Statistical Analysis of Crime Data Using Time Series and Correlation Techniques

By

Megan Stokes

Advisor: Dr. Morris Marx

An Undergraduate Proseminar
In Partial Fulfillment of the Degree of
Bachelor of Science in Mathematical Sciences
The University of West Florida
April 2011

The Proseminar of Megan Stokes is approved:

Morris Marx, Ph.D., Proseminar Advisor — Date

Josaphat Uvah, Ph.D., Proseminar Committee Chair — Date

Accepted for the Department:

Kuiyuan Li, Ph.D., Chair — Date

TABLE OF CONTENTS

Statistical Analysis

By

Megan Stokes

ABSTRACT

In this paper, we explore the theory behind time series analysis, including the different components of a time series: seasonal adjustment, smoothing methods, correlation, random variation, and limitations. Numerous real data sets illustrate the main concepts. We then use this form of analysis to interpret data pertaining to crime and unemployment in the United States of America.

CHAPTER I. INTRODUCTION

A. Statement of Problem

What is statistics? The word “statistics” first appeared in 1749 [1]. Coined by Gottfried Achenwall, a German scholar, “Statswissenschaft” originally referred to the use of demographic information to compare one country with another [1]. Since the term “statistics” was first coined, statistics have been used to explain real-life phenomena. Before long, the definition broadened, and the word came to have two quite different meanings: As a plural noun, statistics is now commonly used to describe any set of data, whatever the source;as a singular noun, statistics refers to a subject – specifically, to a set of mathematically based procedures for collecting, summarizing, and interpreting data [1]. When used in this paper, the term “statistics” will refer to the latter definition, and when referring to the first definition we will say “data set(s).”

Why should we study statistics today? To answer this question all one has to do is turn on a television. The news media and research journals present statistical reports and forecasts concerning:

- Weather forecasts
    - These forecasts are based on models that are built using statistics that compare prior weather conditions with current weather conditions.
- Medical studies
    - Research and testing are necessary to help companies obtain approval for their products from the Food and Drug Administration.

- Politics
    - Whenever there is an election, the news organizations use statistical models to see how a candidate is doing at the time and to predict who is going to win the election.

These are just a few examples of statistics being put to use in everyday life.

Simply stated, the study of statistics is important because it explains what has happened in the past, what is happening in the present, and what may happen in our future. All one has to do is open a statistics textbook and there will be countless examples of real-world problems. The preface to the text *Statistics for Applied Problem Solving and Decision Making* states [1]:

> *Many of the examples are drawn from recent real-life situations in business, finance, management and the sciences.*

In a later book, published by the same authors, they commented on using real-world case studies and historical anecdotes in their statistics work [2]:

> *Our experience in the classroom has strengthened our belief in this approach. Students can better grasp the importance of each area when seen in the context of the other two*

"Real-world problems seldom have simple statements, nor do they lend themselves to clear-cut, straightforward answers" [1]. For this reason, the field of statistics is becoming more sought after, the Bureau of Labor Statistics predicts that [3]:

> *Employment of statistics is projected to grow 14 percent from 2010 to 2020… Growth will result from more widespread use of statistical analysis to make informed decisions. In addition, the large increase in available data from the Internet will open up new areas for analysis.*

These statisticians are expected to perform the following tasks [4]:

- Determine the questions or problems to be addressed
- Decide what data are needed to answer the questions or problems
- Determine the methods for finding or collecting the data
- Design surveys or experiments or opinion polls to collect data
- Collect data or train others to do so
- Analyze and interpret data
- Report conclusions from their analyses

Most people recognize the importance of being able to understand the world around them, and study of statistics is the key to understanding real-world problems.

B. Relevance of Problem

In September of 2009, the Federal Bureau of Investigation made a national press release making the following statement [5]:

> *Unemployment, interest rates, stress—they're all on the rise as the economy is buffeted by a downturn. What's not rising, however, is crime, according to statistics compiled by the FBI that show violent crimes and property crimes declined nationwide in 2008.*

The next sentence of this statement began "The statistics show that…" [5]. That same year at a news briefing when asked President Barack Obama said "real lives, real suffering, real fears" were what was really behind the latest job statistics released concerning the weakening economy of the United States [6]. In February of 2012 the U.S. Department of Labor made the following statement in an article comparing the most recent recession with prior recessions [7]:

> *A general slowdown in economic activity, a downturn in the business cycle, a reduction in the amount of goods and services produced and sold—these are all characteristics of a recession. According to the National Bureau of Economic Research (the official arbiter of U.S. recessions), there were 10 recessions between 1948 and 2011. The most recent recession began in December 2007 and ended in June 2009, though many of the statistics that describe the U.S. economy have yet to return to their pre-recession values.*

What do these two statements have in common? For one, they are both statements made based on analyzing statistics, and two, they are both speaking of the same time period. A conversation with James Parker, the supervising State Attorney for the Santa Rosa County branch, revealed his opinion on crime in the United States: “Crime is not decreasing, from what I have seen over the last few years I would say crime is increasing not just in this area but on a national scale.” Given Mr. Parker’s opinion, the question arose about why the media portrayed one thing and yet he has witnessed the opposite.

This study used data three sets concerning the United States. The first of the three includes annual crime rates for violent crime and property crime, which together are composed of nine different offenses, and the population size of the United States. The second data set lists the unemployment rates for every month, and the last set lists the number of full-time law enforcement officers. The first data set lists the annual volume of crimes and the rate per 100,000 inhabitants from the year 1975 to 2010. Nearly 17,800 city, county, college and university, state, tribal, and federal agencies participated in the UCR Program in 2008 [5]. These agencies represent 94.9 percent of the nation’s population [5]. The unemployment data set lists the monthly unemployment level (in the thousands) and the annual unemployment rates for the years between 1947 and 2011. The last set lists the number of full-time law enforcement officers for the years 1975 to 2010. The data sets are listed in Appendix 1.1.

C. Literature Review

The FBI's Uniform Crime Reporting Program (UCR) collects offenses that come to the attention of law enforcement for violent crime and property crime, as well as data regarding clearances of these offenses [8]. In the FBI's UCR Program, law enforcement agencies can clear, or "close" offenses in one of two ways: by arrest or by exceptional means [9]. To clear a crime by arrest three specific conditions must be met [9]:

1. Arrested.
2. Charged with the commission of the offense.
3. Turned over to the court for prosecution (whether following arrest, court summons, or police notice).

The arrest of one person may clear several crimes, and the arrest of many persons may clear only one offense [9].

To clear a crime by exceptional means the law enforcement agency must meet the following four conditions [9]:

1. Identified the offender.
2. Gathered enough evidence to support an arrest, make a charge, and turn over the offender to the court for prosecution.
3. Identified the offender's exact location so that the suspect could be taken into custody immediately.
4. Encountered a circumstance outside the control of law enforcement that prohibits the agency from arresting, charging, and prosecuting the offender.

An example of an offense being cleared by an exception means include, but are not limited to, the death of the offender; the victim's refusal to cooperate with the prosecution after the offender has been identified; or the denial of extradition because the offender committed a crime in another jurisdiction and is being prosecuted for that offense [9].

In the 1920's, the International Association of Chiefs of Police (IACP) recognized the potential value in tracking national crime statistics [10]. The Committee on Uniform Crime Records of the IACP developed and initiated this voluntary national data collection effort in 1930 and still continues to advise the FBI on the UCR Program progress [10]. During that same year, the IACP was instrumental in gaining Congressional approval which authorized the FBI to serve as the national clearinghouse for statistical information on crime [10]. Since 1930, through the UCR Program, the FBI has collected and compiled data to use in law enforcement administration, operation, and management, as well as to indicate fluctuations in the level of crime in America [10]. The crimes in the UCR program have been chosen because of their seriousness, frequency of occurrence and likelihood of being reported to the police [11]. The offenses presented in the FBI database reflect the Hierarchy Rule [8] which states [12]:

> *There is a significance to the order in which the Part I offenses are presented, with criminal homicide being the highest in the hierarchy and arson being the lowest. The experience of law enforcement agencies in handling UCR data shows that, for the most part, offenses of law occur singly as opposed to many being committed simultaneously. In these single-offense situations, law enforcement agencies must decide whether the crime is a Part I offense. If so, the agency must score the crime accordingly. However, if several offenses are committed at the same time and place by a person or a group of persons, a different approach must be used in classifying and scoring. The law enforcement matter in which*

> *many crimes are committed simultaneously is called a multiple-offense situation by the UCR Program. As a general rule, a multiple-offense situation requires classifying each of the offenses occurring and determining which of them are Part I crimes. The Hierarchy Rule requires that when more than one Part I offense is classified, the law enforcement agency must locate the offense that is highest on the hierarchy list and score that offense involved and not the other offense(s) in the multiple-offense situation. The Hierarchy Rule applies only to crime reporting and does not affect the number of charges for which the defendant may be prosecuted in the courts. The offenses of justifiable homicide, motor vehicle theft, and arson are exceptions to the Hierarchy Rule.*

For example, if one were to enter a store, rob eight customers, and then kill the cashier, only the homicide charge would be reported to the UCR Program. Also, arson was not originally part of the crime reporting process [10]. Arson became the eighth Index crime as the result of a limited Congressional mandate in October 1978 [10]. The Part I offenses are listed in Appendix 1.2.

D. Limitations

The purpose of this study is not to prove/disprove the claims of the statements discussed but merely to discuss how one might determine if these claims are true or not and to consider different methods of statistical analysis. With that in mind, our limitations include that we only have access to the data that the government releases to the public and therefore has been gathered by others. Thus, we rely wholly on the law enforcement and employment offices reporting correct totals. Assuming that the law enforcement offices are reporting all data, it should be noted that participation in the National UCR Program is strictly voluntary [10]. It should also be noted that the findings of a court, coroner, jury, or the decision of a prosecutor are not recorded.

As said earlier, the UCR Program has been functioning since 1930, since then several experts have argued with the programs level of accuracy. Dr. Dean Kilpatrick in a statement to the Senate Committee regarding the reporting of rape cases had this to say about the UCR Program and another program, NCVS, designed specifically to determine estimate the total number of crimes that occur each year in the U.S. including those not reported to law enforcement [13]:

> *Both the NCVS and the UCR have major flaws that result in their being poor tools for measuring rape cases that produce serious underestimates of the total number of unreported and reported rape cases that occur each year…..the bottom line is that the problems with both measures are so serious that they are incapable of providing us with the data needed to determine the proportion of all rape cases that are reported to police as measured by the NCVS or the disposition of those cases reported to police as measured by the UCR…..Congress should demand that changes are*

> *made in the UCR and NCVS to fix this problem so these measures can give us the information we need to determine whether we are making progress in addressing our rape problem.*

In another journal by this same doctor, discussing the accuracy the FBI's figures, he says [14]:

> *"Unfounded" cases, or cases that—according to federal reporting requirements—are presumed to be false or baseless upon investigation, are not included in the UCR totals….reports to authorities that are considered unfounded following investigation also are not counted.*

However, with no way of determining the accurate number of crimes ourselves, the UCR figures will be used in this paper.

CHAPTER II. A TIME SERIES TUTORIAL

The objective of statistics is to make inferences about a large body of data based on information contained in a sample from that population [15]. In other words, "statistics" allow us to interpret complex data. As stated earlier, the topic of crime in America intrigued me and as I began gathering data on crime in America, it became apparent that a time series analysis would be appropriate. In order to obtain a better understanding of the data, one needed to be able to understand the general notion of a time series and the pertinent components in this particular time series. Regression models are not reliable when a data set has variables that are correlated over time, the prediction errors of the regression model will be correlated violating the assumption of independence [16]. The solution to this problem is to construct a time series model [16].

A. What is a Time Series?

In mathematics, a time series has been described as "…a sequence of observations ordered by a time parameter" [17]. Given observations where the distinguishing feature is the nature of the independent variable, if the x_i's represent time, we call the n observations $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$ a time series [1]. Typically the time units are days, months, quarters, or years and the x_i's are evenly spaced. A time series is usually represented by the mathematical equation listing the values of the response as a function of time or, equivalently, as a figure on a graph whose vertical coordinate gives the value of the random response plotted against time on the horizontal axis [15]. Figure 2.1 is an example of a time series where the U.S. treasury bill rates are along the vertical axis and the time in years is along the horizontal axis [15]:

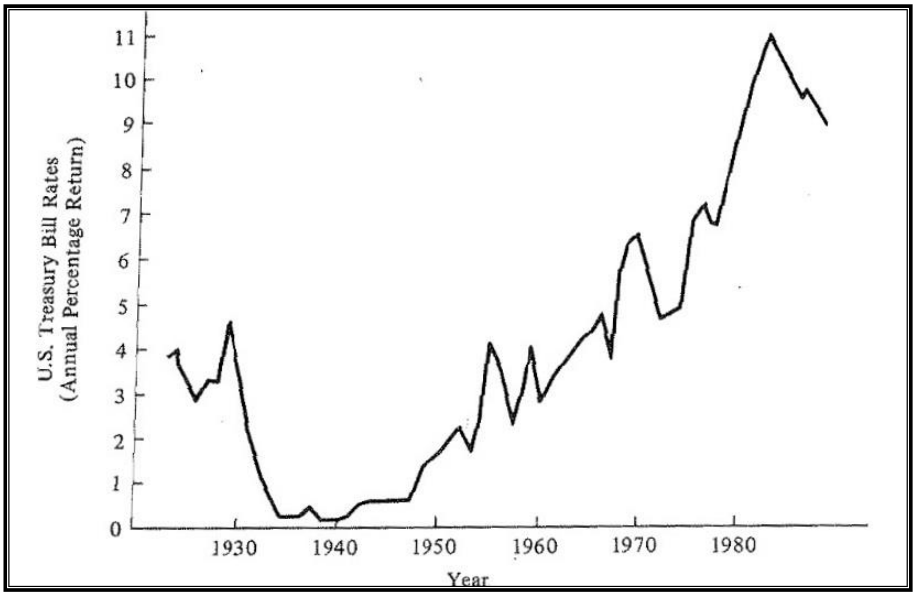


Figure 2.1

Yaffee and McGee tell us that a time series may be measured continuously or discretely [18]. Continuous time series are recorded instantaneously and steadily; for example consider an oscillograph recording harmonic oscillations of an audio amplifier [18]. Figure 2.2 is an example of a continuous time series [19]:

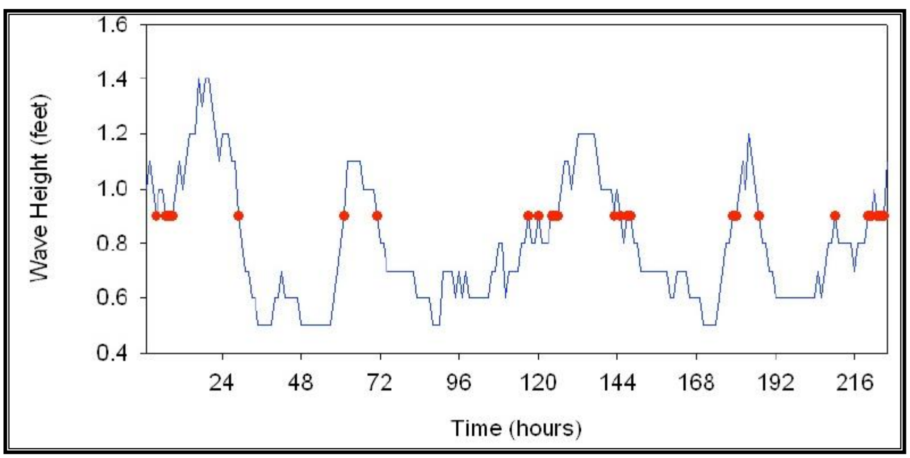


Figure 2.2

Most measurements in the social sciences are made at regular intervals, and these time series data are discrete [18]. Figure 2.3 is an example of a discrete time series [18]:

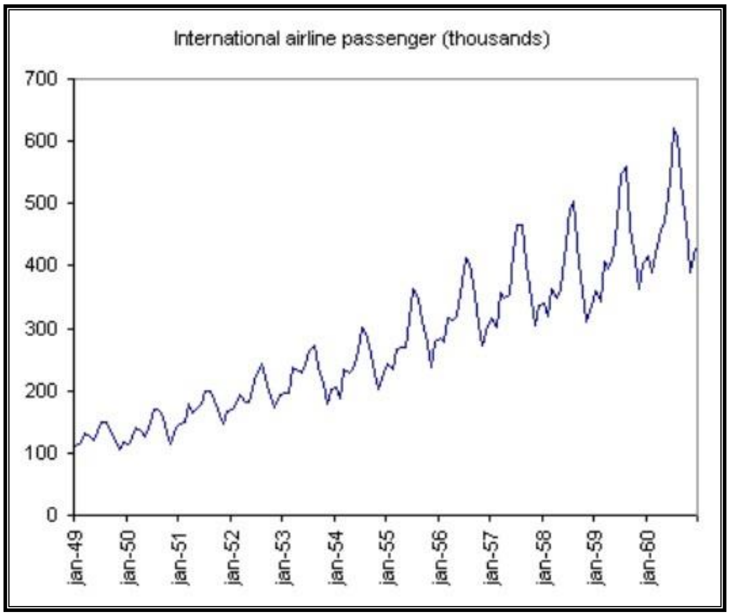


Figure 2.3

The data in this paper is for a discrete time series, where the observations have been recorded regularly every year.

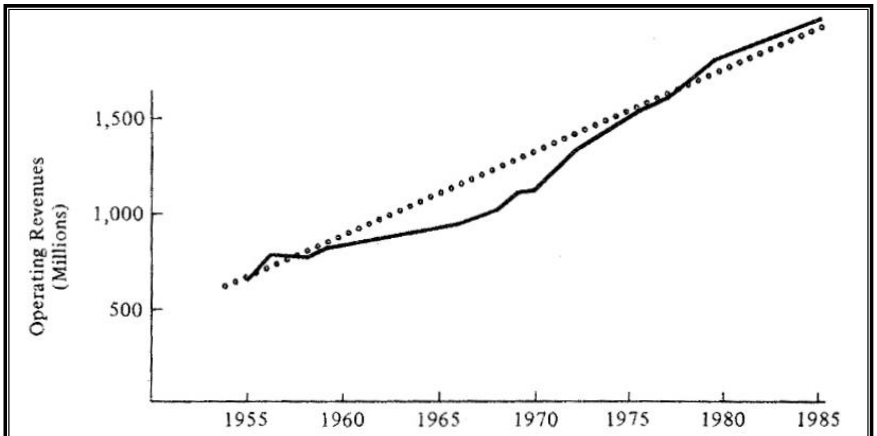
The standard time series model expresses Y_i as a sum of four components:

$$\bar{\mathbf{Y}}_{\mathbf{i}} = \mathbf{g}(\mathbf{x}_{\mathbf{i}}) + \mathbf{s}_{\mathbf{i}} + \mathbf{c}_{\mathbf{i}} + \boldsymbol{\varepsilon}_{\mathbf{i}} \qquad (2.1)$$

where $g(x_i)$ is the trend, s_i is the seasonal component, c_i is the cyclical component, and ϵ_i is a residual effect or a random variable[1]. It should be noted that all four components are not necessarily present – or prominent – in every set of time series data.

B. What are the Components of a Time Series?

A time series has four main components: long-term trends, cyclical effects, seasonal effects, and random variation. Denoted $g(x_i)$, the trend might be a familiar linear or exponential function, but it could be any expression that effectively describes the data's scatterplot [1]. Long-term trends are often present because of a steady increase in population, gross national product, the effect of competition, or other factors that fail to produce sudden changes in response but produce a steady and gradual change over time [15]. In practice, this means that we should like to represent it by a continuous function of time [20]. Trends are classified according to their type (discrete or continuous) and length [18]. Figure 2.4 is a time series with a long-term trend in which the operating revenues of a company are steadily increasing over time in a linear fashion [15]:



Operating Revenues of U.S. Oil Pipeline Companies from 1955 to 1985. Long-term Trend Indicated as Dotted Line

Figure 2.4

Cyclical effects, c_i, in a time series are seen when the response variable "rises and falls in a gentle, wavelike manner about a long-term trend curve." Cyclical effects are generally caused by changes in the demand for a product, by business cycles, and by the inability of supply to meet consumer's demands [15]. Typically, the location, duration, and amplitude of cyclic fluctuations are very difficult to predict [1]. Figure 2.5 is an example of a time series with a cyclical component in which two different data sets (UAH and RSS) are displayed that have recorded the rise and fall of sea surface temperature [21]:

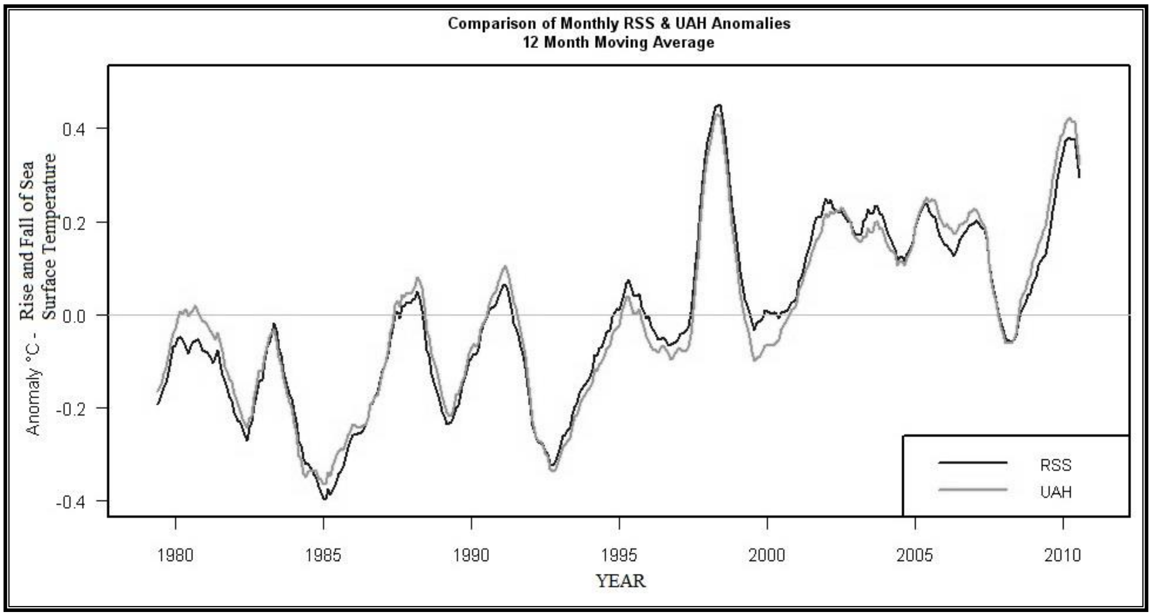


Figure 2.5

Seasonal effects, s_i, are the most commonly mentioned effect in a time series. Seasonal effects in a time series have been described as those rises and falls that always occur at a particular time of the year [15]. Seasonality may follow from yearly changes in weather such as temperature, humidity, or precipitation [18]. For example, propane gas consumption would most likely go down in the summer because of the temperature increase and need for heating decreasing, revenues made at a university cafeteria will be particularly low during the time when students are between semesters, both of these are two examples of when a seasonal effect would be present in a time series. Figure 2.6 is an example of a seasonal effect in a time series in which the unemployment rate of Northwest Florida is shown [22]:

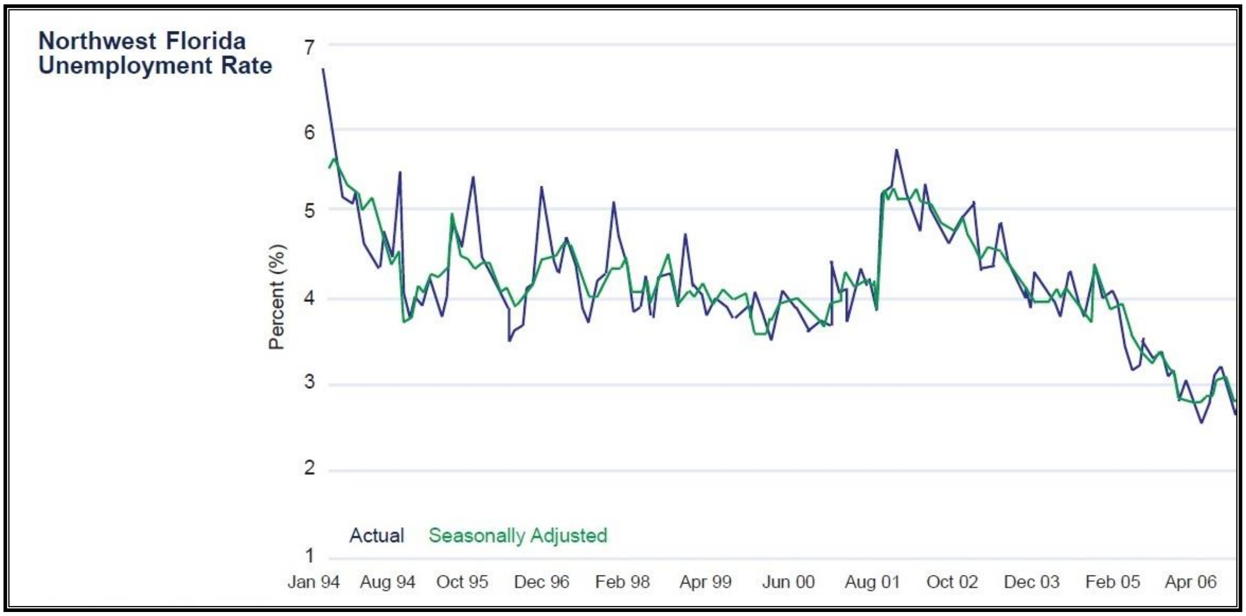


Figure 2.6

The essential difference between seasonal and cyclic effects is that seasonal effects are predictable, occurring at a given interval of time from the last occurrence, while cyclic effects are completely unpredictable [15].

Random variation (ε_i), or the residual component, of a time series represents the random upward and downward movement of the series after adjustment for the long-term trend, the cyclic effect, and the seasonal effect [15]. Suppose we fit the equation

$$\overline{Y}_i = g(x_i) + s_i + c_i + \varepsilon_i \qquad (2.1)$$

to a set of data and find that when i=15, $g(x_{15}) = 133.2$, $s_{15} = -8.5$, and $c_{15} = 3.2$. We would then "expect" $\overline{Y}_{15}$ to equal

$$\overline{Y}_{15} = g(x_{15}) + s_{15} + c_{15} + \varepsilon_{15} = 133.2 - 8.5 + 3.2 = 127.9$$

Suppose the *actual* Y_{15}, though, were *132.7*. The difference, 132.7 – 127.9 = 4.8 (= ε_{16}), is called the residual. It represents, in general, the net effect on *Y* of all factors other than the trend function and the seasonal and cyclical components [1]. Political events, weather, and an amalgation of many human actions tend to cause random and unexpected changes in a time series [15]. All time series contain random variation [15]. The long-term trend and seasonal effect, when identified, can be subtracted from the response values (x_i's), Figure 2.7, in which the monthly sales of a brewing company are depicted, is an example of such [15]:

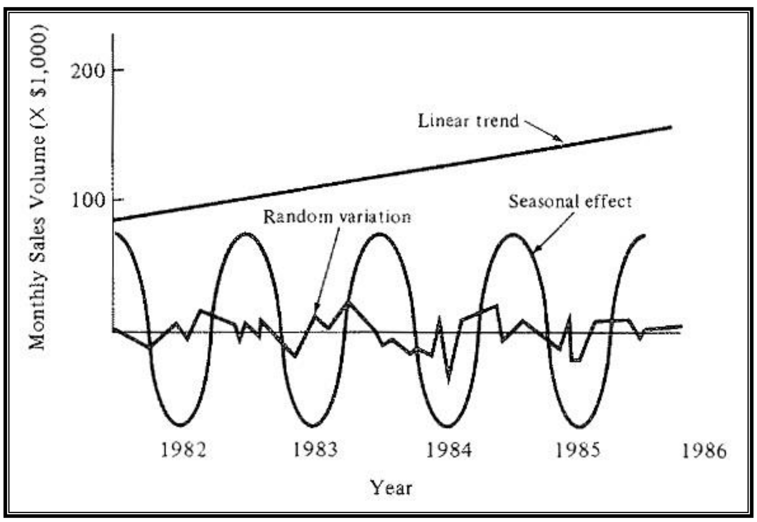


Figure 2.7

The data used in this study has both a linear trend and a seasonal component.

C. Why are Time Series Analyses Important?

Everyone has to make decisions and, usually, time is an important factor in those decisions. For example, when deciding where to invest money one should look at the history of the stock being considered, when looking to purchase a home, one might look at the crime rates or the schools for that area over the last several years. The data that one would need to look at to make these decisions would, hopefully, be arranged in a time series. Mendenhall states that a time series would be appropriate in the following situation [15]:

> *When attempting to estimate the expected value of a random process or to predict a new value at a future point in time after having observed the historical pattern.*

In other words, anyone that is interested in making plans for their future by budgeting time and resources (money, materials, etc..) is concerned with processes that have occurred over time, that is a time series, and what will happen in the future, that is forecasting.

D. Smoothing Methods

Sometimes it is necessary to use another indicator for determining the general movement in a time series other than using the trend function because the trend function can, at times, oversimplify the x,y relationship. An alternative strategy that avoids that shortcoming is a technique known as smoothing [1]. Smoothing techniques provide a method for canceling the effects of random variation in a time series in order to reveal the inherent components of the time series [15]. Like multiple regression analysis, smoothing techniques serve to assist in the *explanation* of a time series or the *prediction* of future realizations of the time series [15].

I. Moving Averages

The simplest smoothing technique is a moving average of the response measurements over a fixed number of time periods [15]. In the method of moving averages, a y_i in the original time series is replaced by an average of *m* (= 2k+1) points [1], each observation y_i is given equal weight [18]. If *m* is odd, the average is computed using y_i itself together with the $\frac{m-1}{2}$ data points that immediately precede y_i and the $\frac{m-1}{2}$ points that immediately follow y_i [1]. The (2*k* + 1)-*point moving average* of an *n*-observation time series replaces the data point y_j with the average

$$\bar{Y}_j = \frac{y_{j-k} + y_{j-k+1} + \cdots + y_{j-1} + y_j + y_{j+1} + \cdots + y_{j+k-1} + y_{j+k}}{2k+1} \qquad (2.2)$$

for each *j*, where $k < j \leq n - k$ [1] and y_j is the process response at response time *j*, y_{j-k} is the process response at response time *j* – *k*, and so forth [15]. The number of observations used for the computation of the mean (i.e., 2*k* + 1) is called the *order* of the series [18]. The net effect is to transform the original time series to a moving average series that is smoother (less subject to rapid oscillations) and more likely to reveal the underlying trend or cycles in the pattern over time [15].

One reason that moving averages are useful is the effect on the random or residual component. Suppose the time series has the general term $x_j + \varepsilon_j$ where ε_j is a random variable. Also suppose for simplicity that each ε_j has variance σ^2 and the ε_j are uncorrelated. Then averaging a span of *m* terms of the time series results in the quantity $\frac{1}{m}\sum_{j=1}^{m}(x_j + \varepsilon_j) = \frac{1}{m}\sum_{j=1}^{m} x_j + \frac{1}{m}\sum_{j=1}^{m} \varepsilon_j$.

Note that the variance of the second term is σ^2 / m, so the variability of the random component of the moving averaged is reduced.

Given the data in Appendix 2.1, and asked to calculate the simple moving average of order 5 for August 1983, the following steps would be performed:

1. Calculate the number of data points preceding (and following) August 1983 to use in the simple moving average

$$\frac{m-1}{2} = \frac{5-1}{2} = 2$$

   Hence, the two months before, and after, August 1983 will be used in the moving average

2. Sum these months

$$June_{83} + July_{83} + August_{83} + September_{83} + October_{83}$$

$$= 30.0 + 28.7 + 33.8 + 25.1 + 22.1$$

$$= 139.7$$

   Note: All of these months are for the year 1983.

3. Divide by the order

$$\bar{Y}_{Aug-83} = \frac{139.7}{5} = 27.9$$

In Figure 2.8 the original values are in blue and the moving average values are in red of a brewing company's sales between the years 1983-1985.

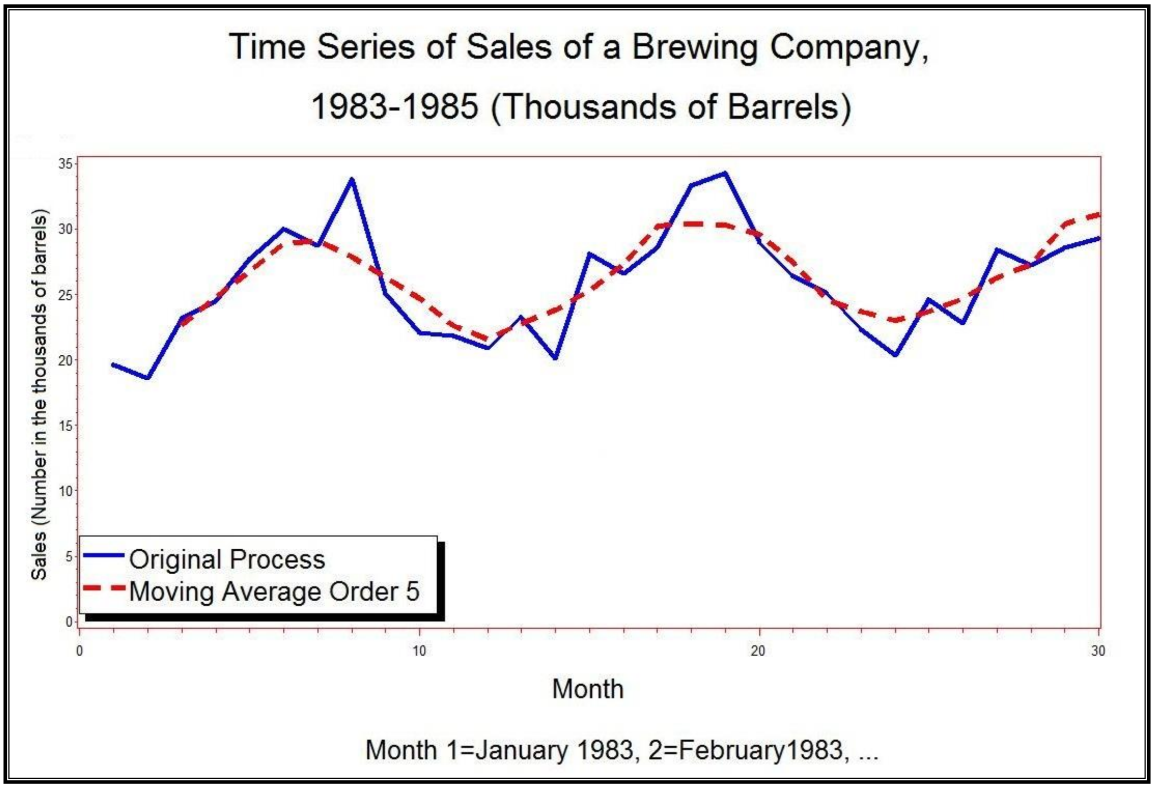
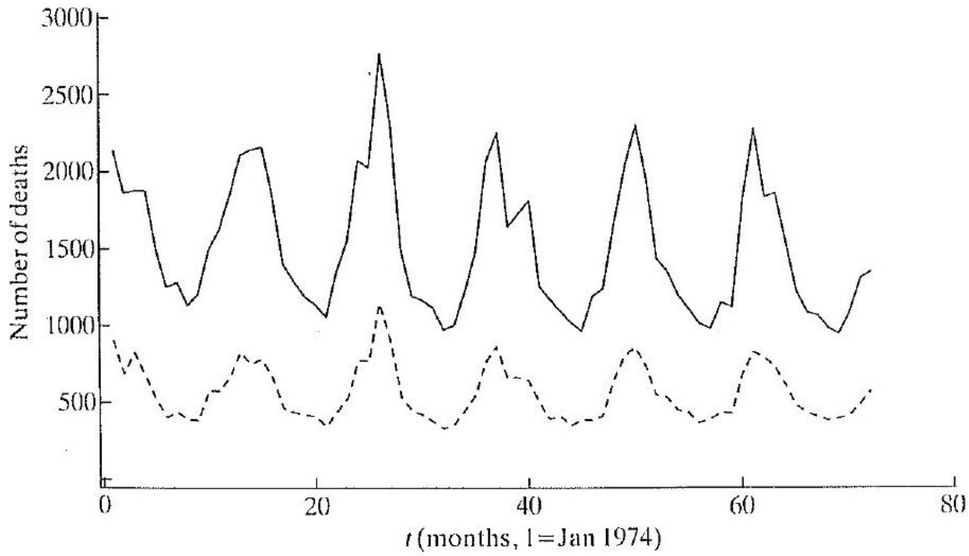


Figure 2.8

Notice, in the original process no time series component is evident, but when the data is smoothed using a moving average of order 5, there is a seasonal component revealed in the graph.

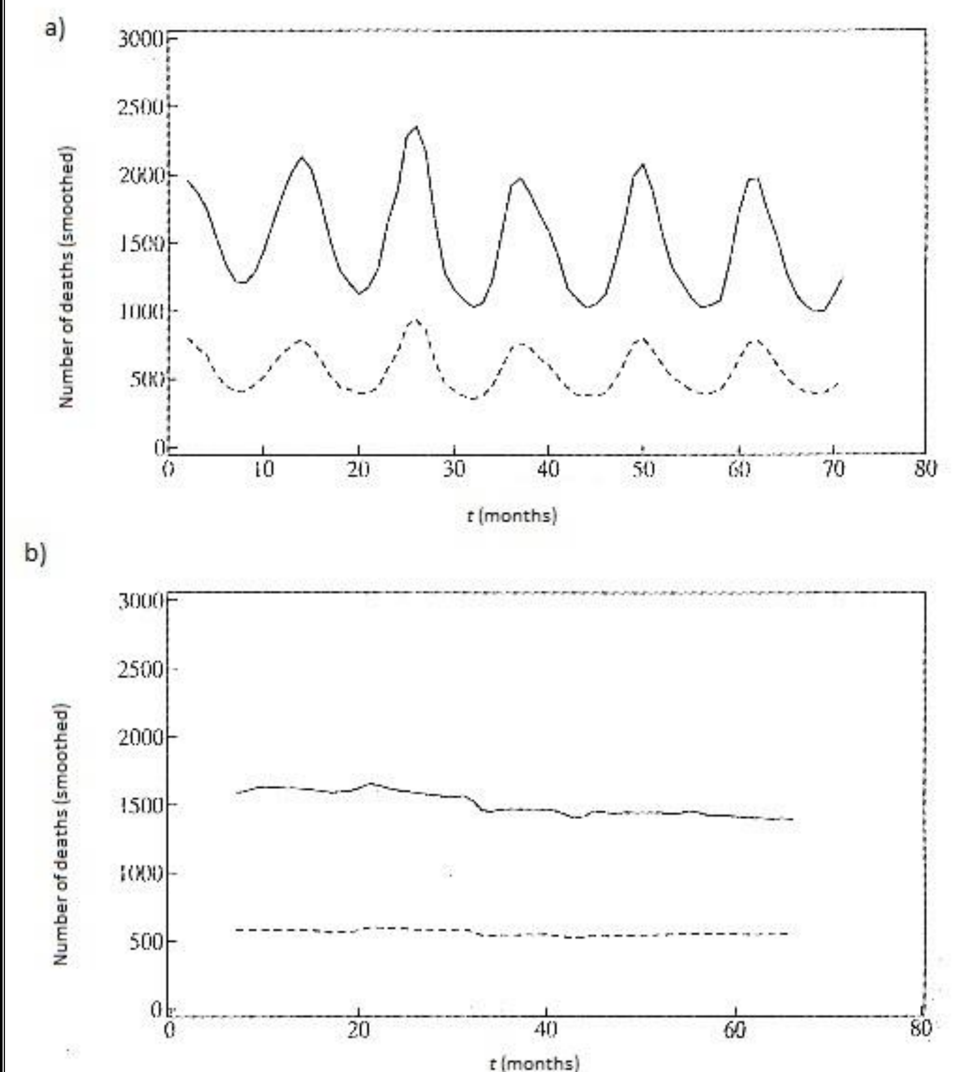
Usually *m* is chosen to be a fairly small number; otherwise, there would be a risk that the "replacement" averages might smooth the time series too much and inadvertently cover up details that *are* important [1]. Figure 2.8 shows the original time series of the monthly returns of death in the UK from bronchitis, emphysema, and asthma over the years 1974-1979: return for males and for females are shown separately [23].



Two time series of monthly returns of deaths in the United Kingdom attributed to bronchitis, emphysema, and asthma over the years 1974 to 1979. —, Males; – – –, females.

Figure 2.9

The following, Figure 2.10a and Figure 2.10b, show the results of applying both a three-point moving average and a thirteen-point moving average to the time series from Figure 2.8 [23]. Notice that a seasonal effect is present, in Figure 2.9a, when a three-point moving average is applied whereas a long-term trend is present, in Figure 2.9b, when a thirteen-point moving average is applied.



Two different moving average smoothers applied to monthly returns of deaths in the United Kingdom attributed to bronchitis, emphysema, and asthma: ——, males; – – –, females. (a) Three-point moving average. (b) Thirteen-point moving average.

Figure 2.10

A moving average on an *even* number of points requires one extra step: since there is no middle y_i to "replace" consecutive moving averages must be processed [1]. Thus it is often constructive to compute the moving average over an odd number of time periods so that we have values of comparison that are actual values [15]. In the example performed about if an even-numbered order had been used, then the response would have been plotted between two months. If the order four were being used, then the months of July, August, September, and October would be used to calculate the moving average, and the response would be plotted between August and September.

$$\bar{Y}_{Aug/Sep-83} = \frac{July_{83} + August_{83} + September_{83} + October_{83}}{4}$$

Hence, there would be no response to compare the *original* value of August (or September) to. In cases such as this, *centering* the moving average solves the problem. The resulting moving average is a mid-value for the first moving average of an even order [18]. For example, after finding the moving average of order 4 and plotting this point between August/September, and then finding the moving average of order 4 for September/October, one would take these two data points and the average of these two would be the centered moving average (of order 4) for the month of September.

$$\bar{Y}_{Sep-83} = \frac{\bar{Y}_{Aug/Sep-83} + \bar{Y}_{Sep/Oct-83}}{2}$$

The primary disadvantage of using a moving average for smoothing is that unless $m = 1$ we do not have a smoothed value corresponding to each response value. In the example above, where $m = 5$, there were no smoothed values corresponding to the first two or to the last two values (months). The choice of m in defining a sequence of moving averages is not always arbitrary [1]. If a time series shows a pronounced seasonal effect and tends to achieve a maximum every *p* periods, then it makes sense to set *m* equal to *p* [1]. Doing so will effectively eliminate all the season-to-season fluctuations [1].

II. Exponential Smoothing

Exponential smoothing is a method, conceived of by Robert Macaulay in 1931 and developed by Robert G. Brown during World War II, for extrapolative forecasting from series data [18]. As stated earlier [18]:

In the method of moving averages, a y_i in the original time series is replaced by an average of m (= 2k+1) points, each observation y_i is given equal weight.

Hence, the effect of the past observations on the future ones is equal. To compensate for the irregular weighting of observations, exponential smoothing is introduced [18].

Given a non-seasonal time series with no systematic trend, $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$ it is natural to take as an estimate of x_{n+1} a weighted sum of the past observations:

$$\hat{x}(n, 1) = c_0 x_n + c_1 x_{n-1} + c_2 x_{n-2} + \cdots \qquad (2.3)$$

where the $\{c_i\}$ are weights [24]. It seems sensible to give more weight to recent observations and less weight to observations further in the past [24]. The speed at which remote responses are dampened (smoothed) out is determined by the selection of the smoothing constant α [15]. An intuitively appealing set of weights are geometric weights, which decrease by a constant ratio [24]. In order that the weights some to one, we take

$$c_i = \alpha(1 - \alpha)^i \qquad i = 0,1, \ldots$$

where α is a constant such that $0 < \alpha < 1$ [24]. For values of α near 1 remote responses are dampened out quickly; for α near 0 they are dampened out slowly [15]. Then equation (1.3) becomes

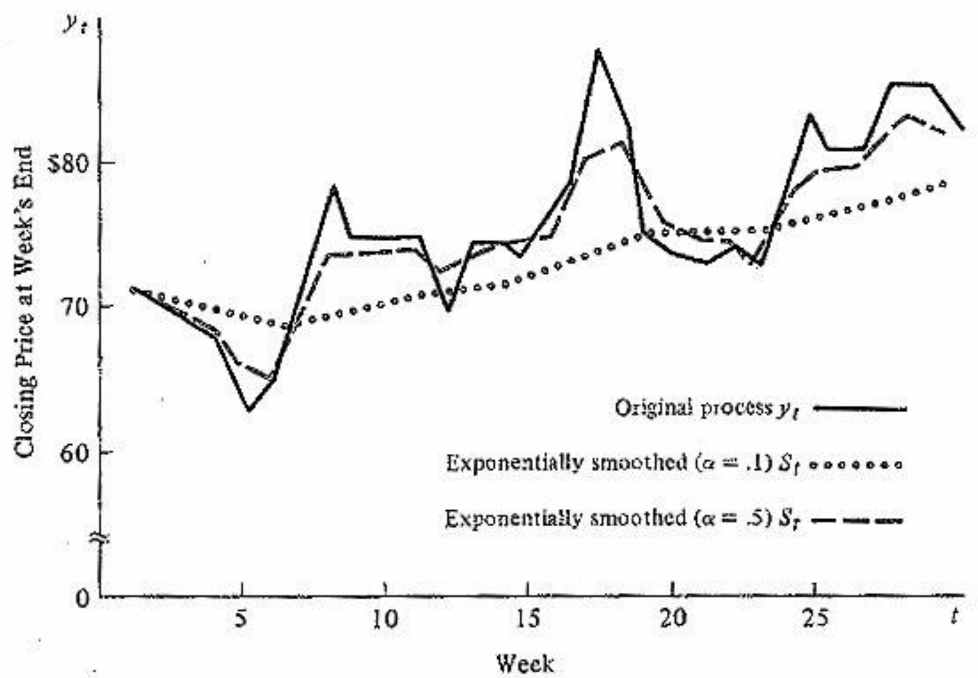
$$\hat{x}(n,1) = \alpha x_n + \alpha(1-\alpha)x_{n-1} + \alpha(1-\alpha)^2 x_{n-2} + \cdots \tag{2.4}$$

This equation implies an infinite number of past observations, but in practice there will only be a finite number [24]. So equation (1.4) is customarily rewritten in the *recurrence* form as

$$\hat{x}(n,1) = \alpha x_n + (1-\alpha)[\alpha x_{n-1} + \alpha(1-\alpha)x_{n-2} + \cdots]$$

$$= \alpha x_n + (1-\alpha)\hat{x}(n-1,1) \tag{2.5}$$

The procedure defined by equation (1.5) is called *exponential smoothing*. The adjective 'exponential' arises from the fact that geometric weights lie on an exponential curve, but the procedure could equally well be called *geometric smoothing* [24]. When the underlying response is quite "volatile" (the magnitude of the random variation is large), then it is best to average out the effects of the random variation quickly [15]. Thus a small (i.e. a value closer to 0)smoothing constant should be selected so that the smoothed value S_t will reflect S_{t-1}, the averaged values from the first (t – 1) time periods, to a great extent than it reflects the "noisy" measurement y_t [15]. Similarly, for a moderately stable process a large (i.e. a value closer to 1) smoothing constant would be selected [15]. The following example illustrates the use of three different exponential smoothing methods on a volatile time series that addresses closing prices for the securities of the Color-Vision Company, a manufacturer of color television sets, over 30 consecutive weeks [15]:



Week's End Closing Security Prices for the Color-Vision Company Over 30 Weeks, with the Two Exponentially Smoothed Processes Superimposed

Figure 2.11

Observe that in both exponentially smoothed time series, the response appears more stable than the original series, and out of the two smoothed time series, the time series with α = 0.1 appears to be much more stable than the other two. This series (with α = 0.1) also suggests the presence of a linear trend with a cyclic effect. Hence the true components of the original series (if a linear trend and a cyclic effect are the true components) most readily become apparent when the original series is smoothed exponentially using a small smoothing constant (α). This is not a generalization; the small smoothing constant happens to yield the best results for the data used here [15].

Perhaps the major advantage of smoothing methods is typified by the old saying that a picture is worth a thousand words [15]. Moving averages and exponentially smoothed time series *sometimes* make trends, cycles, and seasonal effects more visible to the eye and consequently lead to a simple and useful description of the time series process [15]. In this study a moving average of order 11 will be used. There are other types of smoothing methods that may be used for a time series such as index numbers and decomposition models.

E. Adjustment of Seasonal Data

Suppose we are interested in examining short-term trends or the effect of an assumed business cycle on the time series representing business activity, such as the effect of temperature on the sales of water gear. This is a difficult, if not impossible, task if the time series exhibits a pronounced seasonal component, because the seasonal fluctuations could overwhelm the other components [15]. *Seasonal adjustment* may be defined as the removal of seasonality from a time series [18]. If the seasonal component can be removed from the time series, identification, examination, and interpretation of trends and cycles are greatly simplified [15]. Seasonal effects, although they may vary somewhat in their average time of occurrence during the year, have a degree of regularity which other elements of time series (trend, cyclical component) usually do not [20]. There are several different reasons for wanting to examine seasonal effects some of which are [20]:

- To compare a variable at different points of the year as a purely intra-year phenomenon; for example, in deciding how many hotels to close out of season, or at what points to allow stocks to run down.

- To remove seasonal effects from the series in order to study its other constituents uncontaminated by the seasonal component.
- To ‘correct’ a current figure for seasonal effects, e.g. to state what the unemployment figures in a winter month would have been if customary seasonal influences (like Christmas) had not increased them.

These objectives are not the same, and it follows that one single method of seasonal determination may not be suitable to meet them all [20].

The moving average smoothing technique (discussed earlier) is used to remove the seasonal component from a time series. In many time series, the seasonal period is either 4 or 12 months [15]. When the time points are months and the time series is seasonal, the seasonal period is almost always 12 months [15]. Observe, again, the data in Appendix 2.1. If the moving average of order 12 is calculated, the first step (since this is order 12, an even number) is to calculate the simple moving averages of each “month”; again, since *m* (=12) is an even number we will plot our moving average in the center of the two months. For example using the data in Appendix 2.1

$$\bar{Y}_{June/July-83} = \frac{June_{83} + August_{83} + \cdots + December_{83}}{12}$$

$$= \frac{19.6 + 18.6 + \cdots + 20.9}{12}$$

$$= \frac{296}{12} = 24.667$$

Similarly,

$$\bar{Y}_{July/Aug-83} = 24.975$$

Then, to find $\bar{Y}_{July-83}$ the average of the two previous calculations must be taken

$$\bar{Y}_{July-83} = \frac{\bar{Y}_{June/July-83} + \bar{Y}_{July/Aug-83}}{2}$$

$$\bar{Y}_{July-83} = \frac{24.667 + 24.975}{2} = 24.821$$

In Figure 2.12 the original values are in blue and the moving average values (of order 12) are in gray.

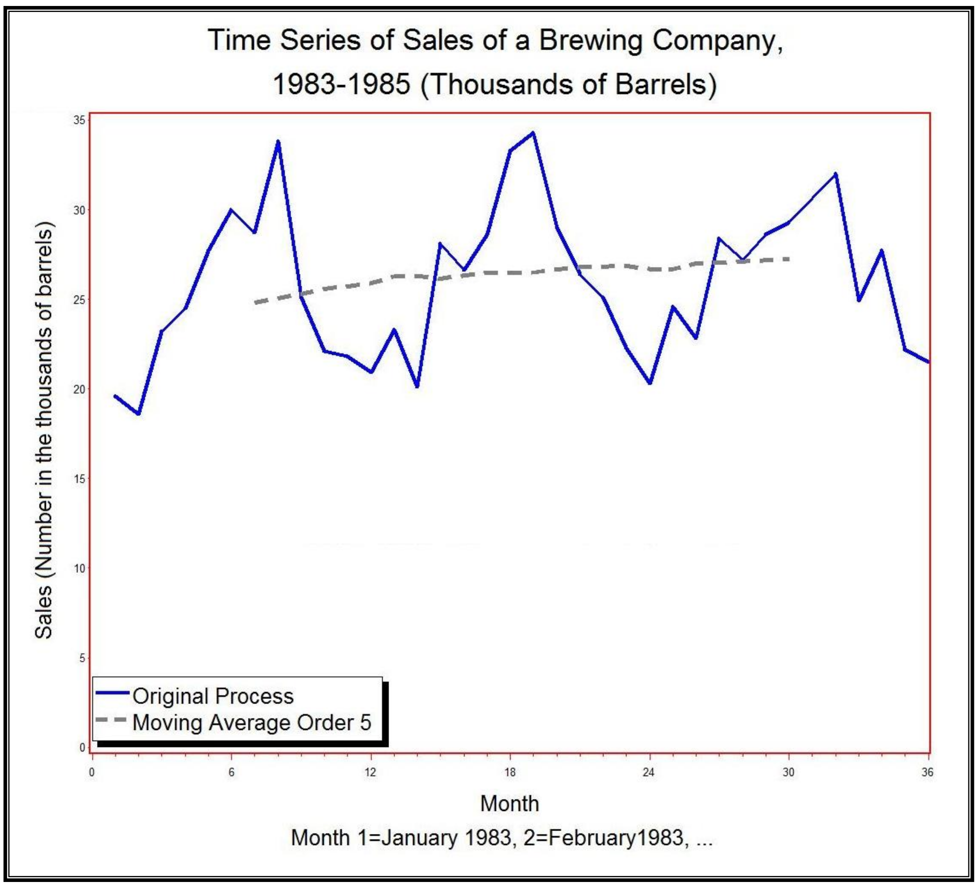


Figure 2.12

In the example earlier when a simple moving average of order 5 was applied a seasonal component was evident (in red) and after applying a simple moving average of order 12 the seasonal component is removed, as desired, and only the trend component is evident (gray), allowing "identification, examination, and interpretation of trends and cycles."

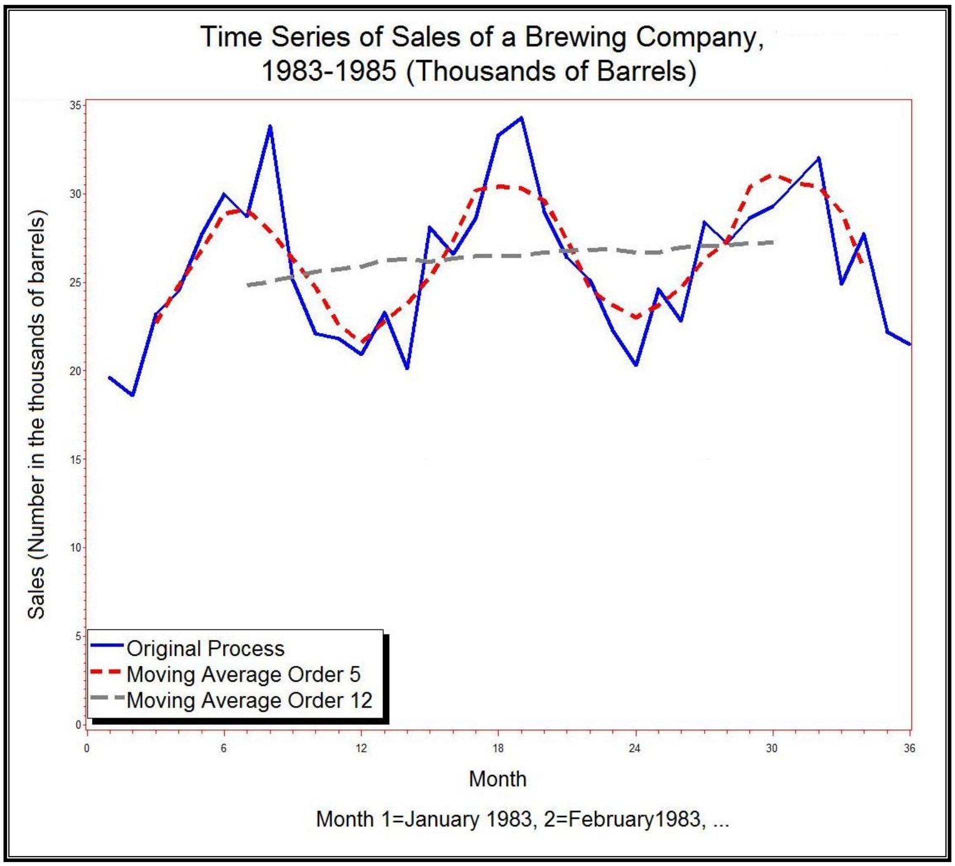


Figure 2.13

There are three types of models, depending on whether the seasonal effect is additive or multiplicative, that are popular for time series that need seasonal adjustment [20]. If m_t is the smooth component of the time series (trend and cyclical effects), s_t is the seasonal component and ε_t the residual term, we may have [20]

$$y_t = m_t + s_t + \varepsilon_t, \tag{2.6}$$

$$y_t = m_t s_t \varepsilon_t, \tag{2.7}$$

or, the multiplicative-seasonal model

$$y_t = m_t s_t + \varepsilon_t. \tag{2.8}$$

The purely multiplicative model (2.7) may be converted to linear form by taking logarithms

$$\log y_t = \log m_t + \log s_t + \log \varepsilon_t. \tag{2.9}$$

In making the transformation (2.9) we assume that ε_t in (2.7) can only take on positive values; otherwise $\log \varepsilon_t$ is undefined [20]. Thus it is convenient to write $n_t = \log \varepsilon_t$, where n_t is a random variable with zero mean [20]. Then (2.9) becomes

$$\log y_t = \log m_t + \log s_t + n_t \tag{2.10}$$

corresponding to

$$y_t = m_t s_t \varepsilon^{n_t}. \tag{2.11}$$

When using the additive model (2.6) to separate the seasonal and trend components, it is reasonable to impose the condition that the sum of the seasonal effects is zero [20]. Thus, for monthly data, the subscript *t* may be written as $t = 12(i-1) + j$, corresponding to the *j*th month of the *i*th year [20]. If it can be assumed that the seasonal effects are the same in different years, then the condition

$$\sum_{j=1}^{12} s_t = \sum_{j=1}^{12} s_j = 0$$

(2.12)

may be imposed since $s_{12i+j} = s_j$ for all i and j [20].

Alternative methods of seasonal adjustment are reviewed by Chatfield [24] which involves eliminating a seasonal effect by differencing and by Wallis [25] using the popular X-11 method which employs a series of linear filters. The data in this paper has already been seasonally adjusted; hence the only effects that we will be inspecting for will be cyclical and trend.

F. Forecasting

A modest objective of any time series analysis is to provide a concise *description* of a historic series [23]. However, a more ambitious task is to *forecast* future values of a series [23]; much time series analyses have been developed to this specific end [23]. We all exist in an environment governed by time. Business organizations, public organizations, and individuals thus have the common goal of allocating available time among competing resources in some optimal manner. This goal is accomplished by making forecasts of future activities and taking the proper actions as suggested by these forecasts [15].

Forecasts can be short-term or long-term. The short-term forecast is usually planned for looking no more than one year into the future. Harrison and Pearce insist at least seven to ten years of historical data are required for long-term forecasting [26]. In business and public administration this involves forecasting sales, price changes, and customer demand, which, in turn, reflect the need for seasonal employment, short-term capital expenditures, and inventory management procedures [15]. The long-term forecast usually looks 2 to 10 years into the future and is used as a planning model for product line and capital investment decisions, as indicated by changing demand patterns [15].

Forecasting should be seen as “an art that becomes more perfect as the forecaster gains experiences and then ability to adapt procedures to meet the changing environment” [15], rather than as an *exact* science. All that can be expected is that the benefits gained by forecasting offset the opportunity cost for not forecasting [15].

CHAPTER III ANALYSIS OF CRIME DATA

As stated earlier, the data in this paper has already been seasonally adjusted, so the only two components left to identify (if they are indeed present) are the cyclical component and a linear trend. It will be most interesting to plot the time series for these crimes on the same model, so this will be done for all of the figures in order to save time and allow for us to better see correlations. Figure 3.1 shows the original process of the crimes for 1975-2009.

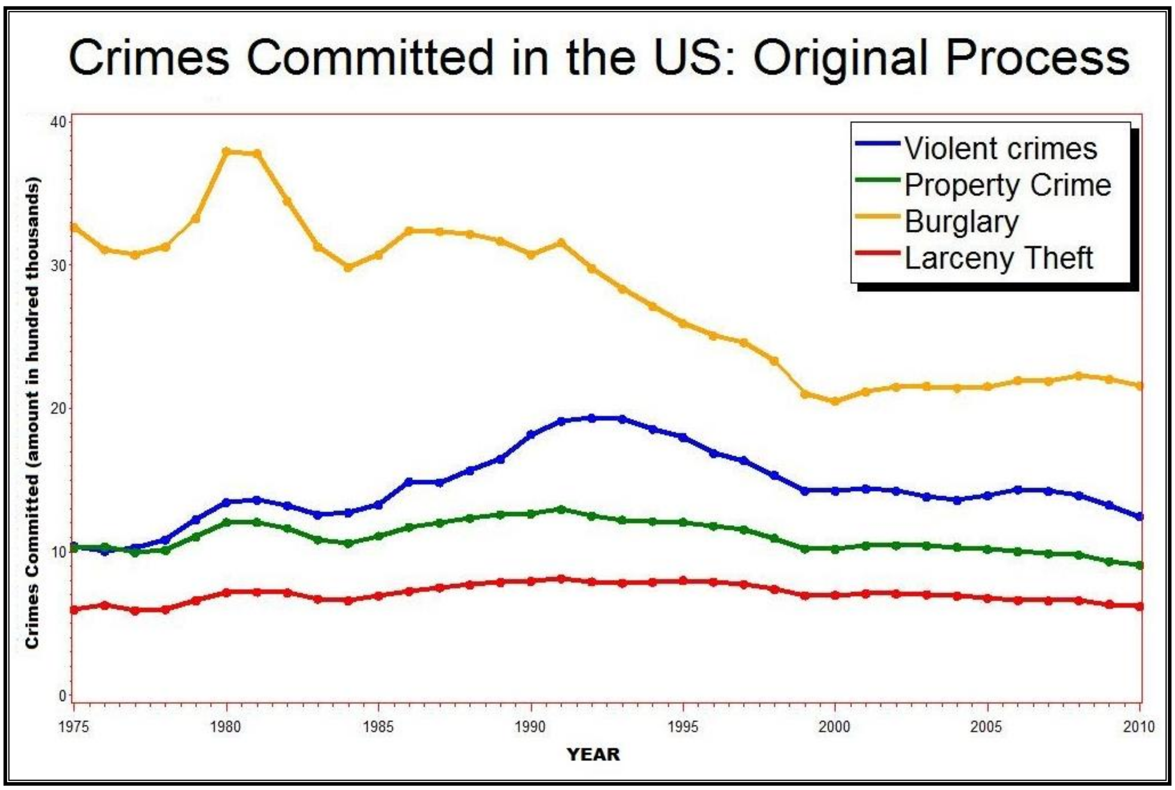


Figure 3.1

Notice that there does not appear to be a linear trend specifically in the burglary time series nor does there in violent crime. However property crime and larceny theft appear to have a linear trend somewhat. After applying a moving average of order 11 (see Figure 3.2), the time series for all four crimes appear to follow a linear trend, except violent crimes which appears to follow a slightly exponential trend.

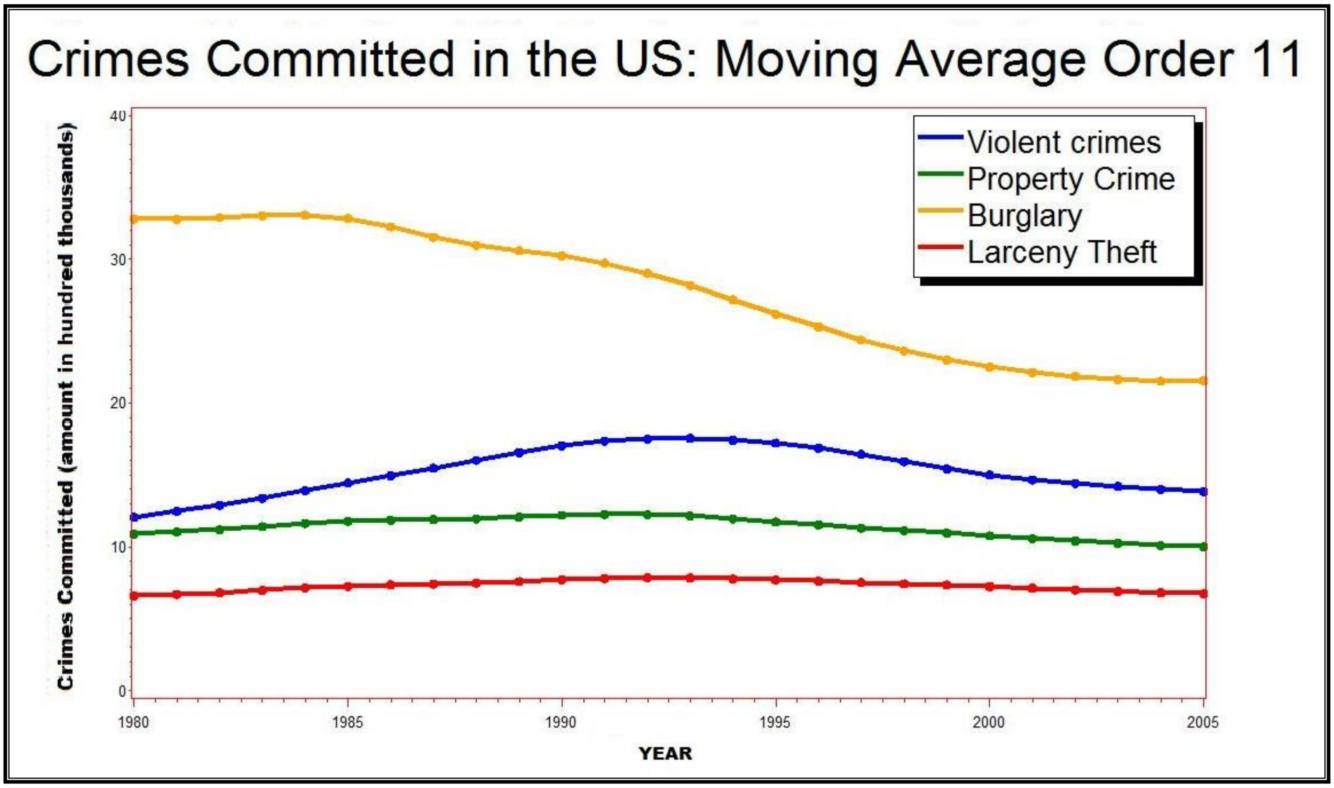


Figure 3.2

Crime does appear to be decreasing for the most part. After running a correlation test in SAS it was found that these crimes are not correlated at all with law enforcement officer employment which leads me to the belief that crime is not effected by law enforcement officer employment. In Figure 3.1 that the two crimes burglary and violent crime do appear to have a cyclical effect. But whenever we take a closer look at the time series (by changing the scale of the y-axis) we find that the linear trend has been taken away in Figure 3.2.

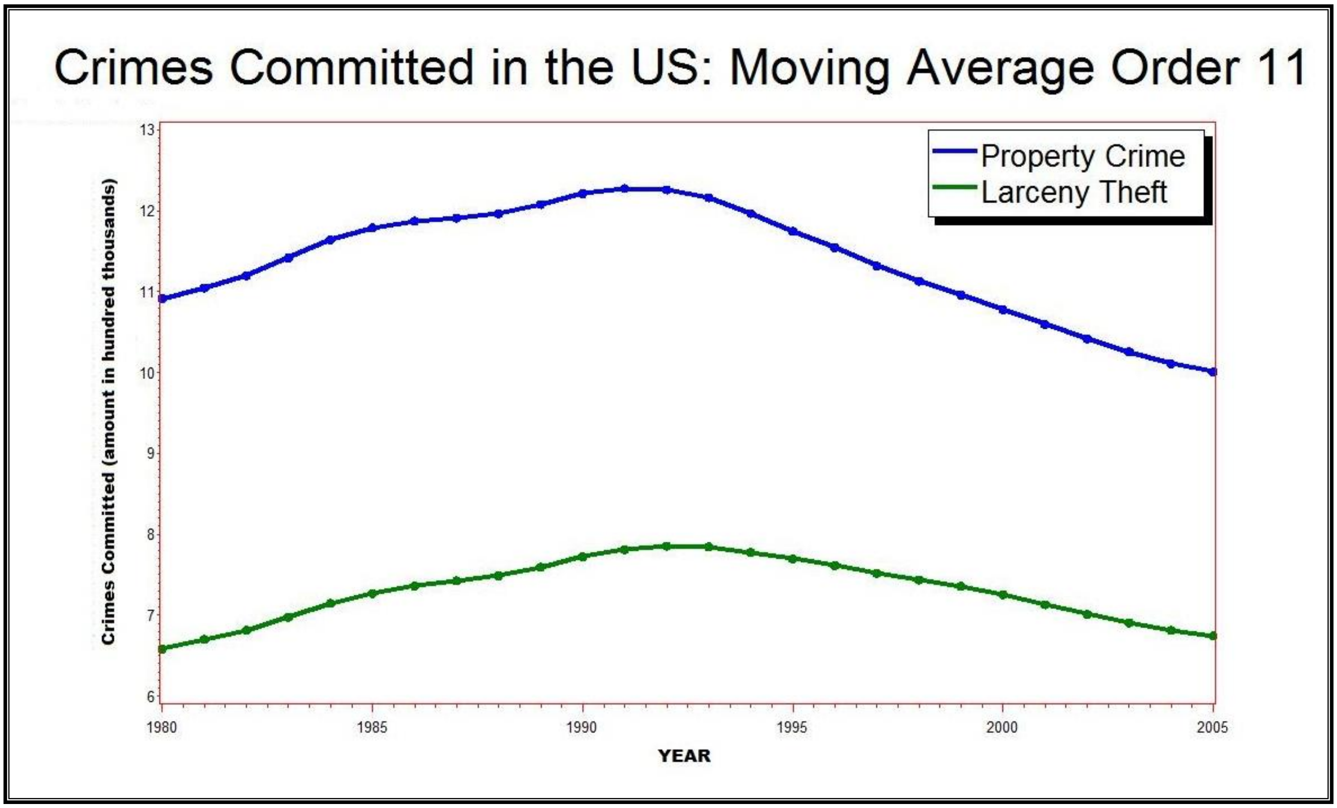


Figure 3.3

Hence, there is a cyclical effect present in all four crimes. The cyclical component of crime time series might well be a cousin to the predator-prey model. As a certain type of crime increases, the law enforcement resources dedicated to that type of crime increases. If this increased attention causes a decrease in the crimes, soon the law enforcement resources would be shifted to other types of crimes. The peaks in crime would correspond roughly to valleys in resources against that type of crime and vice-versa.

For this reason it should be obvious that it would be very easy to show the public a graph such as Figure 3.1 and for them to make the conclusion that crime does follow a linear trend and, hence, crime is indeed decreasing. However, after seeing Figure 3.3 one might argue that crime follows a cyclical pattern and will peak again.

For example, on October 19, 2011, Vice President Biden commented that without more funding for law enforcement [27]:

> *Murder will continue to rise, rape will continue to rise, all crimes will continue to rise…. Go look at the numbers.*

This is a very powerful statement and would cause alarm to any citizen of the United States. Figure 3.4 displays the original time series for both murder and rape (the two types of crime Biden *directly* commented on). Murder and rape are displayed as the amount in hundreds of thousands of offenses committed.

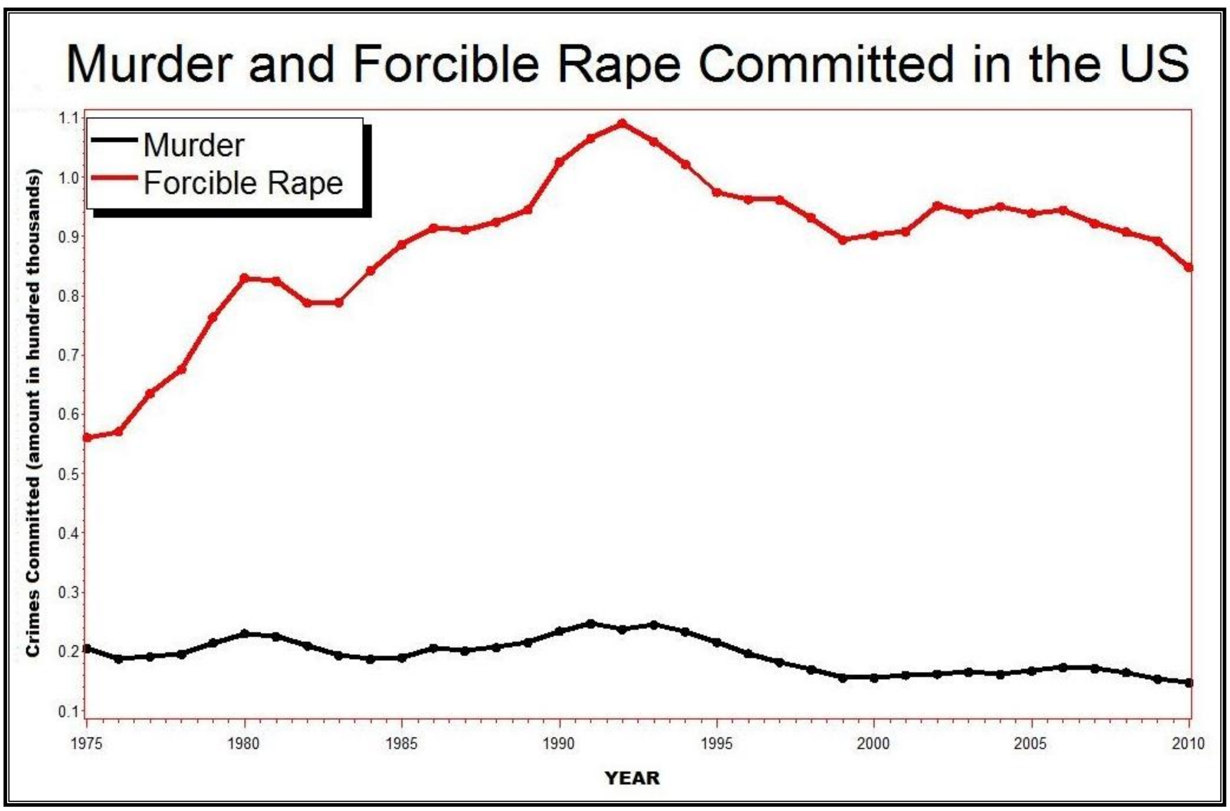


Figure 3.4

Notice that there does not appear to be an increase in rape or murder over the last several years.

To take it a step further, Figure 3.5 displays the same time series but with the number of full-time law enforcement officers overlaid on top. The number of law enforcement officers is in the millions.

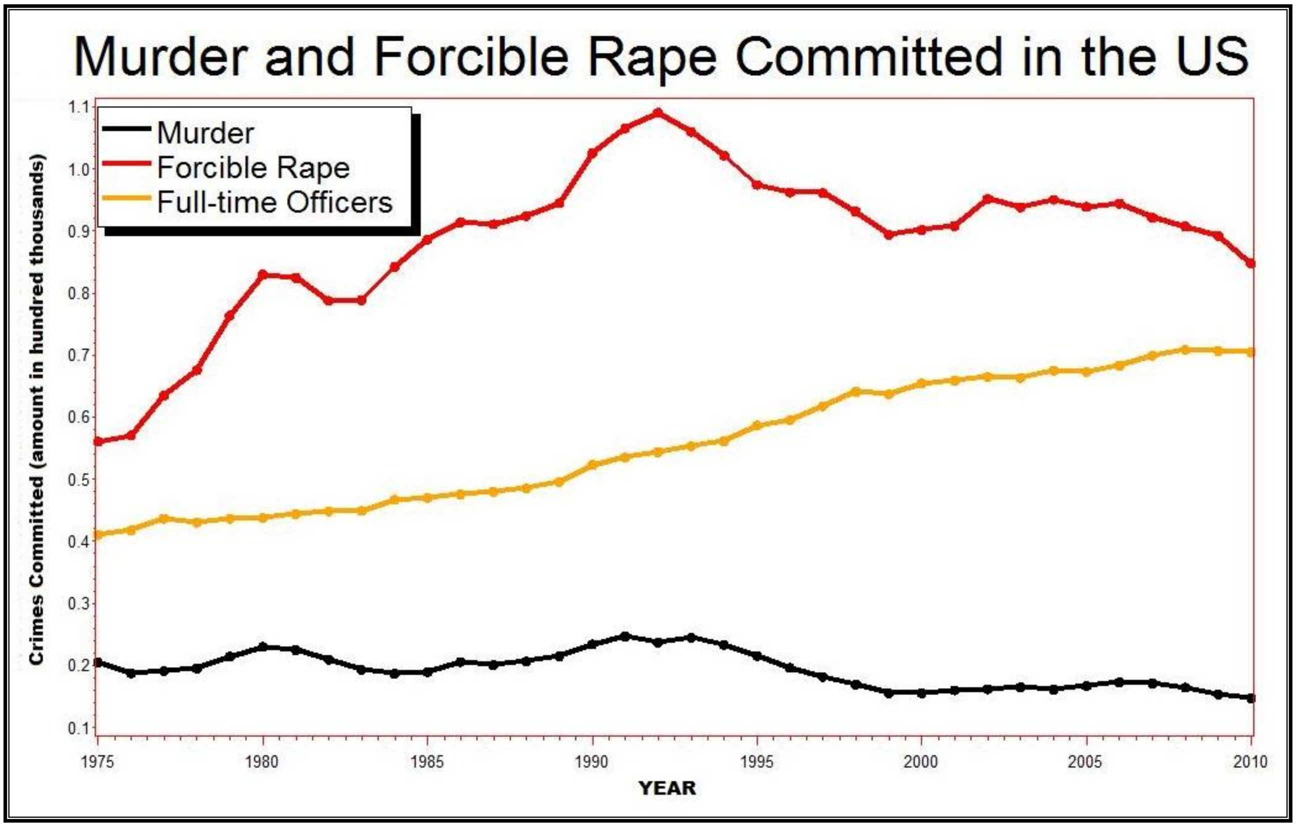


Figure 3.5

The point is not whether Vice President Biden was trying to twist the statistics or not but rather the power of statistics and, more specifically in this case, the power of time series analysis.

A question of interest is why time series from different crimes are correlated. While this is a topic for future study, we can show that this phenomenon occurs, Appendix 3.1 displays the correlation between the different crimes.

CHAPTER IV. CONCLUSIONS

A. Summary

Time series analysis is not only an important study in today's society, it is an extremely powerful one. Time series analysis is a meaningful study for multiple reasons. The analysis of the meaningful data of this data is that it is seasonal.

B. Suggestions for Further Study

The crimes in this study were highly correlated and it would be interesting to find out why the crimes are highly correlated, it could be a spurious correlation, where the data just happens to be going up (or down) at the same rate. It would be very interesting to find out how much crime is not reported, as of right now, no program is instated to record unreported crime which is cause for concern. Perhaps this could be done by studying the number of 911 phone calls made or the number of ER visits at hospitals.

Professor Steven Levitt from the University of Chicago suggests that the following four factors are responsible for the observed decline in crime [28]:

> *Increases in the number of police, the rising prison population, the waning crack epidemic and the legalization of abortion.*

He goes on to say that [28]:

> *Studies on the connection between the number of police and crime in the 1970s and 1980s, as surveyed by Cameron (1988), tended to find an insignificant or negative correlation, because these studies typically failed to account for the endogeneity problem.*

REFERENCES

[1] Larsen, R., Marx, M., and Cooil, B. *Statistics for Applied Problem Solving and Decision Making.* Pacific Grove, CA: Brooks/Cole Publishing Company (1997), pp. xi, .

[2] Larsen, R., Marx, M., and Cooil, B. An Introduction *to Mathematical Statistics and Its Applications, 4th Edition.* Pacific Grove, CA: Brooks/Cole Publishing Company (1997), pp. xi, .

[3] Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2012-13 Edition*, Statisticians. Retrieved on March 31,2012 from the Bureau of Labor Statistics website: http://www.bls.gov/ooh/Math/Statisticians.htm#tab-6.

[4] Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2012-13 Edition*, Statisticians. Retrieved on March 31,2012 from the Bureau of Labor Statistics website: http://www.bls.gov/ooh/Math/Statisticians.htm#tab-2.

[5] The Federal Bureau of Investigation. *"FBI Releases 2008 Crime Statistics".* Retrieved March 29, 2012 from the website:
http://www.fbi.gov/news/stories/2009/september/crimestats_091409.

[6] Lichtelle, Louis. "Jobless Rate Hits 7.2%, a 16-Year High." Retrieved on March 29, 2012 from the *New York Times* printed on January 10,2009 website:
http://www.nytimes.com/2009/01/10/business/economy/10jobs.html?_r=1#.

[7] The United States Department of Labor. *"The Recession of 2007-2009".* Retrieved on March 29, 2012 from the Bureau of Labor Statistics website:
http://www.bls.gov/spotlight/2012/recession/.

[8] The Federal Bureau of Investigation. "Crime in the United States: 2010 Offenses Known to Law Enforcement". Retrieved March 29, 2012 from the U.S. Department of Justice website: http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/offenses-known-to-law-enforcement.

[9] The Federal Bureau of Investigation. "Crime in the United States: Offenses Cleared." Retrieved March 29, 2012 from the U.S. Department of Justice website:
http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/clearances.

[10] Nevada Department of Pulic Safety. "*Frequently Asked Questions*. Retrieved March 29, 2012 from website: http://www.nvrepository.state.nv.us/ucr/forms/FAQsforUCR.pdf.

[11] University of California Los Angeles. *Crime Reporting Standards.* Retrieved March 29, 2012 from website:
http://map.ais.ucla.edu/portal/site/UCLA/menuitem.789d0eb6c76e7ef0d66b02ddf848344a/?vgnextoid=01cca70a079ae010VgnVCM200000dd6643a4RCRD.

[12] The Federal Bureau of Invesitgation. *Uniform Crime Reporting Handbook.* Retrieved March 30, 2012 from the U.S. Department of Justice website: http://www2.fbi.gov/ucr/handbook/ucrhandbook04.pdf.

[13] Kilpatrick, D.G. "What is violence against women? Defining and measuring the problem." *Journal of Interpersonal Violence* .19. 1209-1234. Retrieved on March 31, 2012 from website: http://www.judiciary.senate.gov/pdf/10-09-14KilpatrickTestimony.pdf.

[14] Kilpatrick, G.D., Ruggiero, K. *Making Sense of Rape in America: Where do the Numbers Come From and What Do They Mean?* Retrieved on March 30, 2012 from website: http://new.vawnet.org/Assoc_Files_VAWnet/MakingSenseofRape.pdf

[15] Mendenhall, W., Reinmuth, J., Beaver, R., Duhan, D. *Statistics for Management and Economics, 5th Edition.* Boston, MA: PWS Publishers/Wadsworth, Inc, (1978), pp. 598-606.

[16] McClave, J., Sincich, T. *Statistics, 11th Edition.* Upper Saddle River, NJ. Pearson Education Inc (2009), pp. 722.

[17] Granger, C.W.J., and Newbold, P. *Forecasting Econometric Time Series.* New York, NY: Academic Press, 2005, p. 1.

[18] Yaffee, R., McGee, M. *Introduction to Time Series Analysis and Forecasting.* London, UK:Academic Press (2000), pp. 2, 3,18-23, 50, 510.

[19] Webber, C., Zbilut, J. *Recurrence Quantification Analysis of Nonlinear Dynamical Systems.* Retrieved on March 30, 2012 from the National Science Foundation website: http://www.nsf.gov/sbe/bcs/pac/nmbs/chap2.pdf.

[20] Kendall, M., Ord, J. *Time Series, 3rd Edition.* New York, NY: Oxford University Press (1990), pp. 27,42,.

[21] Day, D. *Comparison of UAH and RSS Time Series with Common Baseline.* Retrieved on March 31, 2012 from Climate Charts and Graphs website: http://chartsgraphs.wordpress.com/2011/03/30/comparison-of-uah-and-rss-time-series-with-common-baseline/.

[22] Morgan, Ash. "The Economic Impact of Diving the USS Oriskany on the Regional Economy." Pensacola, Florida. From the journal *Northwest Florida Economy"* published by the Haas Center for Business Research and Economic Development at the University of West Florida, Winter 2008. Retrieved on April 1, 2012 website: ftp://haasgis.cob.uwf.edu/NW_Florida_Economy/NW_FL_Economy/2008/winter08.pdf.

[23] Diggle, Peter. *Time Series: A Biostatistical Approach.* New York, NY: Oxford University Press (1990), pp. 7, 24.

[24] Chatfield, C. *The Analysis of Time Series An Introduction*. New York, NY: Chapman and Hall (1989), pp 68-69.

[25] Wallis, K.F. "Seasonal Adjustment and Revision of Current Data." From *Journal of the Royal Statististical Society* (1982) Series A. Blackwell Publishing, pp. 145.

[26] Harrison, P.J. and Stevens, C.F. "The Use of Trend Curves as an Aid to Market Forecasting." From *Industrial Marketing Management* (1972), pp. 2, 149-170.

[27] Jones, S. "Biden's Aburd Claims About Rising Rape and Murder Rates." Washington D.C. (2011) From the *Washington Post* retrieved on April 14, 2012 from website: http://www.washingtonpost.com/blogs/fact-checker/post/bidens-absurd-claims-about-rising-rape-and-murder-rates/2011/10/20/gIQAkq0y1L_blog.html.

[28] Levitt, Steven. "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not." Vanderbilt University Nashville, Tennessee. *Journal of Economic Perspectives, Volume 18 Number 1* (Winter 2004), pp. 163, 164, 176. Retrieved on April 17, 2012 from website: http://pricetheory.uchicago.edu/levitt/Papers/LevittUnderstandingWhyCrime2004.pdf.

APPENDIX

Appendix 1.1

Year	Population	Violent crime	Murder	*Forcible rape*	Robbery	Aggravated assault	Property crime	Burglary	Larceny-theft	Motor Vehicle Theft	Number of Full-Time Officers	Unemp. Rate
1975	213124000	10.40	0.21	0.56	4.71	4.93	10.25	32.65	5.98	10.10	411000	8.5
1976	214659000	10.04	0.19	0.57	4.28	5.01	10.35	31.09	6.27	9.66	418000	7.7
1977	216332000	10.30	0.19	0.64	4.13	5.34	9.96	30.72	5.91	9.78	437000	7.1
1978	218059000	10.86	0.20	0.68	4.27	5.71	10.12	31.28	5.99	10.04	431000	6.1
1979	220099000	12.25	0.21	0.76	4.81	6.29	11.04	33.28	6.60	11.13	437000	5.8
1980	225949264	13.45	0.23	0.83	5.66	6.73	12.06	37.95	7.14	11.32	438442	7.1
1981	229146000	13.62	0.23	0.83	5.93	6.64	12.06	37.80	7.19	10.90	444240	7.6
1982	231534000	13.22	0.21	0.79	5.53	6.69	11.65	34.47	7.14	10.62	448927	9.7
1983	233981000	12.58	0.19	0.79	5.07	6.53	10.85	31.30	6.71	10.08	449370	9.6
1984	236158000	12.73	0.19	0.84	4.85	6.85	10.61	29.84	6.59	10.32	467117	7.5
1985	238740000	13.29	0.19	0.89	4.98	7.23	11.10	30.73	6.93	11.03	470678	7.2
1986	240132887	14.89	0.21	0.91	5.43	8.34	11.72	32.41	7.26	12.24	475853	7
1987	242288918	14.84	0.20	0.91	5.18	8.55	12.02	32.36	7.50	12.89	480383	6.2
1988	244498982	15.66	0.21	0.92	5.43	9.10	12.36	32.18	7.71	14.33	485566	5.5
1989	246819230	16.46	0.22	0.95	5.78	9.52	12.61	31.68	7.87	15.65	496353	5.3
1990	249464396	18.20	0.23	1.03	6.39	10.55	12.66	30.74	7.95	16.36	523262	5.6
1991	252153092	19.12	0.25	1.07	6.88	10.93	12.96	31.57	8.14	16.62	535629	6.8
1992	255029699	19.32	0.24	1.09	6.72	11.27	12.51	29.80	7.92	16.11	544309	7.5
1993	257782608	19.26	0.25	1.06	6.60	11.36	12.22	28.35	7.82	15.63	553773	6.9
1994	260327021	18.58	0.23	1.02	6.19	11.13	12.13	27.13	7.88	15.39	561543	6.1
1995	262803276	17.99	0.22	0.97	5.81	10.99	12.06	25.94	8.00	14.72	586756	5.6
1996	265228572	16.89	0.20	0.96	5.36	10.37	11.81	25.06	7.90	13.94	595170	5.4
1997	267783607	16.36	0.18	0.96	4.99	10.23	11.56	24.61	7.74	13.54	618127	4.9
1998	270248003	15.34	0.17	0.93	4.47	9.77	10.95	23.33	7.38	12.43	641208	4.5
1999	272690813	14.26	0.16	0.89	4.09	9.12	10.21	21.01	6.96	11.52	637551	4.2
2000	281421906	14.25	0.16	0.90	4.08	9.12	10.18	20.51	6.97	11.60	654601	4
2001	285317559	14.39	0.16	0.91	4.24	9.09	10.44	21.17	7.09	12.28	659104	4.7
2002	287973924	14.24	0.16	0.95	4.21	8.91	10.46	21.51	7.06	12.47	665555	5.8
2003	290788976	13.84	0.17	0.94	4.14	8.59	10.44	21.55	7.03	12.61	663796	6
2004	293656842	13.60	0.16	0.95	4.01	8.47	10.32	21.44	6.94	12.38	675734	5.5
2005	296410404	13.91	0.17	0.94	4.17	8.63	10.17	21.54	6.78	12.35	673146	5.1
2006	299398484	14.35	0.17	0.94	4.49	8.74	10.02	21.95	6.63	11.98	683396	4.6
2007	301621157	14.23	0.17	0.92	4.47	8.66	9.88	21.90	6.59	11.00	699850	4.6
2008	304059724	13.94	0.16	0.91	4.44	8.44	9.77	22.29	6.59	9.59	708569	5.8
2009	307006550	13.26	0.15	0.89	4.09	8.13	9.34	22.03	6.34	7.96	706886	9.3
2010	308745538	12.46	0.15	0.85	3.68	7.79	9.08	21.60	6.19	7.37	705009	9.6

Appendix 1.2 The Heirarchy Part I Offenses [12]:

1. Criminal Homicide
    a. Murder and Nonnegligent Manslaughter
    b. Manslaughter by Negligence
2. Forcible Rape
    a. Rape by Force
    b. Attempts to Commit Forcible Rape
3. Robbery
    a. Firearm
    b. Knife or Cutting Instrument
    c. Other Dangerous Weapon
    d. Strong-arm—Hands, Fists, Feet, etc.
4. Aggravated Assault
    a. Firearm
    b. Knife or Cutting Instrument
    c. Other Dangerous Weapon
    d. Hands, Fists, Feet, etc.—Aggravated Injury
5. Burglary
    a. Forcible Entry
    b. Unlawful Entry—No Force
    c. Attempted Forcible Entry
6. Larceny-theft (except motor vehicle theft)
7. Motor Vehicle Theft
    a. Autos
    b. Trucks and Buses
    c. Other Vehicles
8. Arson
    a. Structural
    b. Mobile
    c. Other

Appendix 2.1 The Heirarchy Part I Offenses [15]:

Year	Month	Actual Sales	5-Month Moving Average	12-Month Moving Average
1983				
	January	19.6		.
	February	18.6	.	.
	March	23.2	22.7	.
	April	24.5	24.8	.
	May	27.7	26.8	.
	June	30.0	28.9	.
	July	28.7	29.1	24.821
	August	33.8	27.9	25.038
	September	25.1	26.3	25.304
	October	22.1	24.7	25.596
	November	21.8	22.6	25.721
	December	20.9	21.6	25.896
1984				
	January	23.3	22.8	26.267
	February	20.1	23.8	26.3
	March	28.1	25.3	26.152
	April	26.6	27.3	26.333
	May	28.6	30.2	26.479
	June	33.3	30.4	26.475
	July	34.3	30.3	26.504
	August	29.0	29.6	26.671
	September	26.4	27.4	26.796
	October	25.1	24.6	26.833
	November	22.3	23.7	26.858
	December	20.3	23	26.692
1985				
	January	24.6	23.7	26.692
	February	22.8	24.7	26.983
	March	28.4	26.3	27.046
	April	27.2	27.3	27.092
	May	28.6	30.4	27.196
	June	29.3	31.1	27.242
	July	38.3	30.6	.
	August	32.0	30.4	.
	September	24.9	29	.
	October	27.7	25.7	.
	November	22.2		.
	December	21.5		.

Appendix 3.1

Correlation between violent crimes and other crimes in the FBI Heirarchy

Crimes (that are being tested with violent crime)	**r**	**r^2**
aggravated assault	0.96082	0.9232
robbery	0.95416	0.9104
murder and nonnegligent manslaughter	0.93366	0.8717
forcible rape	0.90462	0.8183
larceny-theft	0.89725	0.8051
motor vehicle theft	0.88254	0.7789
property crime	0.87229	0.7609
burglary	0.64768	0.4195