# Do We Have a Linear Trend?

## A Beginner's Approach to Analysis of Trends in Community Health Indicators

Linda L. Remy, PhD
Ted Clay, MS
Geraldine Oliva, MD MPH

October 2005

# TABLE OF CONTENTS

# TABLES AND FIGURES

# INTRODUCTION

Monitoring trends in community health status is an important public health function. Statewide trends may differ from local trends. Monitoring trends also is of value in assessing the impact of public health interventions. It is important that program staff distinguish between significant differences in a number or rate from year to year, and significant trends over 5 or more years. However, few public health managers have the analytic expertise to determine whether a trend may be occurring and, if so, whether it is statistically significant. These guidelines are intended to help program staff and epidemiologists from local health jurisdictions make such determinations. This is NOT intended to teach program managers to do trend analysis. It may help to educate them how to make appropriate trend analysis requests to data analysts or epidemiologists with whom they consult.

Since 1993, the Family Health Outcomes Project (FHOP) has provided training and technical assistance to California's 61 local maternal, adolescent and child health (MCAH) programs. We have helped counties build capacity to use health related data effectively for community health assessment, program planning, and evaluation. Over this period, FHOP has developed written methodologic guidelines for data analysis and a set of automated analytic tools.

To assist local health departments with their data analysis functions, FHOP developed *Guidelines for Statistical Analysis of Public Health Data with Attention to Small Numbers. The Guidelines* describe analytic approaches to address the problem of comparing statistics describing two groups in a given year, or comparing two observations from different years. The *Guidelines* do not discuss approaches to trend analysis.

FHOP also developed a set of Excel Data Templates to provide a tabular overview of a set of key MCAH indicators for a 12 year period and to generate statistical confidence intervals for these data. This enables comparison with state benchmark figures, for which a rate ratio and its confidence interval also are generated. Each template constructs a graph plotting indicator values over time for county data, the state benchmark estimate, and the Healthy People 2010 goals. The graph includes confidence limits for each year of data for the county indicators.

In the past, template users have compared confidence intervals between Point A and Point B. Based upon that comparison, they have made conclusions about whether there had been a significant change in the indicator value between the two time points. Although these differences may be valid, they are inadequate to provide information on *trends.* Consequently, users tend to "eyeball" graphs to see if indicator values appear to be increasing or decreasing without regard to statistical significance. This may cause a problem when analysts present graphs to community groups or decision makers who may be unfamiliar with statistical analysis and then make unwarranted conclusions. Therefore, in response to requests from county MCAH directors, we developed this briefing on linear trend analysis. We suggest that it be used in conjunction with the *Statistical Guidelines* and our Excel Data Templates.

This briefing describes some alternative ways of analyzing trends, and provides guidance for interpreting trend analysis methods. We ask the reader first to become familiar with our *Guidelines for Statistical Analysis* and our *Data Templates.*

All the resources described above are available at our website: http://www.ucsf.edu/fhop/.

We begin with a discussion on how one first approaches trend analysis. Subsequent sections include a number of examples to help the reader understand the issues involved. The Appendices include instructions for using the Trend Templates to calculate trends as well as other statistical information.

# HOW TO APPROACH TREND ANALYSIS

**Look at the Data.** Simplify the picture as much as possible. Look at local data separately from state data. Do not include confidence intervals. Do the data form a generally straight or curved shape? If the shape is generally straight for your jurisdiction **and** for the state (or other comparison), you may use simple trend analysis methods.

If the data curve, simple trend analysis methods may be inappropriate. Using simple trend methods in the face of a complex trend would result in misleading conclusions. If, upon reviewing the data, a complex trend seems likely for either the local jurisdiction or the state, it would be best to obtain help from skilled statisticians.

**Context is Essential.** Trend data bear on a question at the heart of quantitative thinking: "Compared to what?" Consider what one might conclude by simply drawing a graph of county data with one line connecting county rates for two years. A simple comparison of two widely spaced time points often can miss the true picture. Adding a 2-point comparison with nearby counties, the state, or nation would give a better sense of context, but still may mislead. Nothing can replace examining the data in context, from year to year over an adequate time period. Suspected trends can be confirmed by observing linear patterns. The impact of policy or programs can be suggested by changes in direction and curves at critical times. After visually examining the data, and with sufficient numbers of cases, perhaps as few as three time points would essentially convey a 10-year trend. However, to gain this knowledge, you must start by looking carefully at all the data points, first separately then in context.

**Trend Analysis is Time Sensitive**. Sometimes 5-year trends (1991-1995 or 1996-2000) or 3-year rolling averages miss the big picture. Consider a county that introduced an intervention to reduce teen pregnancies. In considering whether to do trend analysis, take care to establish the appropriate intervals. It is important to group time periods and intervals so they can end before the program started and have a sufficient interval after the start to assess effectiveness. In establishing intervals, also carefully compare trends with other similar counties, the region, or the state. Note: It is widely accepted in the field of program evaluation that program effects cannot be assessed until at least 6 months after program initiation.

**Quantitative Data Displays Must Have Graphical Integrity**. The primary motivation for displaying data is to facilitate understanding of complex phenomena. Graphic data displays are critical to conveying unbiased statistical results. Of all methods to analyze and communicate statistical information, well-designed data graphics are usually the simplest and most powerful way to describe, explore, and summarize quantitative data. We strongly recommend that MCAH analysts hone their graphic skills.

# A FEW TECHNICAL TERMS

While we wish it were otherwise, we have to introduce a few technical terms to help readers understand what they are seeing. Regression analysis is the usual statistical technique by which a series of observations are "fit" to a trend line. This line may be straight (the simple case, presented here) or may have some other shape (the complex case).

When we are using regression to examine data for a trend, we are trying to calculate the "best" line to describe the fit between our outcome of interest, referred to as the dependent variable (the rate in a given year, typically shown in a trend analysis formula as "Y") and our independent factor or factors whose impact we are assessing ("X"). In trend analysis the independent variable typically is time, in this case expressed as years. In this situation, three basic statistics are most helpful: the intercept, the slope, and the probability (P-value). It also is helpful to understand the difference between the confidence interval (CI) and the standard error (SE). These are described in the immediately following paragraphs.

**A rate is a measure of some event, disease, or condition in relation to a unit of population, along with some specification of time.** Births, asthma, injury are examples of events, diseases, or conditions. Examples of population units are women age 15 to 44, children age 0 to 4. A year is a typical time period. A constant (100, 1000, 10000, etc) is used to standardize the rate (express it in a common unit) across communities with varying population sizes to allow comparisons. As an example, an asthma rate is calculated by dividing the number of children age 0 to 4 from a given community hospitalized for asthma in a given year (24), divided by the number of children age 0 to 4 in that community that year (5,209), multiplied by a constant (10,000). The result is expressed as 46.1 asthma admissions per 10,000 population of children age 0 to 4.

**The slope is the average rate of change over the years being examined**. The slope is the most important part of a trend model. It represents the rate at which change occurs over time. Interpreting the slope is straightforward. If the value of the slope is zero, there was no trend. The indicator did not change with time. A graph showing a zero slope has a horizontal line. If the slope has a positive value, the rate is increasing. If it is negative, the rate is decreasing. We interpret the slope to mean that, on average, the rate changed by the slope value each year. The issue is whether the slope value is significantly different from zero, i.e., is the P-value less than or equal to 0.05. If it is, we have a linear trend. If it is not, we must conclude there is no meaningful trend. Example 1 illustrates a simple linear trend.

Complex regression analysis adds more factors and/or different mathematical techniques to the basic formula. However, the basic purpose is the same: to "fit" the "best" line to summarize the data trend. Examples 2 and 3 illustrate situations using different mathematical models.

Finally, the data can reflect a more complex picture with a change (or bend) in the slope, a change in the intercept, or both. Example 4 illustrates a change in the slope.

**The intercept is the estimated rate at the start of the period examined**. In the regression formula, the intercept typically is shown as a number, without the X following it.

Under certain circumstances, the intercept can change without an accompanying change in the slope. For example, the percent of infants born with low birthweight could drop from one year to the next in the absence of a trend and remain at this lower value. This change could be statistically significant in the absence of a trend. Example 2 illustrates such a situation. The

trend is flat from 1992 through 1994, then drops and remains essentially flat through the end of the period. This shows as a statistically significant trend, even though it in fact is due entirely to a drop in the intercept.

**The probability value (P-value) reflects the likelihood that the model results are due to chance**. Whether the model is simple or complex, the P-value summarizes the statistical results for both the variables in the model and the final model overall. The interested reader should refer to an introductory statistics text to understand how the P-value is calculated. Here, it suffices to say that a P-value less than 0.05 ($P<0.05$) is statistically significant. In the presence of $P<0.05$, you are safe to reject the null hypothesis that there is no trend or there are no differences among the comparisons. If the P-value is greater than 0.05, then you must accept that there is no trend and/or no differences among the comparison groups. In this latter case, you would conclude that the county results are similar to the state results.

**Standard Error**. Most public health professionals are familiar with the confidence interval (CI), calculated on one year at a time. The CI is useful to help understand if your county rate is different from or similar to the state rate in any given year or compared with **one** other year. However, the CI is not the appropriate statistic to use in trend analysis. In fact, comparing one year to another over a 10-year period without a linear regression analysis could be misleading. Instead, the standard error (SE), or the CI around the slope and intercept is used to determine if the rate is different from its comparison, e.g., the state.

To help readers unfamiliar with regression methods for trend analysis, we developed a Trend Template (shown in Appendix A) which we used to do Example 1, a simple linear trend. The template calculates the basic statistics described above, and includes others that help to get a full understanding of the trend model. This template is inappropriate for curvilinear or any other complex data trend.

# EXAMPLE 1: A LINEAR TREND, THE SIMPLE CASE

For the 10-year period 1991 to 2000, Table 1 shows the annual state and county fertility rate and the county confidence interval (CI) per 1,000 women age 15 to 44. This shows the state fertility rate dropped from 83.5 to 71.9 per 1,000 women age 15 to 44. During this same period, the county fertility rate dropped from 71.4 to 63.6. Figure 1 also shows that the county upper CI always has been below the state rate. Thus we can conclude that the county fertility rate has been significantly lower than the state rate for at least ten years.

| Table 1: County Fertility Rate per 1,000 Women Age 15 to 44 -- 1991 to 2000 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 |
| Rate | State | 83.5 | 82.0 | 80.1 | 78.1 | 76.3 | 74.7 | 72.2 | 71.2 | 70.5 | 71.9 |
| | County | 71.4 | 69.9 | 68.2 | 67.3 | 66.8 | 66.3 | 65.8 | 65.8 | 64.6 | 63.6 |
| County | Lower CI | 68.2 | 66.8 | 65.2 | 64.3 | 63.9 | 63.4 | 62.9 | 62.9 | 61.8 | 60.8 |
| 95% CI | Upper CI | 74.5 | 72.9 | 71.2 | 70.3 | 69.7 | 69.2 | 68.7 | 68.7 | 67.4 | 66.3 |

At this point, it would be helpful to know two things: First, does the county have a statistically significant linear trend? Second, if it does, is the county trend similar to or different from what is happening at the state level?

Using the Trend Template in Appendix A and the Excel methods described in Appendix B, Figure 1 summarizes the results to answer these questions. Instead of a line connecting the year-to-year rates, we show regression lines for the annual county and state rates. Trend analysis focuses on the standard error rather than the CI. In plotting trends, neither CIs nor standard errors are shown. However, results of the statistical tests are summarized either in the plot or in the accompanying text. Finally, we omitted most horizontal lines and all vertical lines to maximize "white space." These visual aids enhance the ability of viewers to focus on the trends.

Figure 1: Fertility Rate per 1,000 Women
Age 15 to 44 -- 1991 to 2000
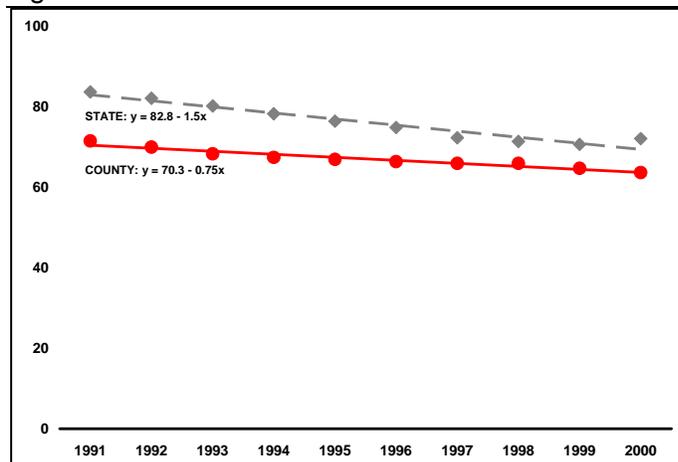


First notice that the regression line is slightly above or slightly below the center for most data points. That line represents the "best fit" for the data we have. The regressions tested if the line slope is significantly different from zero. We put the "best fit" formula of each test below its regression line. This does not show the statistical significance of the results.

In the County, the rate in a given year (y) dropped -0.76 (the slope) from the average rate over the 10 years of 70.3 per 1,000 women (the intercept). The trend template indicates that the slope is significantly different from zero (that is, we have a linear trend) at a P-value <0.0000. Since this P-value is less than 0.05, we can answer our first question: Do we have a statistically significant linear trend? The county fertility rate is declining at the statistically significant rate of about -0.75 annually from an intercept rate in 1991 of 71.1 per 1,000 women age 15 to 44.

Now let's answer the second question: Does the county trend differ from the state? We start by testing if the state trend is significantly different from zero. Starting from a rate in 1991 of 82.8 (the intercept) the state rate is decreasing annually about -1.5 per year (the slope), and the trend is statistically significant at P<0.0000. Thus, we have two statistically significant trends, one in the county and another in the state.

At this point, we can test if the trends are similar. That is, is the difference between the state slope (-1.5) and the county slope (-0.75) statistically significant. The trend template calculated that the difference between these slopes (0.75) has a P-value of 0.0002. This indicates that the state fertility rate is dropping faster than the county fertility rate.

From these analyses, we can conclude the following: The fertility rate for females age 15 to 44 has been declining during the period 1991 to 2000, statewide and in this county. Over this period, the rate of change (the slope) was smaller for the county than the state. The county upper confidence interval (Table 1) was below the state rate in every year, indicating the county rate was lower throughout this period.

# EXAMPLE 2: A CHANGE IN THE INTERCEPT

In this example, our county knew for many years that its percent of infants born weighing less than 2,500 grams was higher than the state average. In 1994, they decided to do an outreach program to increase enrollment in WIC programs for low-income women. Figure 2 below compares the percent of infants born who weighed less than 2,500 grams for the county and state for the 10-year period 1991-2000.

Figure 2: Newborn birthweight less than 2,500 grams (%), 1991-2000



COUNTY: y = 7.5 - 0.13x

STATE: y = 5.9 + 0.04x
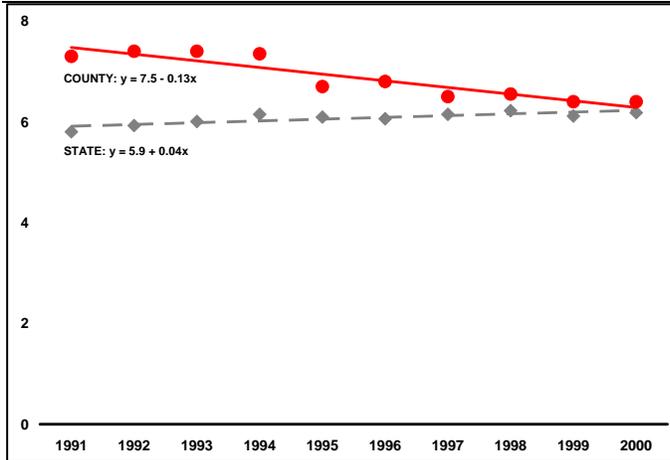
Treating the county line as if it was an uninterrupted trend, regression results show that the county average rate at the start of the period (intercept) was 7.5 and the annual rate of change (slope) was -0.13. The trend template calculated a significant trend (P=0.0001). The state regression results indicate the percent of low birthweight infants increased slightly at 0.04 per year from an intercept of 5.9%. Again the trend template indicates the state increase was statistically significant (P = 0.0028), as was the test for the difference of slopes (P<0.0000).

Trend analysis results reflected in the county formula are statistically significant, and highlight the county's accomplishments. By the end of the period, the county CI overlapped the state rate, indicating the WIC intervention was associated with a change in the county percent of low birthweights into the "average" range. However, the state and county are going in opposite directions. If current trends continue at both levels, the county percent of low birth weights will be below the state average in a few years. However, simple linear regression as used in this example probably is not the appropriate test for these data.

Notice how poorly the county line fits the observed data points. In fact, the county rate was stable from 1990 to 1994, dropped near the state rate when the program was implemented, and since 1995 has not been significantly different from the state rate. In this situation, it would be best to speak with a statistician to obtain a test for difference in the intercept.

# EXAMPLE 3: CURVILINEAR TRENDS

Next, we examine trends for the fertility rate per 1,000 girls age 10 to 14 between 1991 and 2000. As before, Figure 3A shows the observed data points, regression lines and model results using the linear trend template. Notice that both trend lines slope down, indicating the possibility of a trend for the fertility rate to decrease. However, the linear trend test was statistically non-significant for both the state (p = 0.40) and county (p = 30).

## Figure 3A: Fertility Rate per 1,000 Girls Age 10 to 14 -- 1991 to 2000 Linear Trend



STATE: y = 38.8 -0.7478x

COUNTY: y = -0.1973x + 13.201
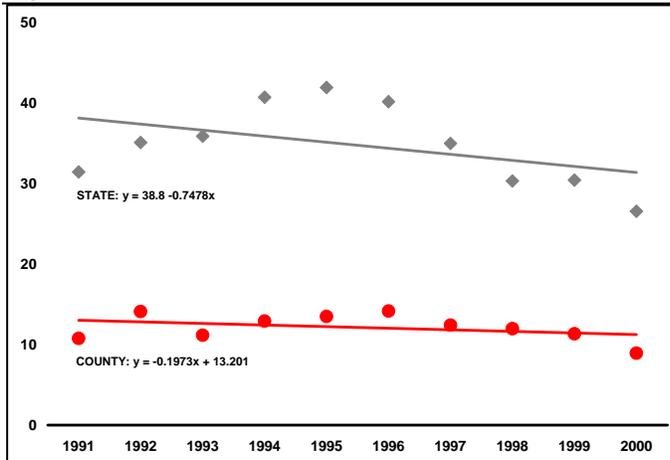
If the straight line is a "good fit", the data points would form no particular pattern relative to the line. Some points would be above the line, some below, in a random pattern as in Example 1. If the straight line is a "bad fit", you usually see a clustered rather than a random pattern. The points above the line would cluster near each other; points below the line also would cluster near each other. Sometimes there may be two clusters above or below the line.

As an example, look more closely at the state data points in Figure 3A, relative to the linear trend line. The 1994, 1995, 1996 and 1997 data points are above the line. Data points below the line are in two clusters (years 1991-1993 and 1998-2000). Clearly in this example, you have a clustered, not random, pattern. You would conclude that the linear model is a "bad fit." In this situation we recommend that you obtain a statistical consultation.

## Figure 3B: Fertility Rate per 1,000 Girls Age 10 to 14 -- 1991 to 2000 Curvilinear



STATE: y =26.9 -0.5x$^2$ + 5.2x

COUNTY: y = 9.8 + -0.15x$^2$ + 1.5x
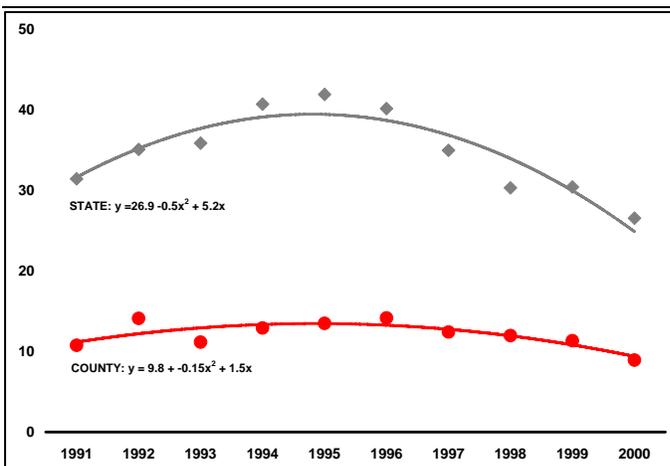
When your data are not explained with the simple linear model, you may find that more complex models fit better. These can be examined using the various alternatives on the "Type" Tab shown in Step 2 of Appendix B. Fitting a "Power" trend line to this data gives you the result shown in Figure 3B. These may make the visual understanding of your data more clear. However, the statistical interpretation of your results becomes more complex and is beyond the scope of this document.

During this 10-year period, both the county and state had what is known as a "polynomial" trend. That is, the best-fitting line is curvilinear, rising through 1995 and declining thereafter. Throughout the period, the county rate was lower, and its curve (-0.15) was flatter than the state curve (-0.5). Based on the R-squared, or the total amount of variance explained by the model (which Excel can give), both curves were statistically significant, when the linear trends were not. When you see this or any other type of complex trend, we recommend statistical consultation to be sure you are interpreting it correctly.

# EXAMPLE 4: ANOTHER COMPLEX TREND

In this example we are looking at mental health hospital admissions for adolescents age 15 to 19. Throughout the period, Figure 4 shows the state rate was relatively stable, and at P = 0.6060 (from the template), there was no trend. However, the county rates display as markedly curvilinear and the Excel regression line that appears to best fit the data is a third degree polynomial. Note that this regression result cannot be obtained with the trend template FHOP developed.

Figure 4: Psychiatric Admission Rate per 10,000
Adolescents 15-19 -- 1991 to 2000

COUNTY: $y = 111.91 - 0.29x^3 + 5.4x^2 - 25.34x$

STATE: $y = 82.6 - 0.27x$
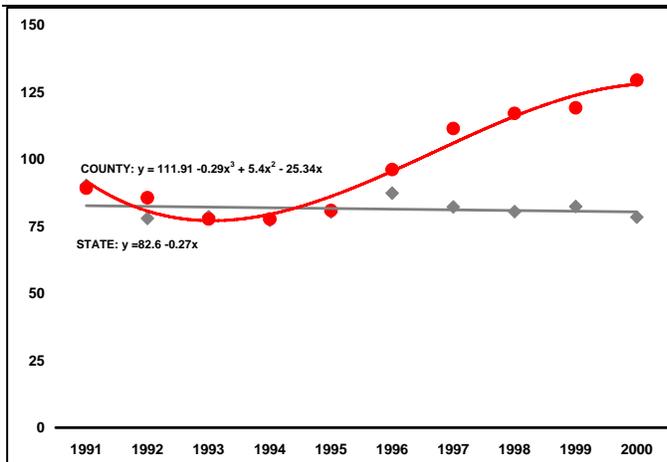
If you examine the data points closely, you will see that the county rate was declining through 1994. After 1995, it increased markedly and in a curvilinear pattern. To tease out this complexity, advanced methods are needed that we do not describe in this monograph. However, we identify this situation so you will know that if you see it, you may wish to seek help from a statistician to see if a spline model might fit or if you should use the National Cancer Institute's JoinPoint software to test these data.

# SUMMARY

This document introduced a few ways to look at linear trends. We started with the statistical concepts of intercept, slope, and probability. Then we showed a table with observed data points and their associated confidence intervals. These were analyzed using the FHOP data template, to test if the data had a statistically significant linear trend. Next, we showed the observed state and county points and regression lines on the same graph. We identified that a linear trend test was an appropriate method for these data. In the discussion for this section, we gave an example of how to compare and interpret results.

The complex curvilinear trend examples provided a few variations of how these look and how they may be interpreted. We wish to highlight that curvilinear trends occur frequently. Because of their frequency, we urge caution when complex curves present themselves, and recommend readers seek help from qualified statisticians in such circumstances. The only trend a non-statistician should be testing using the FHOP trend template is Example One, which was linear for both the county and state. If you see anything else, get a statistical consultation.

Regression can be a useful exploratory tool to understand trends. Exploratory estimates of the rate of change -- the county's intercept and slope – can be helpful in assessing simple trends. Properly interpreted, the results can provide unbiased estimates of the intercept and slope of the pattern of change.

# APPENDIX A: LINEAR TREND TEMPLATE

This shows the results produced by the linear trends template for the analysis of Example 1. The important statistics for purposes of this monograph are <mark>highlighted</mark>. The 12-year trend template is available under Research Tools on the FHOP website at http://www.ucsf.edu/fhop/. This template is inappropriate to use when trends are curvilinear.

## Linear Trend Template Results Example 1

| Year | Calendar | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 |
|------|----------|------|------|------|------|------|------|------|------|------|------|
|      | Start    | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |
| Rate | State    | 83.5 | 82.0 | 80.1 | 78.1 | 76.3 | 74.7 | 72.2 | 71.2 | 70.5 | 71.9 |
|      | County   | 71.4 | 69.9 | 68.2 | 67.3 | 66.8 | 66.3 | 65.8 | 65.8 | 64.6 | 63.6 |
| County | Lower CI | 68.2 | 66.8 | 65.2 | 64.3 | 63.9 | 63.4 | 62.9 | 62.9 | 61.8 | 60.8 |
| 95% CI | Upper CI | 74.5 | 72.9 | 71.2 | 70.3 | 69.7 | 69.2 | 68.7 | 68.7 | 67.4 | 66.3 |

|  |  | State | | County | |
|--|--|-------|--|--------|--|
|  |  | Estimate | Std. Err | Estimate | Std. Err |
| Regression | Intercept | <mark>82.8</mark> | 0.7 | <mark>70.3</mark> | 0.4 |
| Statistics | Slope | <mark>(1.5)</mark> | 0.1 | <mark>(0.8)</mark> | 0.1 |
|  | R2 | 0.94 | | 0.93 | |
|  | F | 118.2 | | 114.9 | |
|  | SE v | 1.3 | | 0.4 | |
|  | DF | 8 | | 8 | |
|  | SS reg | 186.6 | | 46.4 | |
|  | SS resid | 12.6 | | 3.2 | |
| T-Test H0 | T-stat | (10.87) | | (10.72) | |
| Slope = 0 | Prob T | <mark>0.0000</mark> | | <mark>0.0000</mark> | |
|  | Reject | YES | | YES | |
| T-Test H0 | Difference | (0.75) | | | |
| County Trend | Std Err Diff | 0.16 | | | |
| = State Trend | T-Test | (4.86) | | | |
|  | Prob T Diff | <mark>0.0002</mark> | | | |
|  | Reject | YES | | | |

# APPENDIX B:
# USING EXCEL TO OBTAIN LINEAR TREND RESULTS

This example begins when the reader has data organized into a linear graph. As our example, we use county data in the linear trend template. This and subsequent figures are cropped to focus on the point being made.

1. Click on chart, then click on Add Trendline
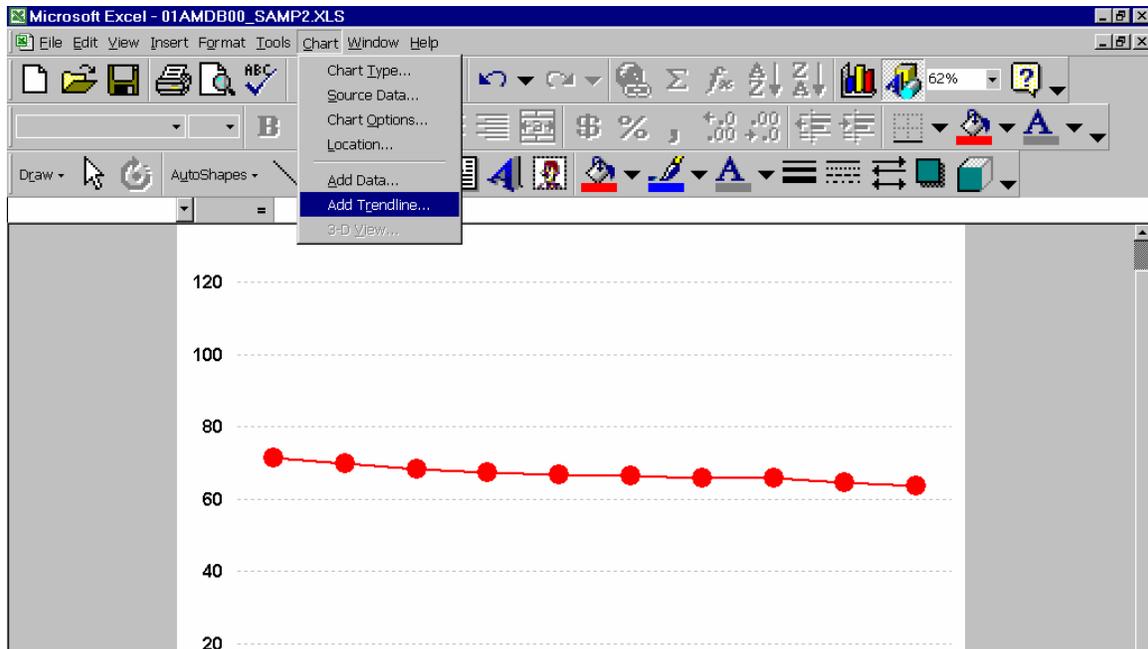
2. Pick the kind of trendline you want to test. In this case, because we think we have a simple linear trend, we picked the first display option. In other examples, we got our best results with other trends, which we tested in turn until we found the "best fitting" model. The "best fitting" model has the highest $R^2$. If two models have similar $R^2$, pick the simpler model.
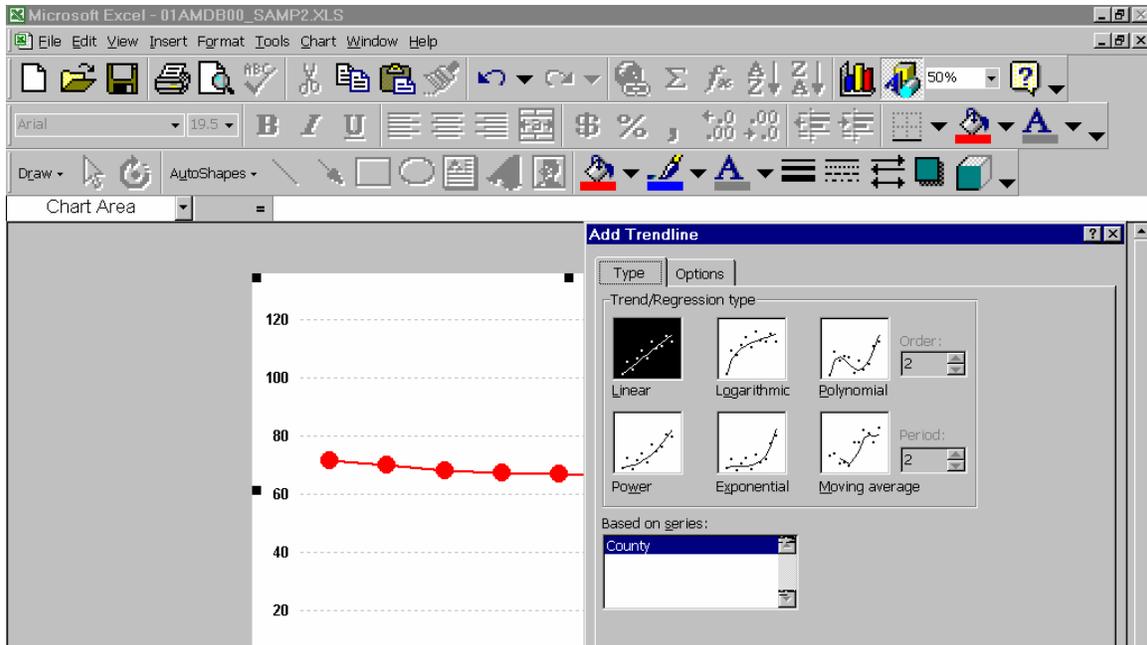


3. Now click on the Options Tab. Select "Automatic" trendline, "Display Equation on Chart" and "Display R-squared value on chart."
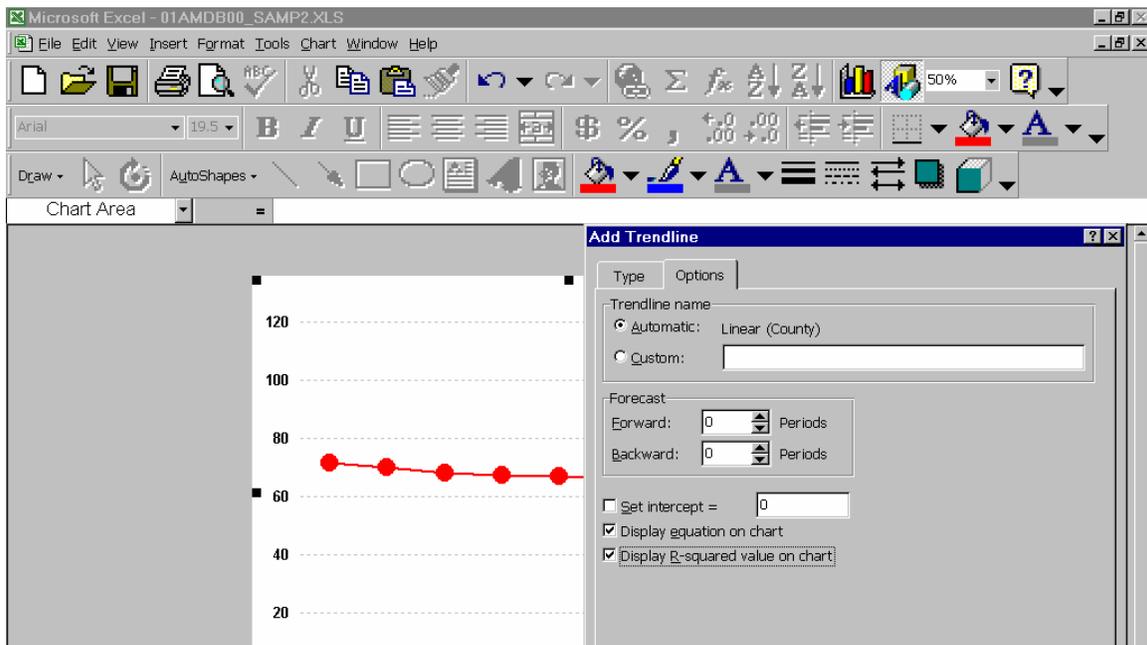
4. This generated a picture with a trendline and the regression results that looks a bit messy. Because the county rate is consistent from year to year, the trendline and the regression results are overlaid on the line connecting the rates. Let's fix this up.
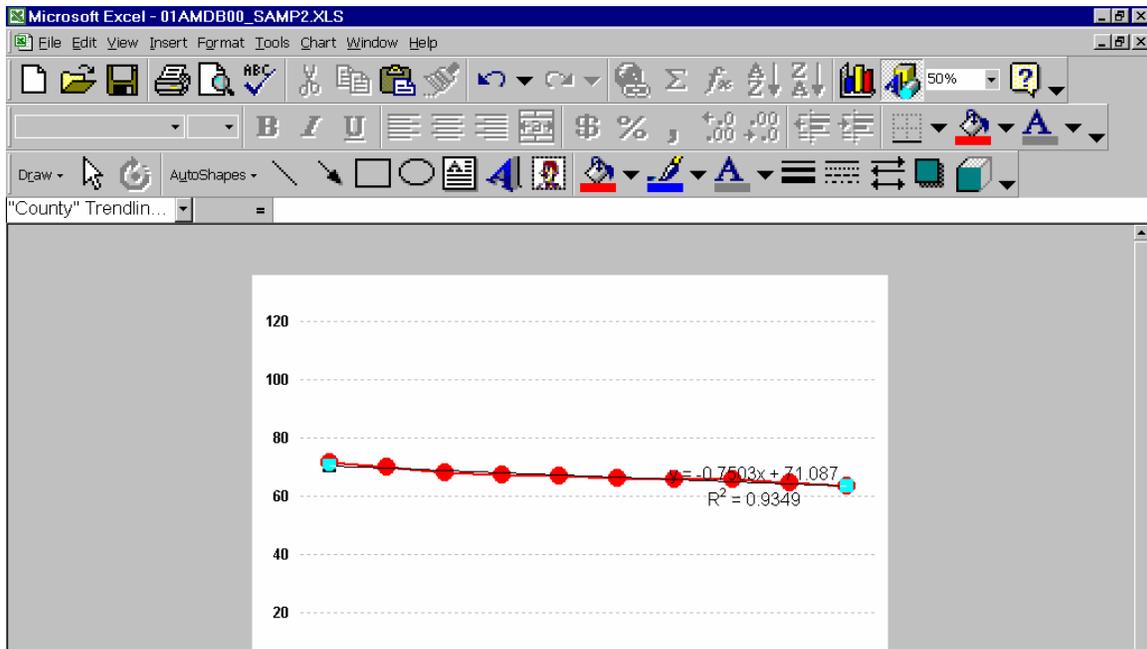


5. In this case, we want to keep the results of the regression, but we do not want it to overlay the graph line. Select the box that contains the equation and move it to the location you want.
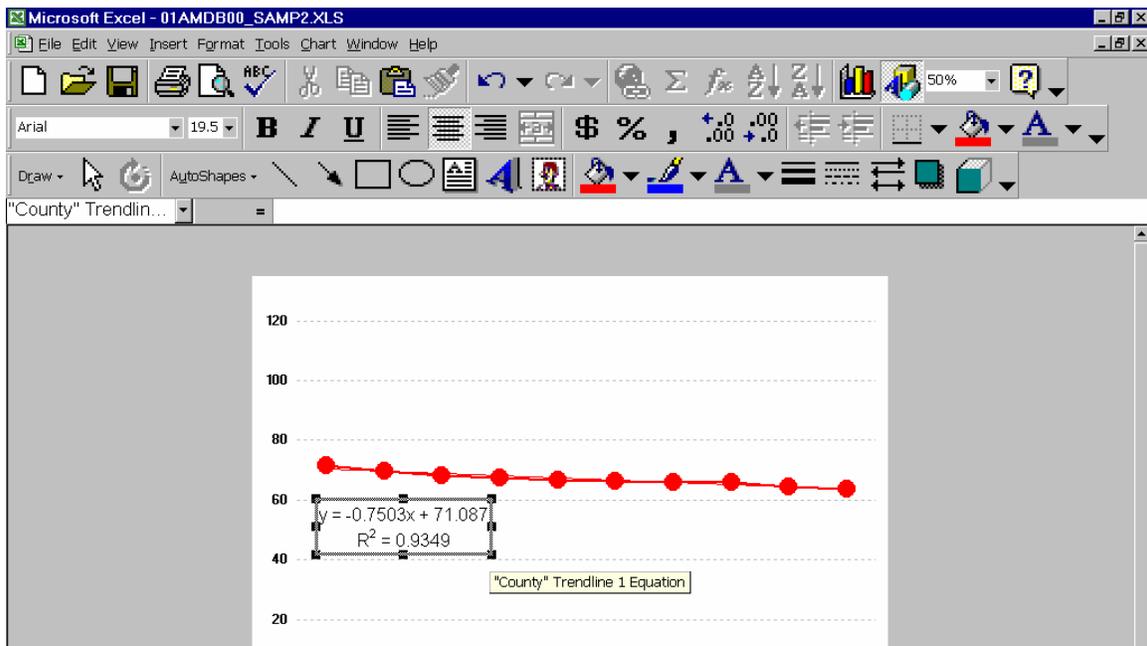
6. You can make the equation a bit more readable by reducing the number of digits that show. In general, there is no need to display more than 1 or 2 decimal points. Remember to round figures when you shorten the displays. Also, you can move the $R^2$ statistic to the same line as the slope and intercept.
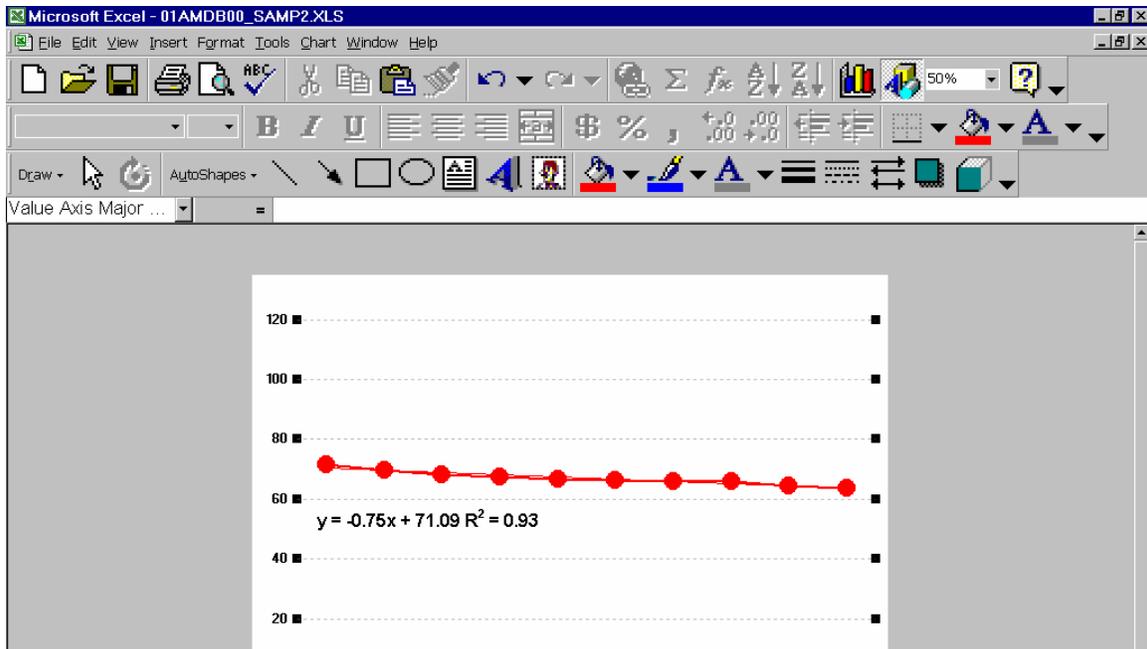
# APPENDIX C: RESOURCES

Healthy People 2010. Available at: http://www.healthypeople.gov. At this site, you will find the national health goals and objectives.

Tufte ER. The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press. 1983. This is the first in a series of three books Tufte wrote that revolutionized how we think about the presentation of data. http://www.edwardtufte.com/tufte/.

Oreglia A: Public Health Rate Program (PHRATE), Version 1.1. Vital Statistics Section, Center for Health Statistics, California Department of Health Services: May 1993. This public release program available for $10 enables an analyst to calculate various simple statistics: dhs.ca.gov/hisp/chs/ohir/products/forms/phrate.pdf.

Singer JD and Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford University Press, 2003. This excellent book for statisticians describes the theory and shows methods underlying longitudinal data analysis. The authors have released programs written to duplicate examples presented in their book for the following statistical software programs: MPlus, MLWin, HLM, SAS, STATA, SPlus, and SPSS. These programs are at: http://www.ats.ucla.edu/stat/stata/examples/alda/.

Example 4 showed an example where the data were best described using two intersecting straight lines rather than a quadratic or higher-degree polynomial, in order to capture the exact point (also called "bend" or "knot") where the trend changed. This type of model is called a "linear spline model," or a first-degree polynomial (straight lines) model. While a single straight line can be fitted using ordinary least squares models, fitting splines requires special methods.

To do Example 4, FHOP staff used an algorithm contained in a macro we developed using SAS. To obtain a copy, please send email to lremy@well.com.

JoinPoint Regression Program. The National Cancer Institute offers this easy-to-use statistical freeware to model longitudinal trends using spline methods. After a simple registration procedure, it can be downloaded from: http://srab.cancer.gov/joinpoint/.

NCI has developed other statistical software to analyze complex trends as described at http://srab.cancer.gov/software/software.html. These various freeware packages can be used for a variety of public health trend analysis needs.

The following are some references showing the use and interpretation of splines. PDF versions of these and other related papers can be obtained by sending email to lremy@well.com.

Pollio DE, Spitznagel EL, North CS, Thompson S, Foster DA. Service use over time and achievement of stable housing in a mentally ill homeless population. Psychiatric Services, Dec. 2000, 51(12), 1536-1543. This article shows the use of spline trends with a small sample (n = 55) followed for 24 consecutive months.

Kim S, Kim K. Personal, temporal and spatial characteristics of seriously injured crash-involved seat belt non-users in Hawaii. Accid Anal Prev. 2003 Jan; 35(1):121-30. This study used police crash data over a 10-year period.

MacNab YC. A Bayesian hierarchical model for accident and injury surveillance. Accid Anal Prev. 2003 Jan; 35(1):91-102. This studied trends for motor vehicle accident injuries to boys aged 0-24 between 1987 and 1996 in British Columbia.