

Teaching Mathematics for Social Justice

Applications and Ideas

Mathematics Teachers

Winston-Salem Forsyth County Schools

Mathematics Education Students and Faculty

Wake Forest University

NCCTM

Fall, 2005

[Jail, Education and Opportunities](#) Deborah Arrington-Martin, *Parkland H.S.*

[Who Does the Lottery Benefit?](#) Justin Allman, *Wake Forest U.*

[The Race of Baseball All-Stars](#) Diana Liberto, *Wake Forest U.*

[Who's Being Left Behind?](#) John Troutman, *Wake Forest U.*

[Analysis of Gas Prices](#) Chris Vaughan, *Reynolds H.S.*

[Poverty](#) Leah McCoy, *Wake Forest U.*

Questions or Comments: mccoy@wfu.edu

I designed this activity to give my Algebra I student an opportunity to explore data from the department of correction. This activity includes the NC Standard Course of Study as well as standards from the National Council of Teachers of Mathematics.

My planned approach included several important concepts. Students had the opportunity to build on their previous mathematical understanding while learning new concepts. This approach helped the students to become progressively more aware of the connection among various mathematical topics and how it relates to our state correctional system. Students modeled mathematical representations while using technology. Students could view and capture mathematical relationships from different perspectives. Students explored conjectures using data tables and graphs relationships. Students then made inferences to predict future data.

To ensure fairness, equity and access to learning for all my students in my class I did several things. First, I arranged the desks in small group format to encourage student interaction. This gave students an opportunity to learn to respect individual and group differences. Secondly, I made sure all students had access to a TI-83 or TI-84 Silver Edition calculator and the web. Lastly, students were exposed to several methods to model mathematical arguments.

As I analysis this lesson students had the opportunity to communicate their ideas with confidence and I was able to give immediate feedback as soon as students made discoveries. Students evaluated their own process of thinking and the method they chose as compared to the one chosen by another student. Student success was immediate and students realized that misunderstanding is part of the learning process. Students explored a range of options and methods to foster their understanding of mathematics. Students gain through this lesson a wide variety of data representations. These representations support different ways of thinking mathematically. Students gained the ability to reason to make conjectures, justify their reasoning and communicate with their peers about the correctional system.

In the future I will be adding the use of excel and power point as part of this activity. I will also be adding a piece for Discrete Mathematics students. Students will explore the dynamics of population growth, statistics and probability as it relates to the correctional population.

OBJECTIVE:

1. Students will research data provided by the state department of corrections.

<http://statelibrary.dcr.state.nc.us/iss/ncdataresources.html>

This web page provides subject specific reports back to 1993 about various categories of crime. Releases tend to run a year or two behind the current one. Different years' data are presented differently. Also, while PDF files are the usual release formats, Excel files are available for some reports in some years. Annual Summary Reports list crime index data for all counties as well as rates for specific crimes for selected cities and counties. (This data is reported from North Carolina as part of the FBI's Uniform Crime Report Separate report covers delinquency arrests. Different specific crime indexes provide statewide data and ten-year trend reports. For instance, the most recent year's crime indexes include murder, property crime, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft, arson and simple assault.

2. Students will collect data to make graphs.
3. Students will analyze collected data and their graphs.
4. Students will make conjectures about the future as it applies to their collected data.

TOPICS:

Here are a few samples.

Who is entering the prison system?

Teens in prison

Women and Crimes

How many prisons should be built in the future?

Why are prisons overcrowded?

Drugs and Jail

Life of an inmate

Salary of an inmate

Crime Clock and future jobs in NC

Where is the money

Essential Idea: North Carolina Crime and the Society

Task 1:

Choose an area of interest.

Research the area.

Collect data

Task 2:

Using the information collected make a graph, make predictions and conjectures about future crimes and job possibilities in the state. Write an essay reporting your findings. Be sure to cite your sources. Make sure you answer the how, why, who, when, what and what next.

Who does the Lottery Benefit?

Justin Allman, Wake Forest University

Level/Course

The social justice goal of this lesson is that students consider some of the arguments against the lottery—mainly that it functions as a regressive tax and is not a profitable investment.

The mathematics of this lesson is most appropriate for a course in Discrete Math or an Algebra II class which covers combinations, permutations, and probability.

Objectives

- The learner will use data analysis to model real-world problems.
- The learner will use logic and deductive reasoning to draw conclusions and solve real-world problems.

Activities

Introduction

The lottery is a current topic of debate in North Carolina. The introduction provides that students have some understanding of the purpose of lotteries as well as one of the arguments against the lottery. The Georgia Lottery is given as an example. An extended research project in which students investigate the lottery debate in North Carolina more deeply is also recommended, though not included here.

Part I: Can playing the lottery be profitable?

Often, people think that paying one or two dollars for the chance to win hundreds of thousands or even millions of dollars is a wise investment. Students are asked to determine probabilities of winning the Fantasy 5 game in Georgia and to analyze the expected value of buying a lottery ticket.

Part II: Who is harmed the most by playing the lottery?

Students will be asked to determine the cost of playing the lottery with a “realistic” chance of winning. Then they will be asked to consider the effect that this spending would have on a families income—comparing and contrasting a family near the poverty line and a family living on Georgia’s median income. Students will be asked to synthesize their findings from Parts I and II and justify an argument of whether or not the lottery functions as a regressive tax.

Assessment

For each part, students are asked to calculate or determine various probabilities and statistics. They must then use these numbers to justify claims about the lottery. Either in groups or individually, they should submit written and/or oral reports of their calculations, findings, and conclusions.

Introduction

1. What is a lottery?

Many states have adopted lotteries whose main purpose is to provide additional state funds to pay for and improve the state's education. For every lottery ticket purchased, a portion of the ticket sale goes toward the state education budget, while the remainder of the money from the ticket sale goes toward the jackpot! Usually, hundreds, thousands, even ten thousands of lottery tickets are sold everyday across any given state, creating huge jackpots. The larger the jackpot gets, the more people that are tempted to play.

2. Example: The Georgia Lottery

- A. How it works:** The state of Georgia uses the revenue generated from its lottery to provide college scholarships to its residents. The Hope Scholarship program offers tuition scholarships to any Georgia resident who has graduated from high school with at least a 3.0 GPA, with one stipulation: the recipient must attend a public institution of higher education (that is, not a private school) in the state of Georgia. The scholarship is renewable for four years provided that the student keeps a 3.0 GPA throughout college.
- B. So what's the catch?** Lotteries ideally provide extra money to improve education within a state. But many critics cite shortcomings of the lottery. Mainly, the lottery displaces wealth regressively. In other words, because poor and working class citizens are particularly attracted to the lottery due to the hope of winning huge sums of money, changing their lives and climbing the economic ladder, they buy more lottery tickets than the average citizen. But the chances of winning are *extremely* small. So the money, which is put into the lottery by families who can afford it the least, gets redistributed to families who mostly do not need financial help in the form of college scholarships. In general, poor citizens pay for middle-class and even rich citizens to go to college. Of course, over time this accentuates the income gap between poor and rich, making improvement of socio-economic status from generation to generation extremely difficult. In short, the lottery destroys the American dream by "taxing the poor" and giving to the rich.
- C. Fantasy 5:** The Fantasy 5 game is one of the most popular lotto games in Georgia. You will use the details of the game to investigate this complaint—that the lottery is not a profitable investment, and it displaces wealth regressively.

Part I: Can playing the lottery be profitable?

1. Fantasy 5—How the game works

The Fantasy 5 game is one of Georgia's most popular lotto games because it allows players to choose the numbers they wish to be entered into the drawing instead of being assigned a number. This ensures that the drawing cannot be fixed! Furthermore, drawings occur everyday with lots of winners (relatively speaking).

How to play

On the Fantasy 5 card, the player selects 5 numbers, all between 0 and 39. Numbers cannot be repeated. Then the player pays \$1 for every time he/she wishes the chosen numbers to be entered in a drawing.

For example, a player selects the numbers 8, 14, 17, 22, and 31. He wants these numbers to be entered in 30 drawings. He must fill out the lotto card with his chosen numbers and then pay \$30—one dollar for each drawing entered.

How to win

Five numbers are drawn at random. Prizes are awarded in the following manner:

Match	Prize Awarded
0 or 1	No prize
2	1 free play
3	3 rd Prize
4	2 nd Prize
5	1 st Prize

The prize amounts are determined by how many players buy tickets each day. If no player wins the first prize, then the money which would have been awarded rolls over to the next day's 1st prize pool. This way, if there is no winner, the jackpot continues to increase.

2. Exercises

a. For the Fantasy 5 lotto game, find:

- i. The number of ways to draw the 5 random winning numbers
- ii. The number of ways to match all 5
- iii. The number of ways to match 4
- iv. The number of ways to match 3
- v. The number of ways to match 2
- vi. The number of ways to match 1 or 0.

b. Calculate the probabilities to match 0 or 1 (one event because both mean you lose!), 2, 3, 4, and 5. Fill in the table.

Matched	Probability
0 or 1	
2	
3	
4	
5	

- c. Use Table I to find the average payout for each of the winning events (matching 3, 4, and 5 of the numbers). Then, determine how much you win or lose on average in each of the five possibilities (matching 0 or 1, 2, 3, 4, or 5) by multiplying the amount you would win in a given event by the probability that the event occurs (Hint: Remember that you paid \$1, regardless of whether or not you win. Also, because matching 2 wins you a free play, consider this event the same as having a payout of \$1).
- d. Now using your results, determine the expected value for one play (entering one drawing for \$1) of the Fantasy 5 lottery. How much on average will you win (or lose) each time you play Fantasy 5? Using your calculations to justify your answer, explain whether or not the lottery is a profitable investment.

Table I

Average Payouts for Fantasy 5 Lotto Game			
	3 of 5	4 of 5	5 of 5
Payout Amounts	\$ 19.00	\$ 400.00	\$ 33,930.00
	\$ 20.00	\$ 515.00	\$ 59,278.00
	\$ 21.00	\$ 514.00	\$ 101,668.00
	\$ 20.00	\$ 399.00	\$ 217,866.00
	\$ 20.00	\$ 377.00	\$ 330,980.00
	\$ 16.00	\$ 314.00	\$ 269,088.00
	\$ 23.00	\$ 571.00	\$ 25,718.00
	\$ 14.00	\$ 249.00	\$ 91,117.00
	\$ 19.00	\$ 399.00	\$ 197,930.00
	\$ 19.00	\$ 335.00	\$ 53,626.00
	\$ 20.00	\$ 467.00	\$ 198,598.00
	\$ 14.00	\$ 251.00	\$ 118,056.00
	\$ 15.00	\$ 328.00	\$ 269,565.00
	\$ 19.00	\$ 457.00	\$ 260,211.00
	\$ 20.00	\$ 502.00	\$ 314,855.00
	\$ 17.00	\$ 324.00	\$ 206,544.00
	\$ 15.00	\$ 368.00	\$ 173,966.00
	\$ 17.00	\$ 297.00	\$ 564,898.00
	\$ 21.00	\$ 508.00	\$ 261,819.00
	\$ 20.00	\$ 427.00	\$ 311,253.00
	\$ 16.00	\$ 291.00	\$ 670,538.00
	\$ 16.00	\$ 292.00	\$ 56,257.00
	\$ 22.00	\$ 438.00	\$ 88,735.00
	\$ 14.00	\$ 243.00	\$ 105,256.00
	\$ 17.00	\$ 452.00	\$ 696,461.00
	\$ 17.00	\$ 314.00	\$ 57,896.00
	\$ 15.00	\$ 318.00	\$ 316,216.00
	\$ 15.00	\$ 276.00	\$ 218,869.00
	\$ 19.00	\$ 405.00	\$ 33,930.00
	\$ 20.00	\$ 482.00	\$ 59,278.00
\$ 19.00	\$ 417.00	\$ 118,056.00	
\$ 13.00	\$ 303.00	\$ 260,211.00	
\$ 13.00	\$ 282.00	\$ 314,855.00	
\$ 16.00	\$ 326.00	\$ 173,966.00	
\$ 16.00	\$ 395.00	\$ 173,966.00	
\$ 17.00	\$ 476.00	\$ 173,966.00	
\$ 16.00	\$ 343.00	\$ 88,735.00	
\$ 17.00	\$ 376.00	\$ 105,256.00	
Averages	\$ 17.55	\$ 379.76	\$204,563.50

Source: The Georgia Lottery Corporation

Part II: Who is harmed the most by playing the lottery?

1. How much does it cost to play the lottery well?

- a. Suppose you want to give yourself a realistic shot at winning a decent prize from the lottery. Because the average payout for matching 3 numbers is only \$17.55, hardly life-changing, you decide this prize is not even worth shooting for. Therefore, you decide that you want to give yourself a 30% shot at winning the first or second prize. In other words, you want to ensure you have a 30% chance to match at least 4 of the numbers. How many times do you need to play Fantasy 5 to ensure a 30% chance of winning the first or second prize? How much would this cost?
- b. Now suppose that 2 families of 4, Family A and Family B, are playing Fantasy 5 as outlined above. That is, they want to ensure a 30% chance of winning first or second prize. Family A has a median income at the poverty line, and Family B has an income equal to the median income for families of four in Georgia. Use Tables II-A and II-B to determine the income for both families. In terms of percentage, how much more must Family A spend to play Fantasy 5 than Family B?
- c. How many times must you play Fantasy 5 to ensure a 30% chance of winning first prize? What about just 5%?

2. Where do the winners live?

- a. Table II-C lists the cities and towns in Georgia where the most recent first prize tickets were sold. Determine the average median income for a city where winning tickets were sold. (Hint: don't forget to include cities/towns that sold more than one ticket the appropriate number of times).
- b. Next, compare the average median income for towns with winning tickets to the median income for families in Georgia. You can find data on Georgia family incomes in Table II-A. Is the average median income of towns with lottery winners more or less than the median income of Georgia families?
- c. Using your calculations from Part I: Can the lottery be profitable?, from the exercises above about ensuring a 30% success rate, and your observations about median incomes in Georgia, comment on whether or not the lottery is a "tax on the poor." Argue whether or not the lottery benefits or harms citizens. Use math to justify your claims!

Table II-A

Georgia Median Family Income			
		Lower	Upper
	Estimate	Bound	Bound
Georgia			
Total:	49745	48971	50519
2-person families	45775	44133	47417
3-person families	49855	47768	51942
4-person families	58060	54857	61263
5-person families	53908	51637	56179
6-person families	52947	44477	61417
7-or-more-person families	51928	46674	57182

Given in 2004 inflation adjusted dollars

Source: US census bureau

Table II-B

2004 Poverty Line	
Size of Family Unit	Weighted Average Thresholds
1-person families	9645
2-person families	12334
3-person families	15067
4-person families	19307
5-person families	22831
6-person families	25788
7-person families	29236
8-person families	32641
9-or-more-person families	39048

Source: US census bureau

Table II-C

**Median Incomes of Cities/Towns where Winning
Tickets were Sold**

City	# First Prizes	Median Income	(Median Income) x (# First Prizes)
Alpharetta	1	71207	71207
Ashburn	1	18702	18702
Athens	1	28403	28403
Atlanta	2	32635	65270
Augusta	1	37194	37194
Austell	1	38933	38933
Cadwell	1	32727	32727
Cohutta	1	41563	41563
Columbus	1	34798	34798
Cumming	1	38237	38237
Dalton	1	34312	34312
Doraville	1	40641	40641
Douglasville	1	45289	45289
Fayetteville	1	55208	55208
Flowery Branch	2	35478	70956
Forest Park	1	33556	33556
Griffin	1	30088	30088
Helen	1	32917	32917
Jackson	1	28472	28472
Jenkinsburg	1	40417	40417
Kingsland	1	41303	41303
Lyons	1	21202	21202
McDonough	1	41482	41482
Morrow	1	46569	46569
Moultrie	1	22193	22193
Peachtree City	1	76458	76458
Rentz	1	25000	25000
Rex	2	52714	105428
Stone Mountain	1	38603	38603
Suwanee	1	84038	84038
Sylvester	1	24114	24114
Trenton	1	34612	34612
Tucker	1	59953	59953
West Point	1	31886	31886
Totals:	37	1350904	1471731

Source: US Census Bureau & The Georgia Lottery Corporation

Who does the Lottery Benefit? Solutions

Part I

2. a.

i. # of ways to draw 5 numbers from 39, without repeating:

$${}_{39}C_5 = \frac{39!}{5!(39-5)!} = \frac{39!}{5!(34!)} = \frac{39 \cdot 38 \cdot 37 \cdot 36 \cdot 35}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 575,757$$

ii. There are 5 "good" numbers and 34 "bad" numbers. # of ways to match all five—must choose all 5 "good" numbers and 0 "bad" numbers:

$$({}_5C_5) ({}_{34}C_0) = (1)(1) = 1$$

iii. # ways to match 4:

$$({}_5C_4) ({}_{34}C_1) = 170$$

iv. # ways to match 3:

$$({}_5C_3) ({}_{34}C_2) = 5610$$

v. # ways to match 2:

$$({}_5C_2) ({}_{34}C_3) = 59840$$

vi. Matching 1 and matching 0 are mutually exclusive events, So the # of ways to match 0 or 1 will be the # ways to match zero *added* to the # of ways to match 1:

$$({}_5C_0) ({}_{34}C_5) + ({}_5C_1) ({}_{34}C_4) = 287,256 + 231,880 = 519,136$$

2.b. The probability to match 0 or 1 is given by:

$$P(0 \text{ or } 1) = \frac{(\# \text{ of ways to match } 0) + (\# \text{ of ways to match } 1)}{\# \text{ ways to choose } 5 \text{ from } 39 \text{ w/ out repeating}} = \frac{519,136}{575,757} \approx 0.886$$

To match 2:

$$P(2) = \frac{59840}{575757} \approx 0.104;$$

and so on for matching 3, 4, and 5.

The completed table should be:

Matched	Probability
0 or 1	0.886
2	0.104
3	0.00974
4	0.000295
5	1.73×10^{-6}

2.c.

-When you match 0 or 1, you lose \$1. You spent \$1 and won nothing. Then, in this event, you win:

$$(-\$1)(0.886) = -\$0.886$$

-When you match 2, you break even. You spent \$1, and you won \$1. Then, in this event, you win:

$$(\$0)(0.104) = \$0$$

-When you match 3, you win (on average) \$16.55. You spent \$1, and you won \$17.55. Then, in this event, you win:

$$(\$16.55)(0.00974) = \$0.161$$

-When you match 4, you win (on average) \$378.76. You spent \$1, and you won \$379.76. Then, in this event, you win:

$$(\$378.76)(0.000295) = \$0.112$$

-When you match 5, you win (on average) \$204,562.50. You spent \$1, and you won \$204,563.50. Then in this event, you win:

$$(\$204,563.50)(1.73 \times 10^{-6}) = \$0.354$$

2.d. The expected value is defined as: $\sum_x xP(x)$.

We've already calculated all the terms of the sum. So the expected value is:

$$EV = \$0.354 + \$0.112 + \$0.161 + \$0 + (-\$0.886) = -\$0.259$$

Because of the negative expected value for a given play, on average, you can expect a return of -\$0.26 every time you spend \$1 on the lottery. Therefore, the lottery is *not* a good investment because you should never expect to win money in the long run. In fact, if you played 100 times, you should expect to lose $100 \times \$0.26 = \26 in terms of equity, not to mention the \$100 you lost in buying tickets.

Part II

1.a. The probability to match 4 or 5 is given by:

$$P(4) + P(5) = 2.95 \times 10^{-4} + 1.73 \times 10^{-6} = 2.97 \times 10^{-4}.$$

Then, the number of times you need to play the lottery to give yourself a 30% chance is given by

$$n = \frac{0.30}{P(4) + P(5)} = \frac{0.30}{2.97 \times 10^{-4}} \approx 1011, \text{ because each play is independent.}$$

1.b. Each play costs \$1, so it would cost \$1011 to get a 30% chance. For Family A, this represents

$$\frac{1011}{19307} \approx 0.052 = 5.2\% \text{ of their yearly income.}$$

For Family B, this represents

$$\frac{1011}{58060} \approx 0.017 = 1.7\% \text{ of their yearly income.}$$

Then Family A spends

$$5.2 - 1.7 = 3.5\%$$

more of their income to have a realistic shot at winning. Or, Family A spends

$$\frac{5.2}{1.7} \approx 3.06 \text{ times more of their income, in terms of percentage, than Family B.}$$

1.c. $n = \frac{0.30}{1.73 \times 10^{-6}} = 173,410$ to give a 30% chance.

To give a 5% chance, you must still play

$$n = \frac{0.05}{1.73 \times 10^{-6}} \approx 28902 \text{ times.}$$

2.a. To get the average median income we must divide the total in column 3 of Table II-C by 37.

$$\frac{\$1,471,731}{37} \approx \$39,776$$

2.b. The median income for families of 4 in Georgia is \$58,060 and the average median income for towns where lottery winners were located is \$39,776—a difference of \$18284. We could assert that there is something systematically different in terms of income between the population of lottery players and the total population of Georgia families.

2.c. Students should touch on the negative expected value, the higher percentage of income that poor families must contribute to have a decent probability of winning (i.e. it acts as a regressive tax), and the fact that the population of lottery players appears to have a lower income than the whole population of Georgia families.

The Race of Baseball All-Stars

Diana Liberto, Wake Forest University

Level/Course: Algebra I

Social Justice Purpose: Students will examine baseball statistics and decide whether race played a part in All-Star Voting.

Objectives:

- Students will use graphs and data analysis to model real-world problems.
- Students will use logic and deductive reasoning to draw conclusions and solve problems.

Part I. Calculating Baseball Statistics

Students will examine baseball statistics of American League All- Star Baseball Players from 2005 and 1975. They calculate the batting average (BA), extra-base power (EB) and total skill score (TSS) for each reserve player who played in the 1975 and 2005 All- Star Game. Students will enter this data into Tables 1A and 1B. This lesson can be enhanced by using Excel and entering equations in the cells.

Part II. Comparing Scores

Students will then compare the scores of the starting players, chosen by the fans, and reserve players, chosen by the coaches. With the total score computed students will be able to discover whether the fans choose the most skilled player for each position.

Students will use Table 2A and 2B to compare the race of the player in each position where the reserve score was higher then the starters score. Students should also compare the years.

Assessment: Students will write a conclusion about race and the all-star voting, using their data.

History of the Major League Baseball (MLB) All-Star Game

All-Star teams were originally selected by the managers and the fans for the 1933 and 1934 games. From 1935 through 1946, managers selected the entire team for each league. From 1947 to 1957, fans chose the team's starters and the manager chose the pitchers and the remaining players. From 1958 through 1969 managers, players and coaches made the All-Star Team selections. In 1970, the vote again returned to the fans for the selection of the starters for each team and remains there today.

The baseball player at each position who gets the most votes starts the game at that position. The All-Star Manager then picks the reserves for each position in the game. The Manager typically picks one back up for Catcher, Second Baseman, Shortstop and Third Baseman. He picks two back ups for the First Baseman and four backups for Outfield.

Fans can vote for players using any standard they prefer. Fans may vote for players by the team they play for, the amount of runs they score, where they went to college, how physically attractive they are, race and the list could go on. One would hope the fans would base on ability but spectators of the sport know that this is not always true.

Your Mission

Using your math talent you must calculate important baseball statistics for the starting and reserve players of the 2005 and 1975 All-Star games. After you have correctly found important baseball statistics you will then decide if fans could have placed their vote by looking at race over skill.

Part One: Calculating Baseball Statistics

All necessary data has been collected for students and placed in Table 1A and 1B: Baseball Statistics. The data was collected on September 18, 2005 at www.mlb.com. All data on the all-star starters from 2005 and 1975 has been entered. It is your job to fill in the statistics for the reserve players. The information not filled in is the reserve's batting average (BA), extra-base power (EB) and their total skill score (TSS).

The formula used to find a player's batting average (BA) is their total number of hits (H) divided by their total number of at bats (AB).

$$BA = H / AB$$

The formula for extra-base power (EB) is calculated by subtracting a player's total number of hits (H) from their total bases (TB) and dividing that by their at bats (AB).

$$EB = (TB - H) / AB$$

There are many ways one could assign an overall score to a baseball player. The formula we will use considers a player's offensive and defensive skill. Our formula for discovering a player's total skill score (TSS) is the batting average (BA) plus the extra-base power (EB) plus the fielding percentage (FP), divided by three.

$$TSS = (BA + EB + FP) / 3$$

Table 1A: Player Statistics 1975

Starters Bold	Race	Position	BA	TB	H	AB	R	EB	FP	TSS
T. Munson	W	C	0.318	256	190	597	83	0.111	0.972	0.467
B. Freehan	W	C		170	105	427	42		0.991	
Tenace	W	1B	0.255	231	127	498	83	0.209	0.998	0.487
G. Scott	W	1B		318	176	617	86		0.989	
C. Yastrzemski	W	1B		220	146	543	91		0.996	
R. Carew	B	2B	0.359	266	192	535	89	0.138	0.973	0.490
J. Orta	H	2B		244	165	542	64		0.978	
Campaneris	H	SS	0.265	168	135	509	69	0.065	0.962	0.431
B. Dent	W	SS		205	159	602	52		0.981	
T. Harrah	W	SS		239	153	522	81		0.963	
G. Nettles	W	3B	0.260	250	155	597	83	0.159	0.964	0.461
D. Chalk	W	3B		177	140	513	59		0.976	
B. Bonds	B	OF	0.270	271	143	529	93	0.242	0.987	0.500
R. Jackson	B	OF	0.253	303	150	593	91	0.258	0.965	0.492
J. Rudi	W	OF	0.278	231	130	468	66	0.216	1.000	0.498
H. Aaron	B	OF		165	109	465	45		1.000	
M. Hargrove	W	OF		216	157	519	82		0.984	
G. Hendrick	B	OF		242	145	561	82		0.983	
F. Lynn	W	OF		299	175	528	103		0.983	
H. McRae	B	OF		212	147	480	58		0.986	
C. Washington	B	OF		250	182	590	86		0.978	

Table 1B: Player Statistics 2005

Starters Bold	Race	Position	BA	TB	H	AB	R	EB	FP	TSS
J. Varitek	W	C	0.284	223	128	451	68	0.211	0.991	0.495
I. Rodriguez	H	C		224	139	500	71		0.995	
M. Teixeira	W	1B	0.297	356	185	622	110	0.275	0.999	0.524
P. Konerko	W	1B		297	157	563	95		0.996	
M. Sweeney	W	1B		232	134	454	60		0.998	
B. Roberts	W	2B	0.301	285	169	561	92	0.207	0.988	0.499
A. Soriano	H	2B		316	166	618	99		0.973	
M. Tejada	H	SS	0.307	331	193	628	85	0.220	0.970	0.499
M. Young	W	SS		336	215	645	113		0.973	
A. Rodriguez	H	3B	0.321	284	186	580	118	0.169	1.000	0.497
M. Mora	H	3B		260	160	569	77		0.960	
V. Guerrero	H	OF	0.317	290	162	511	94	0.250	0.988	0.519
J. Damon	H	OF	0.313	264	188	600	110	0.127	0.985	0.475
M. Ramirez	H	OF	0.288	309	154	534	107	0.290	0.973	0.517
G. Anderson	B	OF		337	157	558	64		0.976	
S. Podsednik	W	OF		170	140	486	78		0.989	
G. Sheffield	B	OF		284	164	561	100		0.987	
I. Suzuki	A	OF		285	196	654	108		0.995	
D. Ortiz	H	DH	0.296	350	171	578	115	0.310	0.976	0.527
S. Hillenbrand	W	DH		267	173	590	90		0.954	

Part Two: What Starters do not measure up to the Reserves?

You will now compare the players in which the reserve had a higher score than the starting player. To help compare these players fill in Table 2. Comparison of Stats with the players at each position where the reserve was better than the starter. In this table you will also enter the race of the starter and the reserve. You will use this information to compare the race of players in which the reserve player had a higher total score than the starter.

Table 2A: Comparison of Stats 1975

Fill in this chart with the Back Up Players that had a better score than the Starters.

Back Up Player	TSS Score	Starter	TSS Score	Back-Up's Race	Starter's Race	Therefore.... (Draw conclusion)

Table 2B: Comparison of Stats 2005

Fill in this chart with the Back Up Players that had a better score than the Starters.

Back Up Player	TSS Score	Starter	TSS Score	Back-Up's Race	Starter's Race	Therefore.... (Draw conclusion)

After examining the information in Table 2A and 2B, write a conclusion with the information you have discovered. Your conclusion should include but not be limited to the following questions:

- What race got chosen over another race more frequently?
- Did any race fail to ever be chosen over another?
- Did the error in picking the most qualified player improve over the 30 year span?
- Do you think race could have played a factor in fan voting?

Solutions:

Table 1A

<i>1975 Starters</i>	<i>Race</i>	<i>Position</i>	<i>BA</i>	<i>TB</i>	<i>H</i>	<i>AB</i>	<i>R</i>	<i>EB</i>	<i>FP</i>	<i>TSS</i>
T. Munson	W	C	0.318	256	190	597	83	0.111	0.972	0.467
B. Freehan	W	C	0.246	170	105	427	42	0.152	0.991	0.463
Tenace	W	1B	0.255	231	127	498	83	0.209	0.998	0.487
G. Scott	W	1B	0.285	318	176	617	86	0.230	0.989	0.501
C. Yastrzemeski	W	1B	0.269	220	146	543	91	0.136	0.996	0.467
R. Carew	B	2B	0.359	266	192	535	89	0.138	0.973	0.490
J. Orta	H	2B	0.304	244	165	542	64	0.146	0.978	0.476
Campaneris	H	SS	0.265	168	135	509	69	0.065	0.962	0.431
B. Dent	W	SS	0.264	205	159	602	52	0.076	0.981	0.441
T. Harrah	W	SS	0.293	239	153	522	81	0.165	0.963	0.474
G. Nettles	W	3B	0.260	250	155	597	83	0.159	0.964	0.461
D. Chalk	W	3B	0.273	177	140	513	59	0.072	0.976	0.440
B. Bonds	B	OF	0.270	271	143	529	93	0.242	0.987	0.500
R. Jackson	B	OF	0.253	303	150	593	91	0.258	0.965	0.492
J. Rudi	W	OF	0.278	231	130	468	66	0.216	1.000	0.498
H. Aaron	B	OF	0.234	165	109	465	45	0.120	1.000	0.452
M. Hargrove	W	OF	0.303	216	157	519	82	0.114	0.984	0.467
G. Hendrick	B	OF	0.258	242	145	561	82	0.173	0.983	0.471
F. Lynn	W	OF	0.331	299	175	528	103	0.235	0.983	0.516
H. McRae	B	OF	0.306	212	147	480	58	0.135	0.986	0.476
C. Washington	B	OF	0.308	250	182	590	86	0.115	0.978	0.467

Table 2A

<i>Back Up Player</i>	<i>Score</i>	<i>Starter</i>	<i>Score</i>	<i>Back-Up's Race</i>	<i>Starter's Race</i>	<i>Therefore....(Draw conclusion)</i>
G. Scott	.501	Tenace	.487	W	W	White got it over White
B. Dent	.441	Campaneris	.431	W	H	Hispanic got it over White
T. Harrah	.474	Campaneris	.431	W	H	Hispanic got it over White
F. Lynn	.516	R. Jackson	.492	W	B	Black got it over White

Solutions (cont.):

Table 1B

2005 Players	Race	Position	BA	TB	H	AB	R	EB	FP	TSS
J. Varitek	W	C	0.284	223	128	451	68	0.211	0.991	0.495
I. Rodriguez	H	C	0.278	224	139	500	71	0.170	0.995	0.481
M. Teixeira	W	1B	0.297	356	185	622	110	0.275	0.999	0.524
P. Konerko	W	1B	0.279	297	157	563	95	0.249	0.996	0.508
M. Sweeney	W	1B	0.295	232	134	454	60	0.216	0.998	0.503
B. Roberts	W	2B	0.301	285	169	561	92	0.207	0.988	0.499
A. Soriano	H	2B	0.265	326	164	618	99	0.262	0.973	0.500
M. Tejada	H	SS	0.307	331	193	628	85	0.220	0.970	0.499
M. Young	W	SS	0.320	338	214	668	113	0.186	0.973	0.493
A. Rodriguez	H	3B	0.321	284	186	580	118	0.169	1.000	0.497
M. Mora	H	3B	0.281	260	160	569	77	0.176	0.960	0.472
V. Guerrero	H	OF	0.317	290	162	511	94	0.250	0.988	0.519
J. Damon	H	OF	0.313	264	188	600	110	0.127	0.985	0.475
M. Ramirez	H	OF	0.288	309	154	534	107	0.290	0.973	0.517
G. Anderson	B	OF	0.281	337	157	558	64	0.323	0.976	0.527
S. Podsednik	W	OF	0.288	170	140	486	78	0.062	0.989	0.446
G. Sheffield	B	OF	0.292	284	164	561	100	0.214	0.987	0.498
I. Suzuki	A	OF	0.300	285	196	654	108	0.136	0.995	0.477
D. Ortiz	H	DH	0.296	350	171	578	115	0.310	0.976	0.527
S. Hillenbrand	W	DH	0.293	267	173	590	90	0.159	0.954	0.469

Table 2B

Back Up Player	Score	Starter	Score	Back-Up's Race	Starter's Race	Therefore....(Draw conclusion)
A. Soriano	.500	B. Roberts	.499	H	W	White got it over Hispanic
G. Anderson	.527	Damon/Ramirez	.475/.517	B	H/H	Hispanic got it over Black
G. Sheffield	.498	Damon	.475	B	H	Hispanic got it over Black
I. Suzuki	.477	Damon	.475	A	H	Hispanic got it over Asian

Who's Being Left Behind?

John R. Troutman, Wake Forest University

Exploring the racial achievement gap on standardized tests

Social Justice Goals: Students will consider commonly reported standardized test statistics and think about some of the institutional causes of these problems.

Course level: Recommended: Discrete Math

Portions applicable to: Algebra I, Algebra II, Advance Functions and Modeling

A. Objectives:

The learners will...

- Represent statistics on student achievement by race on a graph
- Analyze the differences between achievement scores of students from different racial subgroups
- Use TI-83 graphing calculators or Microsoft Excel to find lines of best fit for each of the racial subgroups and discuss the varying rates of change
- Use this information to make predictions about the future of racial achievement gaps in school.

B. Activities:

- Part I: Discover meaning of standardized test scores and how they are calculated
- Part II: Graph recent test scores and analyze the achievement gap between students of different races to discover how it is changing
- Part III: Analyze the impact of economic inequity on those test scores

C. Assessment:

Collect the students' completed handouts, graphs, and final writing assignment.

Lesson packet includes:

- a.) Teacher resources and answer keys
- b.) A sample normal curve
- c.) Examples of Excel Spreadsheets and graphs for part 2
- d.) Student handouts

TEACHER RESOURCES

Who's Being Left Behind? Exploring the racial achievement gap on standardized tests

Social justice motivation:

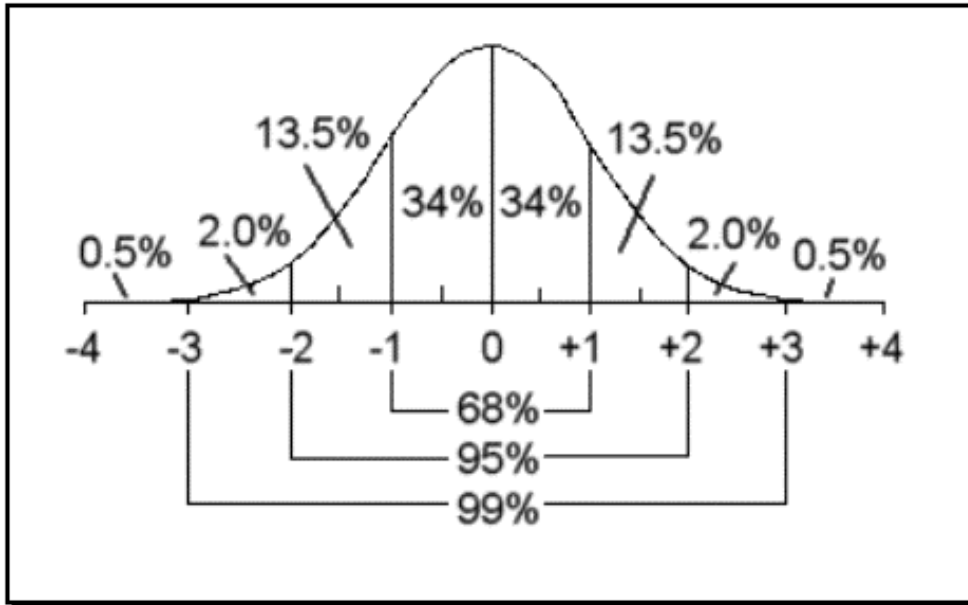
"One of the state's most prestigious high schools, East Chapel Hill High, failed federal testing standards this year. Only 20 percent of East's black students passed an end-of-course reading exam, compared with the goal of 35 percent set by the federal No Child Left Behind testing program. Likewise, 54 percent of black students passed the math exam, compared with the goal of 71 percent." "East Chapel Hill fails U.S. goals", *The News and Observer*, July 23, 2005

In the No Child Left Behind era, communities are inundated with statistics like the ones above. Without a proper understanding of standardized testing methods, statistical reports, and the social contexts that have produced these results, reports such as this one can have damaging impact on students' self concepts. These reports tend to implicitly place blame on the victims of inequitable education systems, further disassociating oppressed students from school communities. With a more complete understanding of standardized test scores, students might be empowered by these reports to advocate against the school structures that serve to reproduce current social stratifications.

Introduction: You may want to provide some background for the lesson by discussing two important characteristics of End of Course Tests.

- These tests are "norm referenced." You may want to explain this concept using the "normal-curve" (see below for an example). This kind of test score is problematic for two reasons: 1) With a normal curve, it is impossible for all students to pass the test. 2) In order to spread scores out across the normal curve, test-makers intentionally include questions on these exams that they think a high percentage of the students taking the test will get wrong. This means that norm-referenced tests are less likely to ask about the basic skills that students need to understand a subject, because these questions would be answered correctly by too many of the students.
- Secondly, the EOC's are "achievement tests" not "aptitude tests". They are (in theory) not meant to test "how smart" a student is, but how much she/he has learned over the course of the semester. This characteristic of the tests means that it is inappropriate to hold students solely responsible for their results on these tests because research has shown that many factors out of the students' control can affect how much they learn in a given course (i.e. quality of instructional materials, quality of instruction, length of time in the course, the structure of the course, etc.).

Normal Curve



Part 1: Students discover how achievement statistics are calculated and what they mean.

- a) Student/teacher discussion on how officials calculated the scores that are reported:
- Students' "raw scores" were determined based on how many questions the students answered correctly; school officials graphed these results and determined which raw scores were considered "passing" by comparing all the scores to one another.
 - Using this number a percentage of passing students for each ethnic group on each test is calculated.
 - Finally, the average percent passing for all the tests was determined by finding the average of from each of the 10 NC End of Course Tests for a given group.
 - This is recorded for each group, and is represented in the chart on the student handout.
- b) Students work through the questions on handouts one-to calculate the percent of students of ethnic group Z who passed the 4 EOC's-and two-looking more closely at actual achievement statistics.

Part 2: In this section students will present achievement data graphically, and determine lines of best fit to model the rate of change experienced by each group. Students will work through handouts 3 and 4.

Note: See answer key for an example of how the data looks graphed in Excel.

Part 3: In this section students will explore the context of those test scores, by looking at statistics about poverty rates in NC and student achievement by economically advantaged and disadvantaged students. Students work through handouts 5 and 6.

Additionally, teachers may want to hand out and discuss the final assignment with students.

Name: _____

Date: _____

**Which Children are Left Behind?
Exploring the racial achievement gap on standardized tests**

Part 1, Where did these numbers come from?

Use the made up numbers in the chart below to determine the average percent of students from ethnic group Z who passed these four NC End of Course Tests in 1999.

Table 1. Group Z Scores

Test	US History	Algebra I	Geometry	Biology
Number of students from Ethnic group Z who took the test in 1999	2341	1050	2290	2106
Number of students from Ethnic group Z who "passed" the test in 1999	2144	1000	2095	1993

Determine the percent of students who passed each test, and write your answer (round to the nearest tenth) in the space provided.

1. US History: _____ %
2. Algebra I: _____ %
3. Geometry: _____ %
4. Biology: _____ %
5. What is the average of these percentages? _____ %

State officials used a similar process with all 10 NC End of Course tests to determine the numbers you will see in the chart on the next page.

Part 1 (cont.)

Table 2. Percentages of Students who passed the North Carolina End of Course Tests in the years 2001-2004, broken down by racial categorizations

Year	White	African-American	Latino/a	Asian-American	Native American
2001-2002	81 %	52 %	60 %	78 %	59 %
2002-2003	83 %	58 %	61 %	80 %	66 %
2003-2004	86 %	61 %	66 %	83 %	69 %

Use the information from the class discussion on EOC test scores and the above chart to answer the following questions.

6. If you randomly selected 100 Native American students from North Carolina, how many would you expect to have received a passing grade on all their End of Course Tests in the 2001-2002 school year? Explain your answer.

7. True or False: The number of Latino/a students who failed the end of course test in 2001-2002 was greater than the number of White students who failed the test that year. Explain your answer.

8. An investigative reporter learns from an anonymous government official that individual African-American students received the highest scores on 10 of the 11 End of Course tests in the 2003-2004 school year. Based on this information, she publishes a story saying the scores of students from that year must have been misreported. Is the reporter justified in forming this conclusion? Why or why not?

Name: _____

Date: _____

Which Children are Left Behind?

Part 2, What do these numbers mean?:

1. Plot the data on a graph with the x-axis representing the years in which the data was collected and the y-axis representing the percentage of students who passed the test that year.

Using a Microsoft Excel or a graphing calculator determine the line of best fit (or trend line) the data points of each racial group. Write the equation of the line in standard form below (round all decimals to the nearest hundredth).

2. White: _____

3. African American: _____

4. Latino/a: _____

5. Asian-American: _____

6. Native American: _____

7. Which line has greatest **slope** largest? What does this tell you about the students in this group?

8. Which line has the largest **y-intercept** largest? What does this tell you about the students in this group?

We can use the lines whose equations you just determined to predict the percentages of students who will pass the End of Course tests in the future if these scores continue to increase in the same way. You can do this by simply looking at your graphs.

Part 2 (cont.)

Using the lines of best fit that you have created, predict the percentage of students in each subgroup who will pass the End of Course tests in the year 2007-2008 school year.

9. White: _____

10. African American: _____

11. Latino/a: _____

12. Asian American: _____

13. Native American: _____

14. Should we expect there to still be a gap between the percentages of African-American students and White students who pass the tests in this school year if student achievement continues as it has over the last 3 years? Explain your answer.

15. Is this gap smaller or larger than it was in the 2003-2004 school year? Explain your answer.

Name: _____

Date: _____

Which Children are Left Behind?

Part 3, What else is involved?:

There are lots of other factors that seem to predict how well a student will do on the End of Course tests. One of these factors is whether or not a student is "Economically Disadvantaged." Using the information in the table below, create a new scatter plot showing the relationships between the percentages of students in different economic groups who pass the End of Course tests.

Table 3. Percentages of Students who passed the North Carolina End of Course Tests in the years 2001-2004, by students' economic status

Year	Economically Disadvantaged	Not Economically Disadvantaged
2001-2002	59	85
2002-2003	63	87
2003-2004	66	82

1. What advantages do students who are not economically disadvantaged have over poor students that play a role in creating this difference?

Below is a table with the percentages of the people in each racial group who are living in poverty. The "poverty line" is determined by government agencies; it is based on the amount of money people need to live in a certain area. People who are living "below the poverty line" do not have enough money to pay for all the things they need (food, shelter, electricity, health care, transportation, etc.).

Researchers determined the percentages in the chart below by executing the following steps for each group:

- They determined the number of people in the group who live in NC
- Using tax reports, they determine the number of people in the group who live below the poverty line
- They divided the second number by the first and multiplied by 100

Table 4. Percentage of the people in 5 racial subgroups who were living below the poverty line in the year 2000 in North Carolina

Racial Category	White	African-American	Latino/a	Asian-American	Native American
Percentage of population below the poverty line	8.4	22.9	25.2	10.1	21

2. Which two racial groups have the lowest percentages of people living in poverty? What can you say (in general) about the achievement on End of Course tests of students from these groups?

3. Which groups have the highest percentages of people living in poverty? What can you say (in general) about the achievement on End of Course tests of students from these groups?

4. Based on the information from the tables in part 3, does it surprise you that higher percentages of African-American, Latino/a, and Native American students fail End of Course tests in NC? Explain your answer.

Final assignment:

Based on the information you have seen, graphs you have created, and questions you have answered, write a paragraph describing the following:

- a.) How students' test scores are related to their income level
- b.) How an individual's income is related to her or his race
- c.) How a student's race is related to her or his test scores

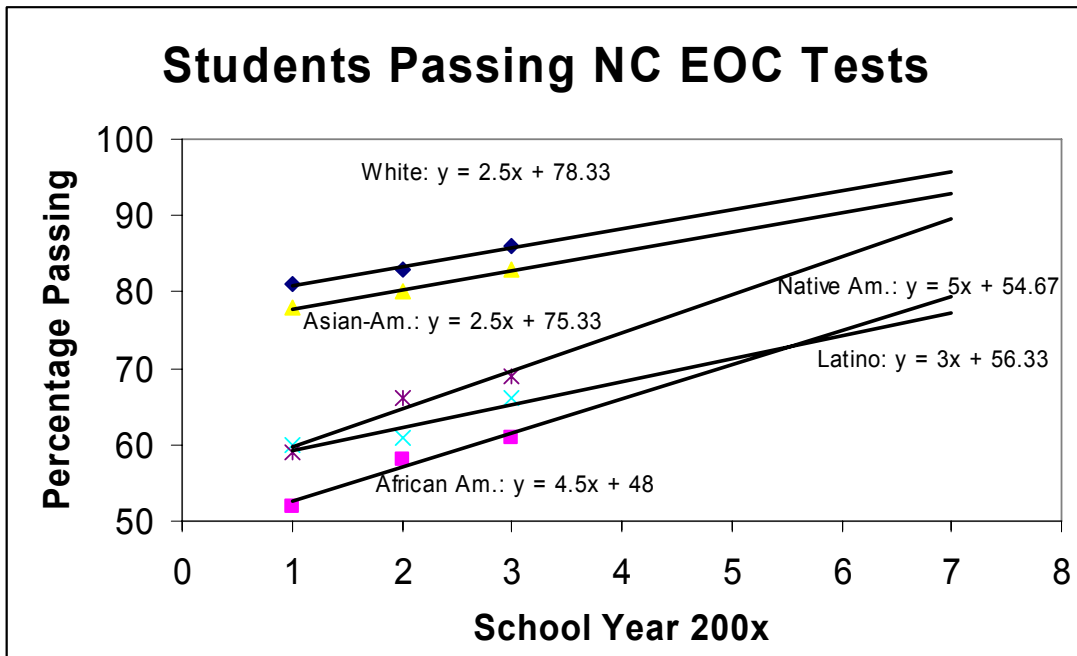
Finally, develop an educated guess about which some of the factors that might be at the root cause of differences in school achievement between students of different races.

Answer Key. Part 1.

1. US History: 91.6%,
2. Algebra I: 95.2%
3. Geometry: 91.5%
4. Biology: 94.6%
5. Average of these percentages? 92.9 %
6. 59 students. This is an average.
7. False. In the 2000 census, people of Latin American descent made up only about 4.7% of NC's population, by contrast Whites made up about 70.2% of the population. Even though the highest percentage of students passed the Algebra I exam the actual number of students who passed was the smallest.
8. No, the reporter would not be justified in making such a claim. First of all, students' actual scores are not factored into this statistic, since it only looks at the numbers of student who passed the test. A student who got all the responses correct would (for the purposes of this statistic) be no different than a student who barely passed the exam. Additionally (and perhaps more importantly in terms of the conversation around test scores with high school students), since so many students took the test it is a fallacy to think that you could use these tests to predict the performance of individual students. In reality, there are thousands of students in each lower achieving racial group who perform much better than White and Asian American students, additionally thousands of students in higher achieving racial groups score much worse than most African Americans, Latino/as, or Native Americans.

Answer Key. **Part 2**

1.



Answer Key. Part 2 (cont.)

2. White: $y=2.50x + 78.33$
3. African American: $y= 4.5x + 48$
4. Latino/a: $y=3x +56.33$
5. Asian-American: $y= 2.5x + 75.33$
6. Native American: $y=5x+54.67$
7. *The line for the Native Americans, this means that the percentage of students who pass the tests each year is increasing more quickly for this group than for any other.*
8. *The line for the White students, this means that the percentage of students who started out passing the tests was higher for this group than for any other.*
9. White: *approximately 95.8 percent*
10. African American: *approximately 79.5 percent*
11. Latino/a: *approximately 77.3 percent*
12. Asian American: *approximately 92.83 percent*
13. Native American: *approximately 89.7 percent*
14. *Yes, according to the lines of best fit, less than 80 percent of African-American students would be passing the test at this point as compared to close around 96 percent of Whites.*
15. *Yes, the prediction for the 2007-2008 school year is smaller than the gap from the data for the 03-04 year. Students can determine this either by looking at their graph and noticing that the White and African-American lines are getting closer together or by subtracting the African-American percentage from the White percentage in each year to see that while 25% more Whites passed the exams in 03-04 only about 16% more passed in the prediction for 07-08.*

Answer Key. Part 3.

1. *Here students may include any of the various advantages that students gain with money, such as ability to afford tutors, technological resources, homes that are generally conducive for studying, plenty of food, etc.*
2. *White and Asian American, students from these groups tend to pass the EOC tests more frequently than students from the other groups.*
3. *African-American, Latino/a, and Native American, these groups tend to pass the EOC tests less frequently than students from other groups.*
4. *This should not be surprising to the students. The poverty statistics suggest that a higher percentage of students in these racial groups would fall into the economically disadvantaged status. Since these students, in general, fail the exams more often, it makes sense that higher percentages of African-American, Latino/a, and Native American students fail the exams*

Answer Key. Final assignment.

In their essays, students should include some version of the following sentences:

- a.) *Students who are economically disadvantaged are more likely to fail NC EOC tests than students who are not economically disadvantaged.*
- b.) *African-American, Latino/a, and Native American individuals are more likely to live below the poverty line.*
- c.) *African-American, Latino/a, and Native American individuals are more likely to fail the NC EOC tests than their White or Asian-American counterparts.*

A Statistical Analysis of Annual Gas Prices from 1976-2005

Level/Course:

This lesson can be used and modified for teaching High School Math, Foundations of Algebra, Algebra I, Algebra II, Discrete Math, and Advanced Functions and Modeling.

Part I: Mean, Median, Mode, & Range

Objective

- Calculate, use, and interpret the mean, median, mode, range for a set of data.

Activity

Have the students find the mean yearly gas prices from 1976-2005 and write their answers in the column named "Annual Mean." Have the students do the attached worksheet titled "Mean, Median, Mode, & Range." You can also have the students create various lists by looking at certain months or a select group of years and analyze the mean or median of each individual month or the median of each year. Teach the students how to use the "SortA(" and "SortD(" functions on the calculator to rearrange the data in either ascending or descending order so that they can more easily find the range and modes.

Assessment

Students will be graded not only on mathematical computations but also on their analysis of interpretation questions written in short answer form.

Part II: Scatter Plots & Linear Regression

Objective

- Collect, organize, analyze, and display data (including scatterplots) to solve problems.
- Approximate a line of best fit for a given scatterplot; explain the meaning of the line as it relates to the problem and make predictions.

Activity

Have students look back at their answers in the "Annual" box beside each year. If you did not do the Part I activity then you can give them Table 2 which includes the needed information. Have the students create a table in either a graphing calculator or a computer program so that when a scatter plot is graphed, the years are on the x-axis and the annual mean is on the y-axis. Instruct each student to look at the scatter plot and determine if they can visual see any pattern or a positive or negative regression line. Then have the students actually plot the linear regression line that best fits the data and determine the projected cost of gas will be in the year 2030 and explain why they feel this may or may not be a good prediction.

Assessment

Students will be graded on their data plots and on their linear regression line. They will also be graded on their use of this data to predict and analysis of results. Finally the students will be graded on their group project which will be presented to the class.

Part III: Box Plots

Objective

- Collect, organize, analyze, and display data (including box plots) to solve problems.
- Calculate, use, and interpret the inter-quartile range for a set of data.
- Identify outliers and determine their effect on the a set of data.

Activity

Take the list of average annual gas prices (either from Table 2 or the chart they received and filled out in Part I) and create a box plot. The students can then trace the box plot to find the minimum, quartile 1, median, quartile 3, and maximum values. From here, have the students find the interquartile range, find outliers and then analyze whether there is bad data and if the scatter plot produces an even distribution or if the graph is skewed.

Assessment

Students will be graded on finding the key points of a box plot graph and their written description of the analysis of results with respect to good and bad data. Finally, the students will be graded on their group project that will be presented to the class.

Part IV: Scatter Plot & Median-Median Line

Objective

- Collect, organize, analyze, and display data (including scatterplots) to solve problems.
- Find the median-median line for a given scatterplot; explain the meaning of the line as it relates to the problem and make predictions.
- Compare two different types of regression lines and which is more accurate for the data.

Activity

The students create a scatter plot putting the years on the x-axis and the annual means on the y-axis and find a median-median line which could be a more consistent regression line. Using the median-median line, they can find the residuals and the root mean square error of the data. They can then use this information to extrapolate data. They should also create a line of best fit and explore the differences between the line of best fit and the median-median line and determine which is actually a more accurate line for the data.

Assessment

Students will be graded on the median-median line and the standard deviation. They will also be graded on their analysis of what the data means and their comparison of the two types of linear regression lines and their prediction of possible future gas prices.

Mean, Median, Mode, & Range Activity

Instructions: Using the attached chart that has the monthly gas prices from 1976-2005, answer the following questions.

1. Look at the chart for the monthly gas averages from 1976-2005. Fill in the column titled Annual with the mean of each year's data.
2. Further Analysis of Mean, Median, Mode and Range.
 - a. Find the median of gas prices for the year of 1982. _____
 - b. Find the median of gas prices for the year of 2001. _____
 - c. Find the range of the Annual gas prices from 1976-2005. _____
 - d. Find the range from January 1976 to August 2005. _____
 - e. Find the mode from January 1976 to December 1979. _____
 - f. Find the mode from January 2000 to August 2005. _____

Short Answer: On another sheet of paper answer the following questions.

3. Compare the means and medians of the years 1982 and 2001. What are the differences and why is there a difference. Which would be better when trying to determine what gas prices were for those two years?
4. Why do you think there is such a big range from January, 1976, to August, 2005? What types of events and reasons could there be for price increases over the 30 year period?

Scatter Plot and Linear Regression Line Activity

Instructions: Using the years as your x-axis and the annual mean as your y-axis, create a scatter plot and linear regression line to answer the following questions.

1. What is the slope of the linear regression line? _____

2. What is the Y-Intercept of the linear regression line? _____

3. What is the equation of the linear regression line in slope-intercept form?

4. Further analysis of your scatter plot and linear regression line.

a. Based on your linear regression line, what would be an estimated cost of gas in 2030?

b. Do you think this will really be the price of gas? Why or why not?

GROUP PROJECT: In your assigned group, research an alternative energy source that can be used to replace gasoline. Be prepared to present to the class the estimated cost of producing and maintaining that form of alternative energy as well as the positives/negatives of using that form of alternative energy.

Box Plot Activity

Instructions: Using the years as your x-axis and the annual mean as your y-axis, create a box plot to answer the following questions about the quality of the data.

1. What is the maximum value (the highest price of gas)? _____
2. What is the minimum value (the lowest price of gas)? _____
3. What is the median gas price? _____
4. What is the cost of gas at quartile one? _____
5. What is the cost of gas at quartile three? _____
6. Is the graph skewed to one side or balanced? _____
7. What is the interquartile range? _____
8. How many years are outliers and what are they? _____

Short Answer:

9. Is there bad data in this scatter plot? Why or why not?

10. Are we living in an outlier year? If so do you think time will prove that this is an outlier year or will there be a new trend in the price of gasoline? Explain your answer.

GROUP PROJECT: In your assigned group find all the outliers and research what events were happening in the United States and the world to determine what was going on that could have led to the price of gas being that much higher or lower than what it should have been. Be prepared to present your findings to the class in a creative manner.

Scatter Plot & Median-Median Line Activity

Instructions: Using the years as your x-axis and the annual mean as your y-axis, create a scatter plot and a median-median line to answer the following questions.

1. What is the equation of the median-median line? _____
2. What are the residuals of each year from 1976-2005?

1976	1977	1978	1979	1980	1981
1982	1983	1984	1985	1986	1987
1988	1989	1990	1991	1992	1993
1994	1995	1996	1997	1998	1999
2000	2001	2002	2003	2004	2005

3. What is the root mean square error? _____
4. What would be the cost of gas in 2030? _____
5. Based on your information, what would the maximum and minimum gas price be in 2030?

Short Answer: If you did the Scatter Plot & Linear Regression Line Activity then pull out your worksheet. If you did not do that activity, find the equation for the line of regression (a best fit line) and answer the following questions. If you need more space finish answers on the back.

5. Compare the possible gas prices from 2030 on the two different linear regression lines. Which price do you think will be more accurate and why?

6. Do you think whichever equation you picked as the best equation (line of best fit or median-median line) for this problem will always be a better regression line for all data analysis? Why?

TABLE 1: AVERAGE GAS PRICES BY MONTH FROM 1976-2005

Table comes from the U.S. Department of Labor: Bureau of Labor Statistics web site: <http://www.bls.gov>

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual Mean
1976	0.605	0.600	0.594	0.592	0.600	0.616	0.623	0.628	0.630	0.629	0.629	0.626	
1977	0.627	0.637	0.643	0.651	0.659	0.665	0.667	0.667	0.666	0.665	0.664	0.665	
1978	0.648	0.647	0.647	0.649	0.655	0.663	0.674	0.682	0.688	0.690	0.695	0.705	
1979	0.716	0.730	0.755	0.802	0.844	0.901	0.949	0.988	1.020	1.028	1.041	1.065	
1980	1.131	1.207	1.252	1.264	1.266	1.269	1.271	1.267	1.257	1.250	1.250	1.258	
1981	1.298	1.382	1.417	1.412	1.400	1.391	1.382	1.376	1.376	1.371	1.369	1.365	
1982	1.358	1.334	1.284	1.225	1.237	1.309	1.331	1.323	1.307	1.295	1.283	1.260	
1983	1.230	1.187	1.152	1.215	1.259	1.277	1.288	1.285	1.274	1.255	1.241	1.231	
1984	1.216	1.209	1.210	1.227	1.236	1.229	1.212	1.196	1.203	1.209	1.207	1.193	
1985	1.148	1.131	1.159	1.205	1.231	1.241	1.242	1.229	1.216	1.204	1.207	1.208	
1986	1.194	1.120	0.981	0.888	0.923	0.955	0.890	0.843	0.860	0.831	0.821	0.823	
1987	0.862	0.905	0.912	0.934	0.941	0.958	0.971	0.995	0.990	0.976	0.976	0.961	
1988	0.933	0.913	0.904	0.930	0.955	0.955	0.967	0.987	0.974	0.957	0.949	0.930	
1989	0.918	0.926	0.940	1.065	1.119	1.114	1.092	1.057	1.029	1.027	0.999	0.980	
1990	1.042	1.037	1.023	1.044	1.061	1.088	1.084	1.190	1.294	1.378	1.377	1.354	
1991	1.247	1.143	1.082	1.104	1.156	1.160	1.127	1.140	1.143	1.122	1.134	1.123	
1992	1.073	1.054	1.058	1.079	1.136	1.179	1.174	1.158	1.158	1.154	1.159	1.136	
1993	1.117	1.108	1.098	1.112	1.129	1.130	1.109	1.097	1.085	1.127	1.113	1.070	
1994	1.043	1.051	1.045	1.064	1.080	1.106	1.136	1.182	1.177	1.152	1.163	1.143	
1995	1.129	1.120	1.115	1.140	1.200	1.226	1.195	1.164	1.148	1.127	1.101	1.101	
1996	1.129	1.124	1.162	1.251	1.323	1.299	1.272	1.240	1.234	1.227	1.250	1.260	
1997	1.261	1.255	1.235	1.231	1.226	1.229	1.205	1.253	1.277	1.242	1.213	1.177	
1998	1.131	1.082	1.041	1.052	1.092	1.094	1.079	1.052	1.033	1.042	1.028	0.986	
1999	0.972	0.955	0.991	1.177	1.178	1.148	1.189	1.255	1.280	1.274	1.264	1.298	
2000	1.301	1.369	1.541	1.506	1.498	1.617	1.593	1.510	1.582	1.559	1.555	1.489	
2001	1.472	1.484	1.447	1.564	1.729	1.640	1.482	1.427	1.531	1.362	1.263	1.131	
2002	1.139	1.130	1.241	1.407	1.421	1.404	1.412	1.423	1.422	1.449	1.448	1.394	
2003	1.473	1.641	1.748	1.659	1.542	1.514	1.524	1.628	1.728	1.603	1.535	1.494	
2004	1.592	1.672	1.766	1.833	2.009	2.041	1.939	1.898	1.891	2.029	2.010	1.882	
2005	1.823	1.918	2.065	2.283	2.216	2.176	2.316	2.503					

TABLE 2: 1976-2005 ANNUAL AVERAGE CHART

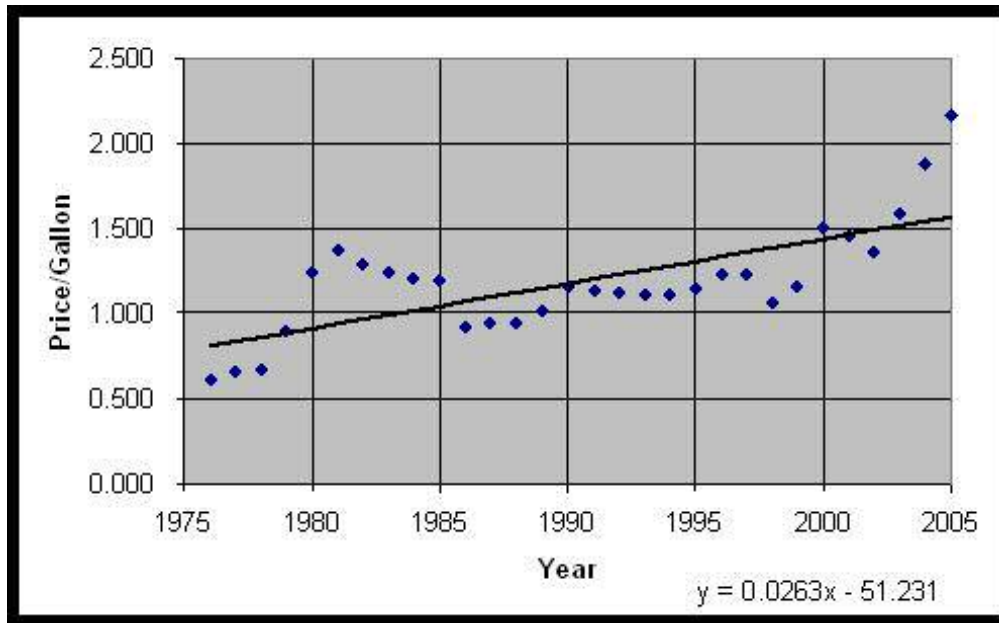
Year	Annual Mean
1976	0.614
1977	0.656
1978	0.670
1979	0.903
1980	1.245
1981	1.378
1982	1.296
1983	1.241
1984	1.212
1985	1.202
1986	0.927
1987	0.948
1988	0.946
1989	1.022
1990	1.164
1991	1.140
1992	1.127
1993	1.108
1994	1.112
1995	1.147
1996	1.231
1997	1.234
1998	1.059
1999	1.165
2000	1.510
2001	1.461
2002	1.358
2003	1.591
2004	1.880
2005	2.163

Teacher's Answer Key

Part I: Mean, Median, Mode, & Range

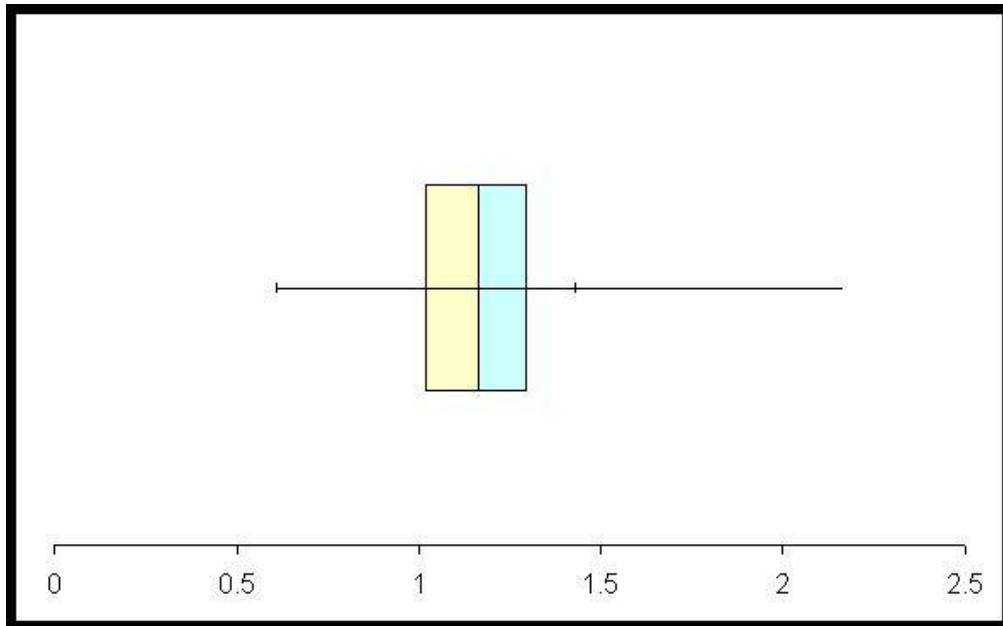
- 1) The answers for the chart can be found by looking at Table 2.
- 2)
 - a. 1.301
 - b. 1.477
 - c. 1.548
 - d. 1.911
 - e. 0.665
 - f. No Mode

Part II: Scatter Plots & Linear Regression



- 1) 0.0263
- 2) -51.231
- 3) $y = 0.0263x - 51.231$
- 4) a. 2.2306

Part III: Box Plots



- 1) 2.163
- 2) 0.614
- 3) 1.1645
- 4) 1.022
- 5) 1.296
- 6) Skewed Right
- 7) 0.411
- 8) 2004 and 2005

Part IV: Scatter Plot & Median-Median Line

- 1) $y = 0.010125x - 18.9116$
- 2)

1976	1977	1978	1979	1980	1981
-0.4814	-0.4485	-0.4457	-0.2228	0.1091	0.23198
1982	1983	1984	1985	1986	1987
0.13985	0.07473	0.0356	0.01548	-0.2697	-0.2588
1988	1989	1990	1991	1992	1993
-0.2709	-0.205	-0.0732	-0.1073	-0.1304	-0.1595
1994	1995	1996	1997	1998	1999
-0.1657	-0.1408	-0.0669	-0.074	-0.2592	-0.1633
2000	2001	2002	2003	2004	2005
0.1716	0.11248	-0.0007	0.22223	0.5011	0.77398

- 3) 0.279
- 4) about 1.6421
- 5) Maximum = 1.9211 and Minimum = 1.3631

Level/Course

The social justice goal of this lesson is that students consider poverty.

There are multiple parts that might be utilized at different mathematics levels:

- Part I is very simple mathematics and would be appropriate in middle grades or in any high school course.
- Part II involves graphing data and reasoning. It would be particularly appropriate in Discrete Mathematics or in other high school courses.
- Part III is more advanced and involves using the Trapezoid Rule for approximation of the area under a curve (definite integral), which would fit into Pre-Calculus or Advanced Functions and Modeling.

Objectives

- The learner will use graphs and data analysis to model real-world problems.
- The learner will use logic and deductive reasoning to draw conclusions and solve problems.

Activities

The activities in these lessons use all five of the NCTM process standards. The context is a real-world connection in a problem-solving format. Analysis involves reasoning with multiple representations of data and communication of conclusions.

Part I. What Is Poverty?

The reality of poverty is introduced and students are asked to consider whether they could live below the poverty level. They are asked to create a budget and to relate their own life to that of a family below the poverty line.

Part II. Who Are the Poor?

Students are given two tables from the U.S. Census Bureau and asked to analyze this data, to create graphs, and summarize their conclusions. The tables give the poverty level in the 50 states and the District of Columbia, and demographic characteristics of people in poverty such as age, race, and education.

Part III. How Has Poverty Changed?

Students are introduced to the Gini Coefficient to quantify the inequality of income across the population. Data is examined for the past 30 years and students are asked to explain results.

Part IV. What Can Be Done?

Students are given a number of sources to learn more about poverty.

Assessment

For each part of these lessons, students are asked to consider the real-world data. Either in group projects or individually, they analyze mathematically, reflect on their results, and submit written and/or oral reports of their findings and conclusions.

I. What is Poverty?

- A. About the Hurricane Katrina disaster in New Orleans, Senator Barak Obama said

Whoever was in charge of planning was so detached from the realities of inner city life in New Orleans ... that they couldn't conceive of the notion that every American couldn't "load up their family in an SUV, fill it up with \$100 worth of gasoline, stick some bottled water in the trunk and use a credit card to check into a hotel on safe ground."
(Newsweek, Sept. 19, 2005)

- B. Food, shelter, clothing, health care, transportation - these are only the beginnings of the basic necessities of modern American living. Each year, the federal government calculates the minimum amount of money required by families to meet these basic needs. The resulting calculation is what is commonly referred to as the "poverty line." For 2004, the government set the poverty guidelines as follows.

Table I. 2004 Poverty Line

Size of Family Unit	Weighted Average Thresholds
One person	9,645
Two persons	12,334
Three persons	15,067
Four persons	19,307
Five persons	22,831
Six persons	25,788
Seven persons	29,236
Eight persons	32,641
Nine persons or more	39,048

SOURCE: U.S. Census Bureau.

- C. View the documentary "**Tour Poverty USA**" from the United States Conference of Catholic Bishops at

<http://www.usccb.org/cchd/povertyusa/tour.htm>

- Do you think your family could live with this income?
- How much would it bother you to give up many of the "extras?"
- Create a budget estimating what you spend for these "extras" in a normal year.
- How does your life relate to that of a family below the poverty line?

II. Who Are the Poor?

- A.** Table II-A gives the percentages and ranks of people below the poverty level in the 50 states and the District of Columbia. Select the ten states which are highest in percent of people below the poverty level. Using a spreadsheet (or a graphing calculator or graphing by hand), construct a line graph of states (horizontal axis) and percentage of people below the poverty level (vertical axis). Then on the same graph, plot the percentages of children below the poverty level. Label the graphs of total people and of children. What can you say about these variables?
- B.** Still using Table II-A, repeat the exercise in Part A above for the bottom-ranked ten states (lowest in poverty level). What can you say about these variables? What can you say about the comparison of the top and bottom ten states?
- C.** Table II-B gives percentages of several subgroups of people who are very poor (income at less than 50% of the poverty level). Select one characteristic, and construct a bar graph comparing the percentages of those subgroups who are "very poor." What can you say about your results? Repeat for two other characteristics. Write a summary of your conclusions.

Table II-A. Percent of People Below Poverty Level - 2004

State	Percent of People Below Poverty Level	Rank for People Below Poverty Level	Percent of Children below Poverty Level	Rank for Children Below Poverty Level
United States	13.1		18.4	
Alabama	16.1	9	23.3	8
Alaska	8.2	49	11.2	48
Arizona	14.2	16	20.3	17
Arkansas	17.9	5	25.9	5
California	13.3	20	18.9	21
Colorado	11.1	32	14.5	34
Connecticut	7.6	50	10.5	50
Delaware	9.9	41	13.8	38
District of Columbia	18.9	4	33.9	1
Florida	12.2	27	17.7	24
Georgia	14.8	13	21.3	12
Hawaii	10.6	38	14.4	35
Idaho	14.5	14	19.6	18
Illinois	11.9	29	16.8	28
Indiana	10.8	36	14.8	32
Iowa	9.9	41	12.4	44
Kansas	10.5	39	12.5	42
Kentucky	17.4	7	25.0	6
Louisiana	19.4	2	30.0	3
Maine	12.3	25	17.1	27
Maryland	8.8	46	11.4	47
Massachusetts	9.2	44	12.5	42
Michigan	12.3	25	17.6	25
Minnesota	8.3	48	10.7	49
Mississippi	21.6	1	31.0	2
Missouri	11.8	30	16.2	30
Montana	14.2	16	19.2	19
Nebraska	11.0	33	13.1	40
Nevada	12.6	23	18.8	22
New Hampshire	7.6	50	9.7	51
New Jersey	8.5	47	11.8	45
New Mexico	19.3	3	27.7	4
New York	14.2	16	20.7	15
North Carolina	15.2	12	21.9	11
North Dakota	12.1	28	15.5	31
Ohio	12.5	24	18.3	23
Oklahoma	15.3	11	20.7	15
Oregon	14.1	19	19.1	20
Pennsylvania	11.7	31	16.8	28
Rhode Island	12.8	22	21.0	14
South Carolina	15.7	10	22.8	10
South Dakota	11.0	33	14.8	32
Tennessee	14.5	14	21.1	13
Texas	16.6	8	22.9	9
Utah	10.9	35	13.3	39
Vermont	9.0	45	11.7	46
Virginia	9.5	43	12.9	41
Washington	13.1	21	17.2	26
West Virginia	17.9	5	24.4	7
Wisconsin	10.7	37	14.0	36
Wyoming	10.3	40	14.0	36

Source: U.S. Census Bureau, 2004 American Community Survey

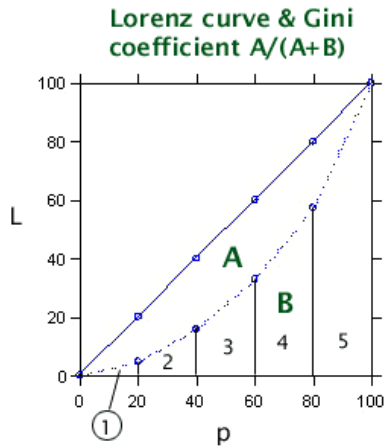
**Table II-B. Selected Characteristics of People at Specified Levels of Poverty
2004**

Subject	All People for Whom Poverty Status is Determined	People at Less than 50% of the Poverty Level
Population for whom poverty status is determined	284,577,956	5.7%
GENDER		
Male	139,214,726	5.0%
Female	145,363,230	6.4%
AGE		
Under 18 years	71,810,759	8.3%
18 to 64 years	178,561,896	5.3%
65 years and over	34,205,301	2.1%
RACE OR ETHNICITY		
White	215,298,360	4.3%
Black or African American	34,576,665	12.6%
American Indian and Alaska Native	2,137,754	11.4%
Asian	12,076,732	5.6%
Native Hawaiian and Other Pacific Islander	401,425	8.6%
Some other race	14,733,087	8.2%
Hispanic or Latino origin	40,219,766	8.6%
HOUSEHOLD TYPE		
In married-couple family households	180,844,134	1.9%
In other households	103,733,822	12.2%
EDUCATIONAL ATTAINMENT		
(Population 25 years and over)	186,534,177	
Less than high school graduate	29,976,049	8.0%
High school graduate (includes equivalency)	55,055,121	4.2%
Some college or associate's degree	51,091,603	3.1%
Bachelor's degree or higher	50,411,404	1.9%
CITIZENSHIP STATUS		
Native	250,346,192	5.6%
Foreign born	34,231,764	6.6%
DISABILITY STATUS		
With disability	37,771,428	7.2%
No disability	226,399,338	5.1%
WORK STATUS		
(Population 16 to 64 years)	186,589,012	
Worked full-time, year-round	88,904,929	0.3%
Worked less than full-time, year-round	59,720,615	6.8%
Did not work	37,963,468	15.2%
HOUSING		
In owner-occupied housing units	198,711,099	2.5%
In renter-occupied housing units	85,866,857	13.0%

Source: U.S. Census Bureau, 2004 American Community Survey

III. How Has Poverty Changed?

In economics, one measure of inequality is called the **Gini Coefficient**. This statistic allows us to quantify the distribution of income across a population. The Gini Coefficient ranges from 0 to 1, where 1 is perfect inequality (one part of the population has all the income and the rest have none) and 0 is perfect equality (all in the population have equal shares). The general shape of the graph is



By definition, the **Gini Coefficient** is the ratio of the area between the 45 degree equality line and the **Lorenz Curve** which is the graph of our population (P) and income (L) data. This area represents the amount of inequality. So, in this graph the **Gini Coefficient** is equal to Area A/(Area A + Area B). The closer the Lorenz curve is to the line of perfect equality, the smaller the Gini coefficient, and the less the inequality.

Since we don't know the equation of the curve, we will use the **Trapezoid Rule** to approximate the area (B) under the curve by finding the areas of the polygons:

Polygon 1 is a triangle, $((b * h)/2)$

Polygons 2-5 are trapezoids, $(h * (b1 + b2)/2)$.

Note: the bases are vertical lines.

(Area A + Area B) will always be $(100*100)/2$

Area B will be (Area 1 + Area 2 + Area 3 + Area 4 + Area 5)

Area A will be (Area A + Area B) - Area B

The Gini Coefficient can also be represented by the following formula

$$G = \left| 1 - \sum_{k=1}^{k=n} (X_k - X_{k-1})(Y_k + Y_{k-1}) \right|$$

where k= the number of partitions.

We will compare data in 10 year increments from 1970 to 2000.

Share of Aggregate Income Received by Households

Year	Lowest fifth	Second fifth	Third fifth	Fourth fifth	Highest fifth
2000	3.6	8.9	14.8	23.0	49.8
1990	3.8	9.6	15.9	24.0	46.6
1980	4.2	10.2	16.8	24.7	44.1
1970	4.1	10.8	17.4	24.5	43.3

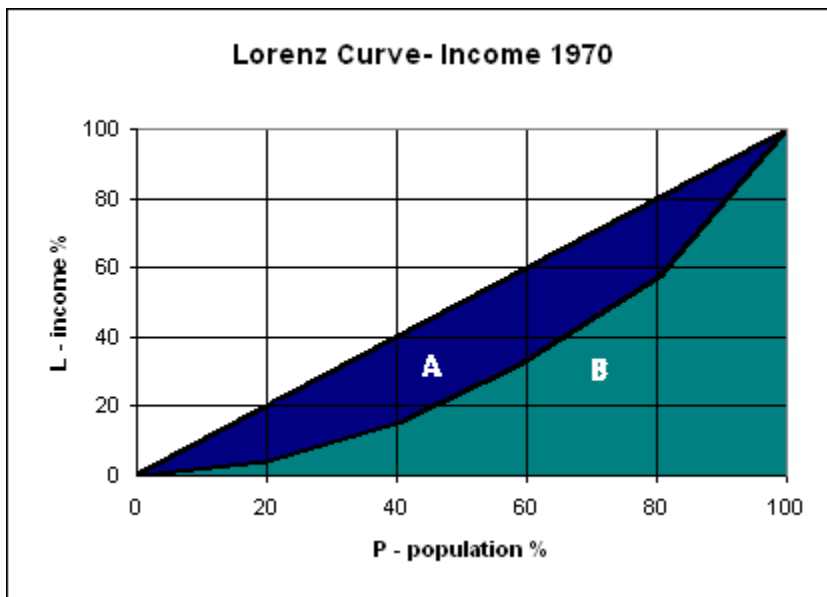
Source: U.S. Census Bureau

For the year 1970:

First, we will arrange the data in a working table and calculate **P** = the cumulative share of the population (%) and **L** = cumulative share of income (%) for each fifth of our population year. (Spreadsheet is recommended, but it could be done by hand or with calculator.)

Gini 1970

Income Category	Share of Total Income %	P %	L %
Bottom Fifth	4.1	20	4.1
2 nd Fifth	10.8	40	14.9
3 rd Fifth	17.4	60	32.3
4 th Fifth	24.5	80	56.8
Top Fifth	43.3	100	100.1



The next task is to calculate.

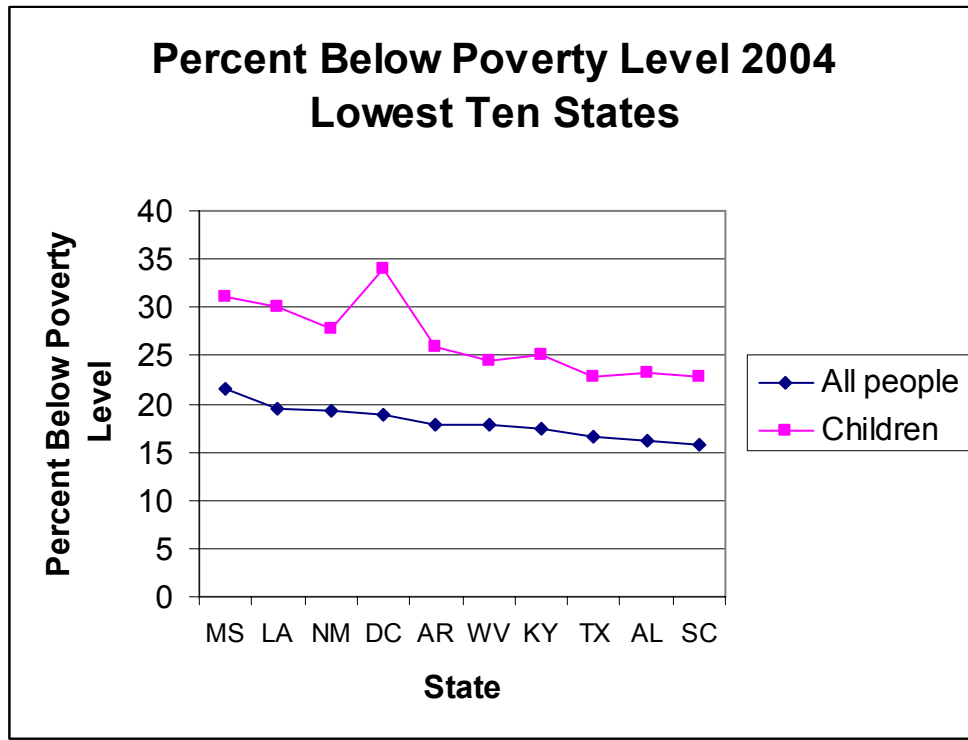
Area A + Area B	$100 \cdot 100 / 2$	5000
Area 1 = $((b \cdot h) / 2)$	$20 \cdot 4.1 / 2$	41
Area 2 = $(h \cdot (b_1 + b_2) / 2)$	$20 \cdot (4.1 + 14.9) / 2$	190
Area 3 = $(h \cdot (b_1 + b_2) / 2)$	$20 \cdot (14.9 + 32.3) / 2$	472
Area 4 = $(h \cdot (b_1 + b_2) / 2)$	$20 \cdot (32.3 + 56.8) / 2$	891
Area 5 = $(h \cdot (b_1 + b_2) / 2)$	$20 \cdot (56.8 + 100.1) / 2$	1569
Total Area B = (Area 1 + Area 2 + Area 3 + Area 4 + Area 5)		3163
Area A = (Area A + Area B) - Area B	$5000 - 3162$	1837
Gini Coefficient = $(\text{Area A} / (\text{Area A} + \text{Area B}))$	$1837 / 5000$	0.367

1. Complete the graphs and calculations for the Gini Coefficients for 1970, 1980, 1990, and 2000. (You may either put them on the same graph or create separate graphs.) Make a table of the Gini Coefficients for all the years. Look at the graphs and the Gini Coefficients. What can you say about your results? Is there a trend? What conclusions can you draw?

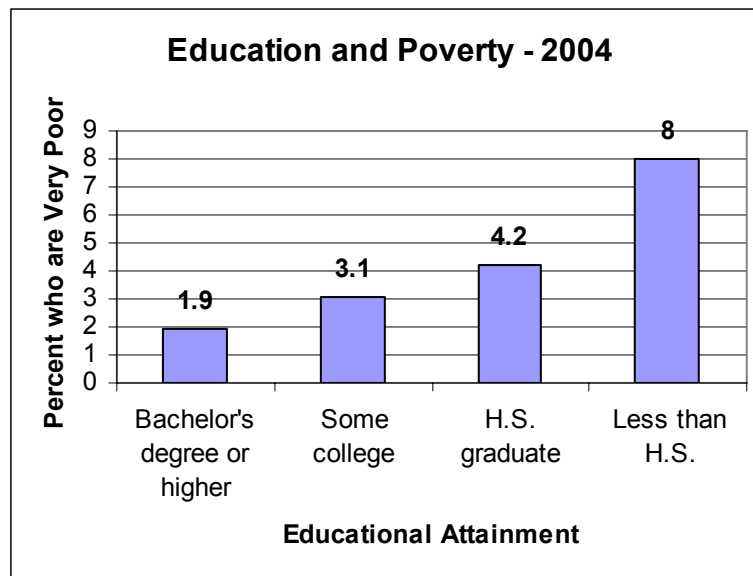
IV. What Can Be Done? Learn More. Care More. Do More.

- Children's Defense Fund
 - <http://www.childrensdefense.org/>
- Educate to End Poverty in America
 - <http://www.usccb.org/cchd/povertyusa/involved.shtml>
- Fight Hate and Promote Tolerance
 - <http://www.tolerance.org/>
- How to End Poverty (Time Magazine)
 - <http://www.time.com/time/covers/1101050314/>
- National Center for Children in Poverty (Columbia University)
 - <http://www.nccp.org/>
- Race, Poverty, and Katrina (NPR)
 - <http://www.npr.org/templates/story/story.php?storyId=4829446>
- U.S. Census Bureau Poverty Data
 - <http://www.census.gov/hhes/www/poverty/poverty.html>

II-A.



II-C. Example:



III.

y70		y80		y90		y00	
P	L	P	L	P	L	P	L
20	4.1	20	4.2	20	3.8	20	3.6
40	14.9	40	14.4	40	13.4	40	12.5
60	32.3	60	31.2	60	29.3	60	27.3
80	56.8	80	55.9	80	53.3	80	50.3
100	100.1	100	100	100	99.9	100	100.1

	y1970	y1980	y1990	y2000
A + B	5000	5000	5000	5000
A1	41	42	38	36
A2	190	186	172	161
A3	472	456	427	398
A4	891	871	826	776
A5	1568	1559	1533	1504
Ar B	3163	3114	2996	2875
Ar A	1837	1886	2004	2125
Gini	0.367	0.377	0.401	0.425

