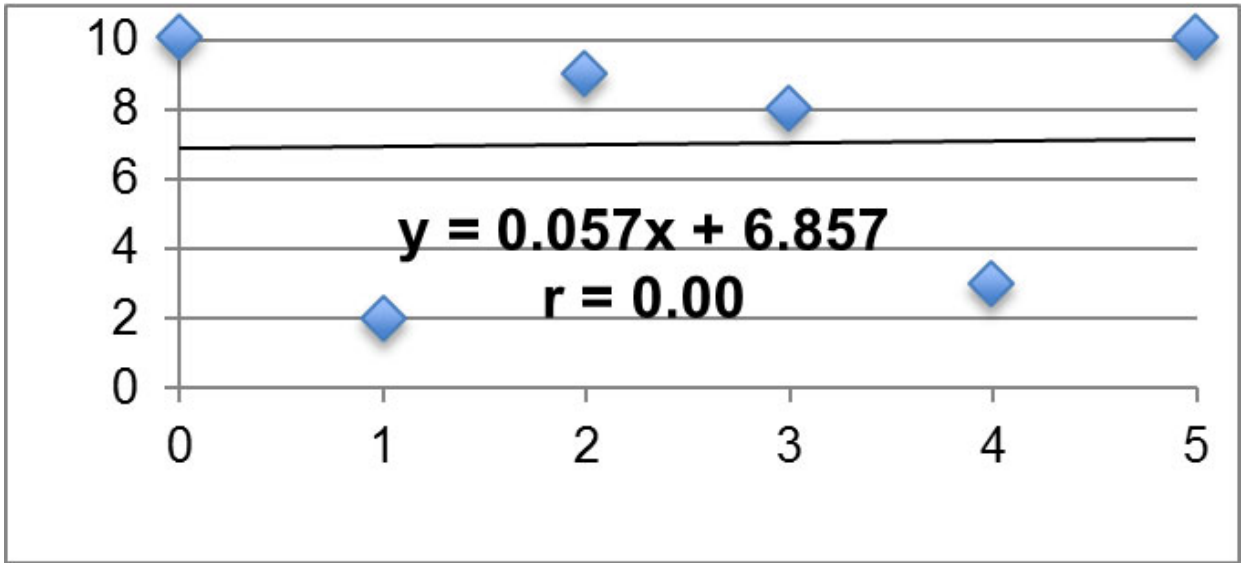
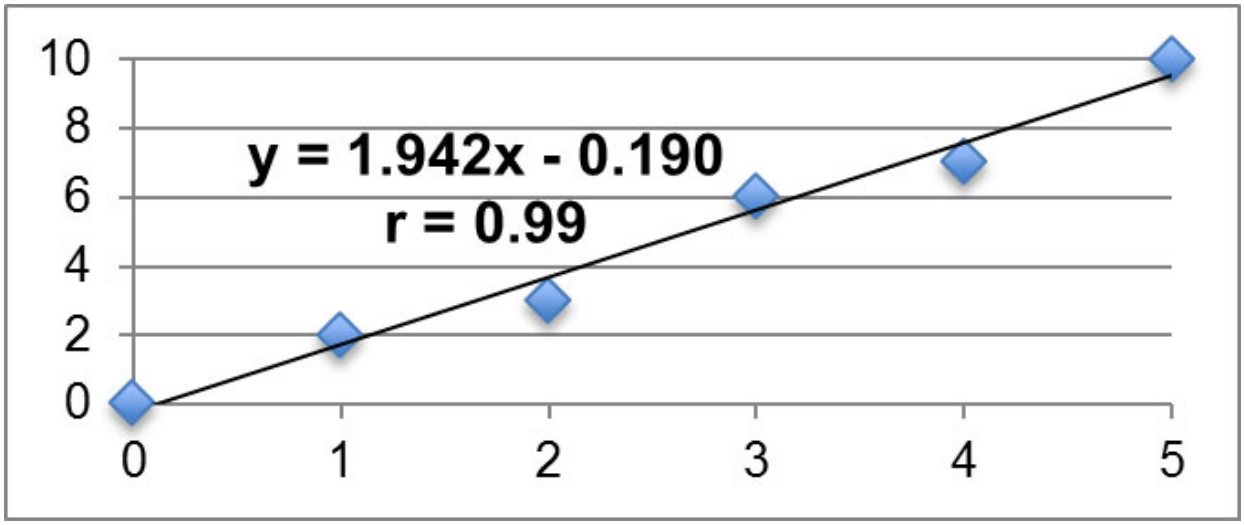


Note that L_1 values were used for the x-axis and L_2 values were used for the y-axis (or the height of each bar).

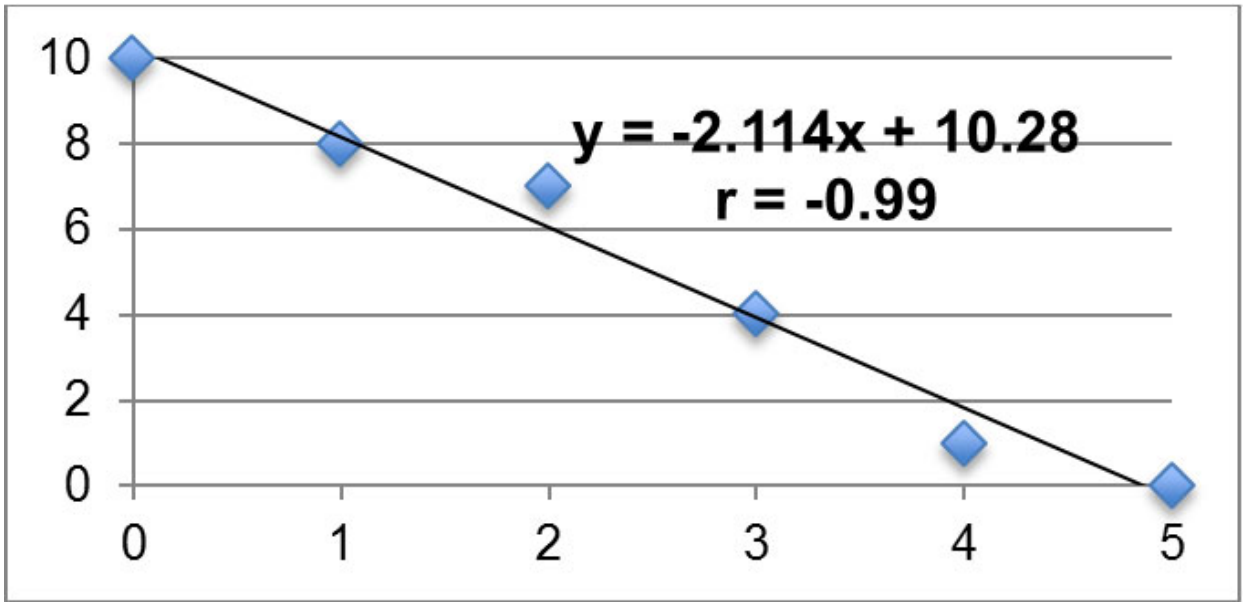
Often times we want to statistically analyze two data sets at a time and determine if there is a relationship between them. We call such data sets bivariate (as opposed to univariate). The relationship we are looking for between bivariate data sets is referred to as the correlation of the data sets. In this unit, we will look for linear correlation, or when the two data sets are related by a constant rate of change. We will look for linear correlation in data sets by creating scatter plots and calculating the product-moment correlation coefficient, or *r*. If *r* is equal to zero, we say there is no linear correlation between the two sets. The closer *r* is to positive or negative one, the stronger the linear correlation (positive or negative, respectively). It is imperative to remember that a correlation in data sets does not imply a causation, however, as the correlation could be a result of sheer coincidence or possible lurking variables. A common example I might give for this in my class is that when ice cream sales go up in the neighborhood, so does violence in the neighborhood. This does not necessarily imply that higher ice cream sales are the cause of more violence. Instead we might claim that we

can establish a correlation between the two, but we would want to investigate other variables that might cause the increase in violence (my students typically respond that we should look into temperature). We utilize linear regression models (using a method called least squares regression) to fit data sets that have a strong linear correlation and algebraically explain the relationship between the independent and dependent variables. We will primarily utilize technology such as Microsoft Excel or Texas Instrument Graphing Calculators to find the equation for the linear model as well as the *r*-value. These models are useful in explaining observed patterns (interpolation) in the data set as well as making predictions outside of the observed data set (extrapolation). Table 3 below summarizes key examples of linear correlation, with the linear model relating the dependent variable *y* to the independent variable *x*; the *r*-value is also displayed on each. Each example was created on Microsoft Excel using my own data sets.

Table 3: Summary of Key Examples of Linear Correlation

No Linear Correlation	
Strong Positive Linear Correlation	

Strong Negative Linear Correlation



Health in Chicago

When studying Chicago, we break the 237 square miles of the city into 77 community areas, some of which are single neighborhoods and some of which are multiple neighborhoods grouped together (there are over 100 neighborhoods total). Each of these community areas consists of approximately 35,000 of the city's total 2.7 million residents[3] . The neighborhoods of Chicago reflect the diversity within the city, but also the segregation of the city by ethnicity, race, socioeconomic class, and in many ways, health. Much of this unit will focus on Community Area 61, or New City – an area consisting of the Canaryville neighborhood and the Back of the Yards neighborhood (where my school is located and many of my students reside).

In August 2011, the Chicago Department of Health rolled out "Healthy Chicago," a multi-faceted plan to address health problems in the city. Utilizing available health statistics, this initiative set out to inform and recruit public, private, and community-based organizations to address pertinent health issues. Twelve main focus areas were decided upon: tobacco use, obesity prevention, heart disease and stroke, HIV prevention, adolescent health, cancer disparities, access to care, healthy mothers and babies, communicable diseases control and prevention, healthy homes, violence prevention, and public health infrastructure. With the ultimate goal of making Chicago the "healthiest city in the nation," this program has worked over the past four years to remedy issues within each of the twelve key priorities, releasing progress updates and annual reports regularly[4] . As of the 2013 annual report, 92% of the work outlined in the Healthy Chicago initiative was in progress and/or had been completed. The success of the program at that point was attributed to the strong partnerships in the community (public and private), the variety of priorities that have been addressed in the policies, the positive role of social media and technology in informing the public, and the strength of other public awareness tools[5] . The June 2015 Implementation update provides examples of recent Healthy Chicago work, including the "Step Up. Get Tested" campaign's goal to test 5,000 individuals at risk for HIV across the city during the month of June (addressing the HIV Prevention key priority). Under the Obesity Prevention key priority, it was noted that Chicago Public schools recently celebrated "Fresh Attitude Week" in which fresh fruits and vegetables, along with various vegetarian options, were provided in school lunches. Local farmers made visits to some schools in the city as well to increase student understanding of the local food system. Finally, this update detailed the work on "Healthy Chicago 2.0," the update to the Healthy Chicago initiative with a rollout scheduled for 2016 [6] .

Beginning Analysis of New City Community Area

The Healthy Chicago program has helped to address and improve upon some of the health issues in community areas such as New City, yet there are still significant issues to be studied and addressed. First and foremost, it is important to understand the demographics of this population[7] . The New City area population has fallen 17% from 51,721 to 44,377 between 2000 and 2010. Almost half of the current residents are under the age of 25. Roughly one-third of the residents are living below poverty level, and 12.2% live in crowded housing situations. Approximately 57% of the residents are Hispanic, 30% are Black, 11% are White, and 2% are Asian[9] . This community area has a dependency rate of 42.0%, and 42.4% of the population has no high school diploma. The per capita income is $12,524 (well below the city average of roughly $28,000) and unemployment is 17.4%. When ranked and compared to other community areas in the city by these demographics, New City is definitively one of the most at-risk community areas. In fact, this community area has the 7th highest hardship index (91) in Chicago, a measure used to summarize the demographics listed above[7] .

Moving to a discussion of the health of this population, upon examining death rates, New City has rates that are especially high compared to the city average in assault (homicide), liver disease and cirrhosis, firearm-related, injury (unintentional), lung cancer, and diabetes-related. Overall, there is a higher death rate in males in New City, and the leading causes of death are cancer (all sites, but specifically lung cancer) and coronary heart disease. Shifting the focus to the prevalence of infectious diseases in New City, the number of cases reported for both chlamydia and gonorrhea in females per 100,000 is significantly higher than the rate reported for the city of Chicago as a whole[7] . It is important to note that there is a wide range of health-related data available for all Chicago community areas, particularly New City, on the City of Chicago's official website[8] . Much of the data above will be utilized in instructional activities to support students in gaining comfort with measures of central tendency and dispersion, as well as measuring frequency. Furthermore, this data will serve to prompt students to identify areas of interest in their community that they will study more in depth with their teams.

If we examine the rate of lung cancer in the New City area further, there are many possible factors that might influence this high rate. Figure 1 below represents the prevalence of lung cancer in all 77 of the Chicago community areas, with the New City community area labeled with one of the highest rates[7] .

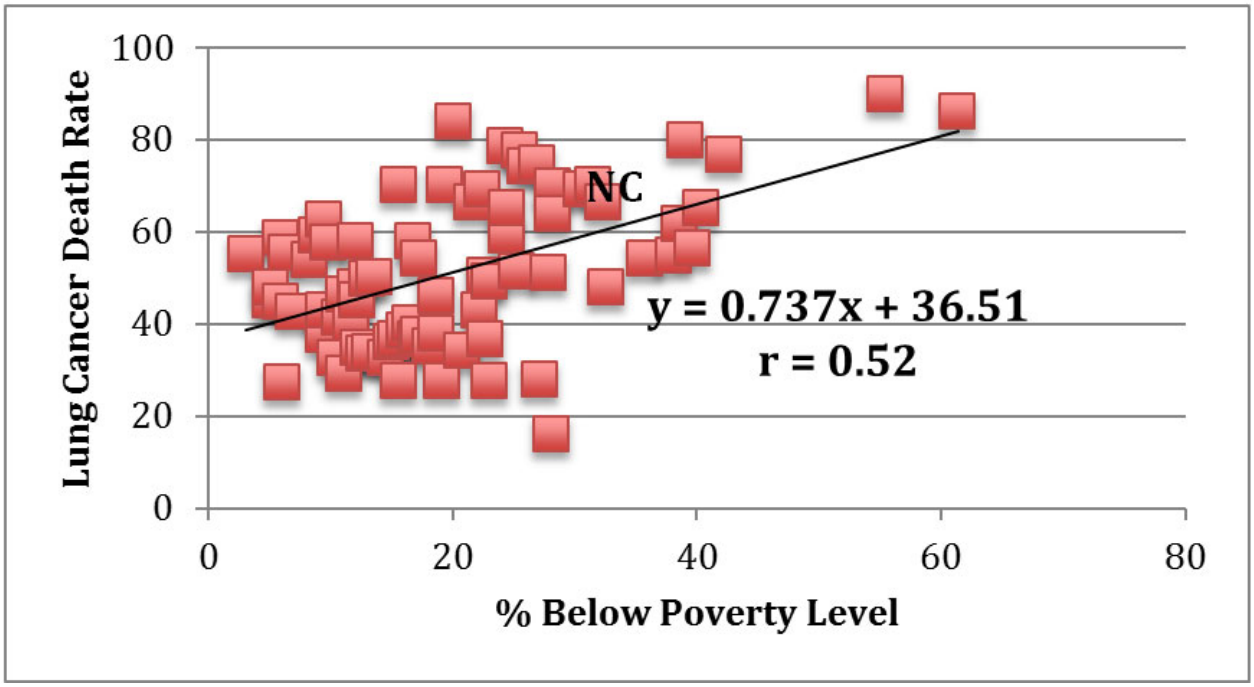
Figure 1: Adjusted Lung Cancer Death Rates per 100,000 Residents in 77 Chicago Community Areas (2006 – 2010)

(image 8)

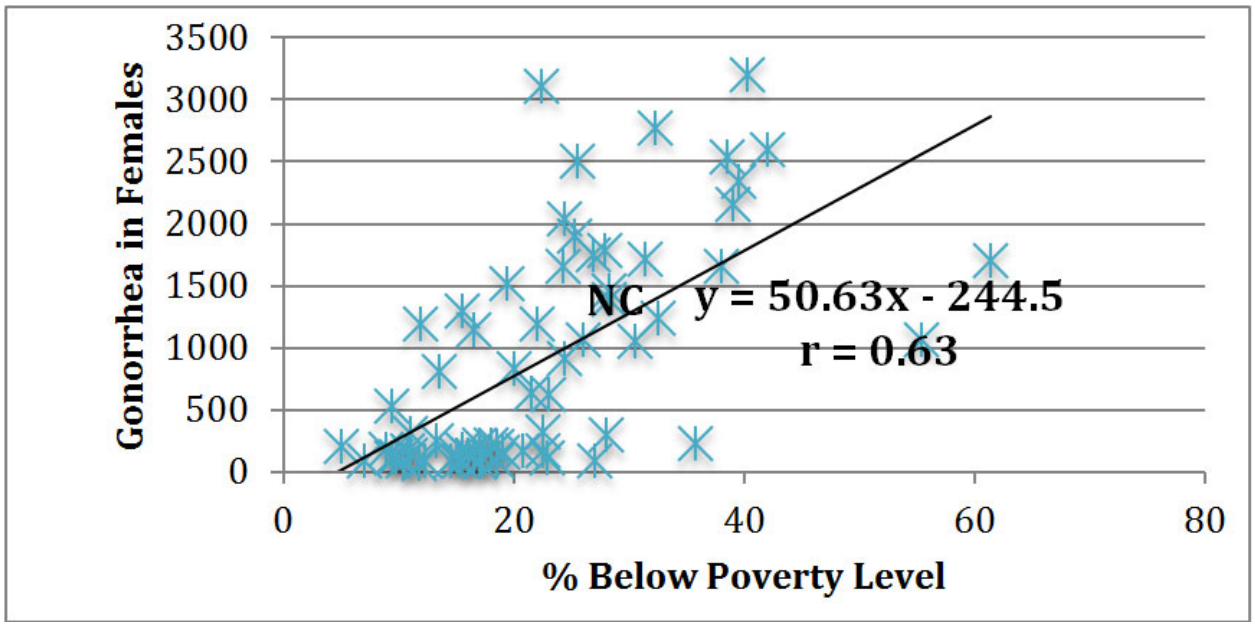
When comparing this map with the map of percent below poverty level for each community area, I noticed similarities in shading patterns. Thus, I completed a linear regression analysis for a scatter plot of lung cancer deaths as compared to percent below poverty level for the 77 community areas in Chicago. Table 4 below illustrates this analysis, as well as additional analyses looking at gonorrhea cases in females and childhood blood lead level screenings as compared to percent below poverty level due to seemingly similar patterns upon investigating the prevalence of each. The approximate data point representing New City is marked "NC" in each graph. These analyses were performed by Excel utilizing data from the City of Chicago Data Portal[8] .

Table 4: Linear Regression Analyses for Lung Cancer Death Deaths per 100,000 vs. Percent Below Poverty Level (1), Gonorrhea Cases in Females per 100,000 vs. Percent Below Poverty Level (2), and Childhood Blood

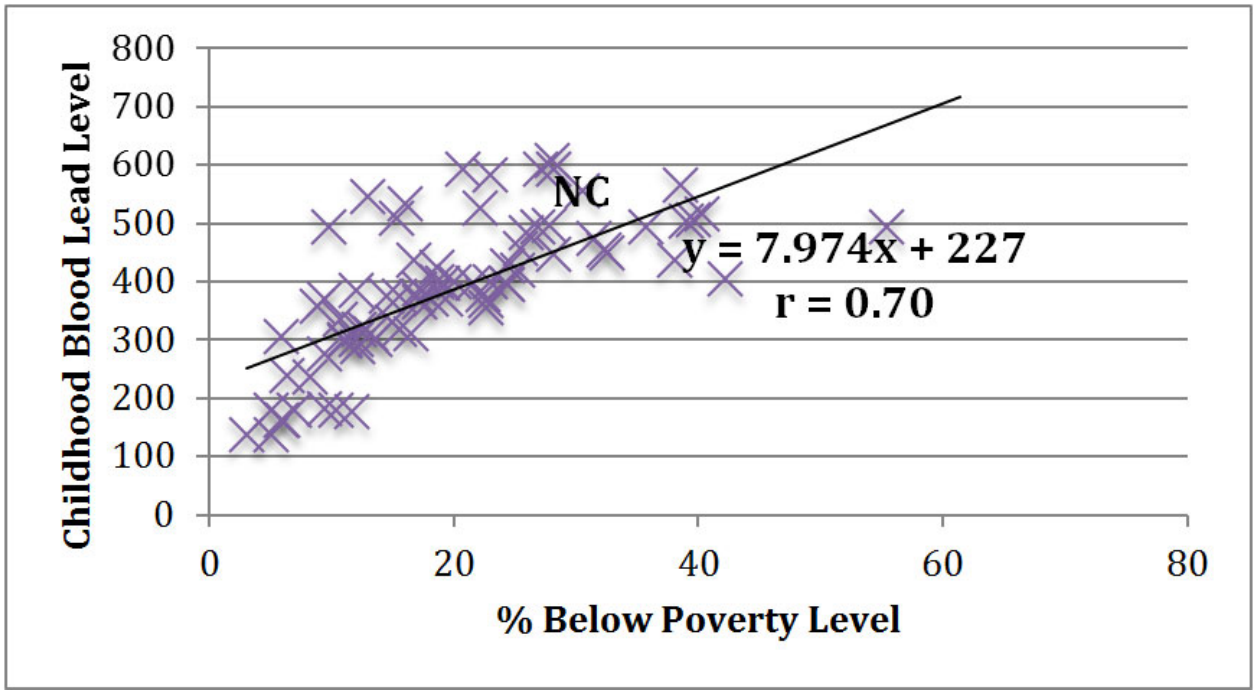
1. Lung Cancer Deaths per 100,000 vs. Percent Below Poverty Level



2. Gonorrhea in Females per 100,000 vs. Percent Below Poverty Level



3. Childhood Blood Lead Level Screening per 1,000 (0 – 6 yrs) vs. Percent Below Poverty Level



As evidenced in the linear regression analyses above, there is not necessarily a statistically strong linear correlation between percent of a community area below poverty level and lung cancer death rates (r = 0.52). There does, however, seem to be a general trend that as poverty level rises, so does lung cancer death rate. As a statistical comparison, there is a stronger positive linear correlation between poverty level and gonorrhea cases reported in a community area (r = 0.63), and an even stronger positive linear correlation between poverty level and child blood lead level screenings in a community area (r = 0.70). While we cannot claim that any of these three health issues are necessarily caused by poverty, it does seem reasonable to assert that living in a higher poverty neighborhood such as Back of the Yards in the New City community area puts one at a higher risk for death from lung cancer, contracting gonorrhea, and exposure to lead in children.

While the Healthy Chicago key priority of cancer disparities is seemingly focused on disparities in breast cancer (and therefore working to increase access to mammograms in impoverished community areas), the key priorities of tobacco use, adolescent health, access to care, and healthy homes should seemingly be targeting high incidence of lung cancer death rates. Depending on interest in lung cancer and other respiratory diseases, I might prompt students to investigate these priorities for resources and solutions. I will likely also prompt students to investigate pollution and toxicity in their neighborhood in addition to the areas listed above. The city of Chicago, and particularly neighborhoods such as Back of the Yards, are well known for many pollution issues both past and present.

It is important to note that the background information detailed above on the health of Chicago and the Back of the Yards neighborhood in the New City community area is not meant to be exhaustive. It is merely meant to set the stage for student inquiry and research into relevant health topics of their choice that they find to be prevalent in their neighborhoods. It will, of course, be important to discuss the quality of the data and the meanings of the measures used to describe and quantify health with students.

Teaching Strategies

End Goal

In order to explain the teaching strategies I will employ in this unit, I find it necessary first to describe the summative assessment for this unit. Keeping my end goal in mind for this set of students in this initial unit will be integral in selecting the most effective methods and strategies.

Students will ultimately work in teams of three to research health problems that are prevalent in their neighborhoods. They will first use measures of central tendency, measures of dispersion, and frequency diagrams to summarize the demographics and health statistics they find. They will then identify possible factors contributing to these health problems and utilize a least squares regression method to analyze the linear correlation between these factors and the problem. Students can examine both physiological and social factors affecting these issues, though there will be a greater emphasis on the social factors given their limited biology knowledge at the beginning of the year. Finally, students will examine current and potential resources and solutions for these health problems in their neighborhood. They will be expected to work as a team to select and employ the most effective mathematical tools to examine these resources and solutions and make recommendations for the future health of their communities. This component of the summative assessment will represent my students' growing abilities to apply core statistical methods and other mathematical skills to issues that are relevant to their lives.

The presentation of these statistical analyses to each other also represents an important end goal for this unit. Students will need to effectively communicate their findings to the class, beginning to practice their roles as health advocates in their communities with each other. We will compile the class findings, and the class will decide the most effective way to share this information with the school community. In addition, we will work together to write questions addressing the biological underpinnings of these health problems as they are discussed. Students will take on the responsibility of working to understand the answers to these questions throughout their studies in 11th Grade DP Biology. Reports will be revisited in our math class at the conclusion of the school year to reflect on progress in understanding the physiological factors of health problems as well as any updates to resources and solutions within their neighborhoods. This component of the summative assessment will represent my students' growing ability to communicate with mathematics fluently and effectively, and to understand health issues and advocate for their right to adequate health resources.

Initial Classroom Structures

We will work together at the beginning of this unit to establish initial classroom structures that make students responsible for their own learning in this unit and subsequent units. As 11th grade students accepted into the rigorous Diploma Program at our school, this group brings a wealth of positive strategies and skills to the classroom that I intend to tap into. Through a series of individual, team, and class questioning activities, we will establish classroom routines and procedures, as well as consequences that we will all uphold and enforce. Some examples of these classroom protocols are detailed below.

Weekly Agendas

Upon establishing the core components of each week (team explorations, direct instruction, team practice problems, individual practice problems, individual quiz or exam), we will work together to build our typical

weekly agenda in terms of sequence of activities and the duration of each. I will use this template to then provide a detailed weekly plan to the students on a regular basis so that they can prepare accordingly. Students will have choice as to when they submit each assignment based on when we schedule the activity. For example, if team practice problems are scheduled for Thursdays, we might agree they are routinely due every Friday along with the individual practice problems. I have found that the student choice allowed in setting the weekly agenda structure helps students better understand their role and responsibility in class, and it often increases their investment in class.

Team Explorations

Team explorations will occur at least weekly and will be integral in engaging students in the material, as well as leading students to make connections to prior knowledge, making the math more accessible to them. Teams will need to work independently from other teams and the teacher for a specified amount of time, and then present their findings at the end. Correct answers will not be emphasized, and in fact, explanations of incorrect solutions will be encouraged so that we can better understand the process each team went through in their attempt to solve the problem set.

Direct Instruction

Direct Instruction will be limited to 15-minute intervals maximum during this unit. Upon completion of the instructional objectives, students will be given time to update and compare notes, as well as to ask relevant questions from the lesson in order to ensure that they have a complete and thorough record to use in later activities.

Team and Individual Practice Problems

These problems will be a mix of textbook problems as well as application problems I compile related to health statistics in Chicago. When we determine our weekly agenda template, we will set aside one day of “study group” time, where students will meet with the same team for one full class period. We will work to set out student and teacher expectations for this time in class. I have found this study group day to be integral in higher-level mathematics studies in giving the students time for regular, supported practice of difficult content. We will also work to lay out a procedure for checking in these practice problems through study group teams – both for completion and correctness. I will typically not take the lead on checking in these types of practice problems.

Individual Quiz or Test

Students will know that as individuals they will be responsible for demonstrating an understanding of certain mathematical skills and content on a weekly basis in the form of a quiz or test.

Ideas for Progress

Think-Pair-Share

It is often extremely difficult for students to feel comfortable participating in mathematical discourse. As we discuss statistical methods and health topics, I will routinely turn to Think-Pair-Share time in order for students to gain comfort in participating in discussion. This time will be built into and around direct instruction and team exploration activities, as well as class warm-ups and other relevant times.

Double Entry Journals

Individual and team practice problem sets will require students to map out selected solutions in double-entry journal style – mathematical work on the left side of the page and written explanations of the steps and their thought process on the right. I will also utilize this format during direct instruction from time to time to model the correct format as well as the utility of this practice. Through this strategy, I hope to increase student understanding and ability in communicating mathematical processes (both correct steps and missteps).

Reflection and Self-Assessment

Students will be expected to assess and reflect formally and informally on their performance in this class on a regular basis. I will alternate the focus between grades, participation, team collaboration, and understanding of material. We will use this assessment and reflection time to refine our classroom protocols as necessary.

Student-Led Warm-Ups

Every class will start with a 5-10 minute warm-up of sorts to prepare for the day, particularly while I take care of any administrative tasks. Initially, I will develop the preview or review problems utilized as the warm-up, and volunteers will be taken to present. As we progress through the unit, we will begin to assign roles in the warm-ups so that they truly become student-led. Some students will be responsible for researching relevant health statistics, others for writing application problems related to these statistics, and others for presenting their solutions once the class has had a chance to attempt each problem. We will rotate roles to ensure that everyone takes on each role at least once by the end of the unit.

Classroom Activities

Beginning of Unit - Team Exploration – Health Surveys

During the first week of this unit, students will work in teams to design and conduct a health survey in their neighborhood. Teams will be responsible for compiling the data they collect and communicating it to the class effectively. This exploration will serve as an introduction to the unit as well as a statistics pre-test to gauge students' prior knowledge and skill level.

We will begin this activity with a Think-Pair-Share addressing the questions: "What does it mean to be healthy? What does it mean to be unhealthy?" We will work as a class to develop a working definition of health. Students will then work in teams to write ten survey questions that measure the health of a person. They will be required to write both quantitative and qualitative questions. Teams will share their questions and get feedback as to how to revise them. Next we will work as a class to identify a reasonable sample population to represent their neighborhood. Teams will then administer the survey to their identified audience. Teams will reconvene with their data sets and work (with or without technology) to summarize their data graphically. Each team will present their findings to the class. As audience members, students will be expected to rate each presentation on clarity and completeness using a rubric, and these ratings will be coupled with my own assessment to assign a grade to each team. We will conclude the activity with another Think-Pair-Share addressing the questions: "How much of your health is under your own control? What are the major health problems you may face in your lifetime?" as well as any other follow-up questions the class may have. This

activity will help me to refine future activities in the unit to better address student ability and interest.

Middle of Unit - Team Practice Problems – Reading a Statistical Report

In the middle of this unit, after we have learned basic statistical methods, teams will be asked to practice applying these methods to various reports on health data sets in Chicago. This activity will serve as an important step in students identifying a focus for their final reports for this unit as described in the teaching strategies section above. Teams will identify a particular health problem of interest and apply appropriate methods to process and summarize the available data. Teams will also present their findings to the class. This activity will serve as mid-unit assessment of student's progress on applying statistical methods as well as communicating effectively with mathematics to an audience.

I will begin by familiarizing students with the various reports and resources I used in my discussion of Health in Chicago above. Students will then choose a particular health problem that is prevalent in the city of Chicago. Teams will pull the necessary reports that break the prevalence of this problem down in each of the 77 community areas in Chicago. Teams will work to summarize the prevalence of this health problem by calculating all measures of central tendency and dispersion for these 77 data points. They will also interpret the meaning of each of these measures of central tendency and dispersion in context. Teams will then work to graphically represent this set of data by creating all frequency diagrams – they will create a frequency and cumulative frequency table, a cumulative frequency diagram, a box-and-whiskers plot, and a histogram. In each representation, they will identify where their community area of New City falls. Finally, teams will identify at least three factors that they believe could be related to this health problem in the community areas where it is most prevalent. These factors will be the independent variables they investigate in their linear correlation analyses in future activities. Each team will present their findings to the class. As audience members, students will be expected again to rate each presentation on clarity and completeness using a rubric, and these ratings will be coupled with my own assessment to assign a grade to each team.

Appendix: Alignment to Standards

In the interest of making this unit most useful to the widest audience, I have identified the Common Core Math Standards that align to this unit below.

Measures of Central Tendency and Dispersion Standards

CCSS.MATH.CONTENT.HSS.ID.A.2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

CCSS.MATH.CONTENT.HSS.ID.A.3. Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

Grouping Data and Measuring Frequency Standard

CCSS.MATH.CONTENT.HSS.ID.B.5. Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data.

Linear Correlation Standards

CCSS.MATH.CONTENT.HSS.ID.C.7. Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

CCSS.MATH.CONTENT.HSS.ID.C.8. Compute (using technology) and interpret the correlation coefficient of a linear fit.

CCSS.MATH.CONTENT.HSS.ID.C.9. Distinguish between correlation and causation.

Resources

Note that all online resources below were referenced on 7/30/15.

Notes/Bibliography

1. This website provides additional information on IB programs: http://www.ibo.org/
2. Student text *Mathematics Standard Level for the IB Diploma* was utilized to develop the scope and sequence of statistical methods for this course. Citation: Fannon, Kadelberg, Woolley, and Ward. *Mathematics Standard Level for the IB Diploma*. Cambridge: Cambridge University Press, 2012.
3. This website serves as a resource for general facts and statistics on the City of Chicago: https://www.cityofchicago.org/city/en/about/facts.html
4. This website serves as a resource for additional health statistics on the City of Chicago: http://www.cityofchicago.org/city/en/depts/cdph/provdrs/healthychicago.html
5. The following link provides access to the 2013 Annual Report published by the Chicago Department of Public Health: https://www.cityofchicago.org/content/dam/city/depts/cdph/CDPH/HealthyChicagoAnnualReport2013.pdf
6. The following link provides access to the June 2015 monthly update on the Healthy Chicago initiative: http://www.cityofchicago.org/content/dam/city/depts/cdph/CDPH/HCUpdateJune2015.pdf
7. This link provides access to an interactive health atlas focused on the City of Chicago: http://www.chicagohealthatlas.org/
8. The following website provides additional public health statistics for the City of Chicago: http://www.cityofchicago.org/city/en/depts/cdph/provdrs/pol_plan_report/svcs/office_of_epidemiologydataanalysisrequests.html
9. The following provides demographical information on the New City community area: https://en.wikipedia.org/wiki/New_City,_Chicago
10. The following provides a link to the Back of the Yards College Preparatory High School website: http://boycp.org/

Technology Help Websites

TI-84 Plus Graphing Calculator

If you need assistance with using the TI-84 Plus Graphing Calculator in your instruction, I recommend this website: https://education.ti.com/en/us/products/calculators/graphing-calculators/ti-84-plus/features/features-summary. Note the menu on the left that provides options such as Downloads (particularly the guidebook), Support Resources, and Tools for Teachers.

Microsoft Excel

If you need assistance with using Microsoft Excel in your instruction beyond the Help function, I recommend visiting this abstract: http://academic-publishing.org/pdfs/01c-xl-stats_extract.pdf and perhaps ordering the text if you find it useful.