TIME SERIES ANALYSIS OF CRIME RATE IN ZIMBABWE 2008 – 2014

BY

TOGARA GUMBOCHUMA

B1231803

HBSCED MATHEMATICS

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS OF BACHELOR OF SCIENCE EDUCATION HONOURS MATHEMATICS

APRIL 2015
APPROVAL FORM

To be completed by the student:

I certify that this dissertation meets the preparation guidelines as presented in the Faculty guide and instructions for typing dissertations

_________________________________________   ........................................

(Signature of Student)                                                                          Date

To be completed by the supervisor:

______________This dissertation is suitable for submission to the Faculty.

______________This dissertation should be checked for conformity with the Faculty guidelines

_________________________________________   ........................................

(Signature of Supervisor)                                                                          Date

To be completed by the chairperson of the department.

I certify, to the best of my knowledge, that the required procedures have been followed and preparation criteria have been met for this dissertation.

_________________________________________   ........................................

(Signature of the Chairperson)                                                                          Date
RELEASE FORM

NAME OF AUTHOR: GUMBOCHUMA TOGARA

TITLE OF PROJECT: TIME SERIES ANALYSIS OF CRIME RATE IN ZIMBABWE 2008 – 2014

PROGRAMME FOR: BACHELOR OF SCIENCE EDUCATION

WHICH PROJECT HONOURS DEGREE IN MATHEMATICS WAS PRESENTED

YEAR GRANTED: 2015

Permission is hereby granted to Bindura University of Science Education Library to produce the single copies for private, scholarly or scientific research ONLY. The author reserves other publication rights. Neither the dissertation nor extensive extracts from it may be permitted or otherwise reproduced without the author’s written permission.

SIGNED:........................................

PERMANENT ADDRESS: House Number 6131 Granary Park
Phase 3.
Harare

Cell +26377 6272413, +26371 5005326
Abstract.
This research work sought to fit a time series model that can be used to forecast the future patterns of crime rate (HBT) and reveal trends in Zimbabwe. Using the time series data (total monthly crime HBT) obtained from the Zimbabwe Statistical Agency, Quarterly digest and the Compendium of statistics, a tentative ARIMA model was obtained. Empirical results from the study indicate that the model can be used for forecasting, since the forecasted values falls within the 95% confidence interval. It was concluded that the ARIMA(2 1 0) model was statistically significant based on the hypothesis tests by means of the Portmanteau test (Ljung-Box and Pierce) and the unit root tests, coupled with the analysis of the residual plots as well as the penalty statistics (Akaike Information Criteria) based on the principle of parsimony. I recommend the use of Time Series in modeling crime rates in Zimbabwe based on the results obtained in this research.
Acknowledgements
I wish to express my deepest appreciation to my supervisor, Mr K. Basira, who also doubles as Chairman Mathematics and Physics department BUSE for meticulously reading, supervising and making valuable contributions in the realization of the success of this work. My special thanks go to the Librarian, Zimbabwe Statistical Agency (ZimStat) for providing the necessary data, without whom the project would not have been successful. In addition, my appreciation goes to Mr J Mafodya of UZ, Department of Statistics for assisting me with statistical software for calculating unit root tests of stationarity. Furthermore, I thank all lectures in the Mathematics and Curriculum departments, Bindura University of Science Education for the immense academic knowledge imparted.
Dedication.
I graciously and dutifully dedicate this piece of work to the Almighty God who by his grace and mercies endowed me with knowledge and strength to undertake this academic journey successfully. I again, respectfully dedicate it to my lovely wife Lucia, and to the whole family.
List of Tables
4.1 Monthly HBT crime data
4.2 Unit root and stationary tests
4.3 Unit root and stationary tests for differenced data

List of Figures
2.1 Graphical representation of trend variation 9
2.2 Graphical representation of cyclic variation 11
4.1 Trend analysis of HBT CRIME Series 27
4.2 Histogram of HBT crime series data 27
4.3 ACF plot of HBT crime series data 28
4.4 PACF of HBT crime series data 29
4.5 Time series plot of differenced data 31
4.6 ACF plot of differenced data 33
4.7 PACF plot of differenced data 33
4.8 Residuals ACF plot of tentative model (2,1,0) 33
4.9 Residuals PACF plot of tentative model (2,1,0) 34
4.10 Residuals versus observation data 36
4.11 Histogram of residuals 37
4.12 Normal score plot 37
4.13 Residuals versus fitted values
4.14 Time series plot of actual, fits and forecasts

**List of Exhibits**
4.1 Autocorrelation function
4.2 PACF function of figure 4.2 above
4.3 Final Estimates of parameters
4.4 Ljung-Box test
4.5 Forecasting using model ARIMA (2,1,0)

**List of Appendices**
Appendix A: Raw data
Appendix B: ACF PACF figures 4.15 - 4.21
Appendix C: Minitab exhibits
Appendix C: ACF and PACF plots of tested different ARIMA models

**List of acronyms.**
ACF- Autocorrelation function
AIC- Akaike information criteria
AR- Autoregressive
ARIMA- Autoregressive integrated moving average
ARMA- Autoregressive moving average
BTFV- Burglary Theft from Vehicle
BUSE- Bindura University of Science Education

CCTV- Close circuit television

EGARCH- Exponential general autoregressive conditional heteroscedasticity

GARCH- Generalized autoregressive conditional heteroscedasticity

HBT- House breaking and theft

HOM- Hazardous Organic Mishap

IMA- Integrated moving average

KPSS- Kwiatkowski-Phillips-Schmidt-Shin

MA- Moving average

MAPE- Mean absolute percentage error

MSE- Mean square error

PACF- Partial autocorrelation function

SARS- Severe acute respiratory syndrome

TARCH- Threshold general autoregressive conditional heteroscedasticity
# Contents

APPROVAL FORM ........................................................................................................... i

RELEASE FORM ............................................................................................................. ii

Abstract ........................................................................................................................... iii

Acknowledgements ......................................................................................................... iv

Dedication ......................................................................................................................... v

List of Tables ..................................................................................................................... vi

List of Figures .................................................................................................................. vi

List of Exhibits ................................................................................................................. vii

List of Appendices .......................................................................................................... vii

List of acronyms ............................................................................................................... vii

Chapter 1 ......................................................................................................................... 1

  1.0 Background to the study ......................................................................................... 1

    1.1 Statement of the problem .................................................................................... 2

    1.2 Research Questions ......................................................................................... 2

    1.3 Assumptions ...................................................................................................... 2

    1.4 Significance of the study .................................................................................. 3

    1.5 Limitations ....................................................................................................... 3

    1.6 Definition of Terms .......................................................................................... 3

CHAPTER 2 ..................................................................................................................... 5

  2.0 Introduction ............................................................................................................. 6

    2.1 Literature Review of Related Research ......................................................... 6

    2.2 Theoretical Review .......................................................................................... 9

    2.3 Components of Time Series .......................................................................... 9

      2.3.2 Seasonal Variation ................................................................................... 10

      2.3.2 Cyclical Variation ................................................................. 11

      2.3.3 Irregular variation .................................................................................. 11

    2.4 Univariate Time Series Models ....................................................................... 12

    2.5 Common Approaches to Univariate Time Series ........................................ 12

      2.5.1 Decomposition .................................................................................... 12

      2.5.2 The spectral plot ................................................................. 12
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.3 Autoregressive (AR) Models.</td>
<td>12</td>
</tr>
<tr>
<td>2.5.4 Moving Average (MA) Models.</td>
<td>13</td>
</tr>
<tr>
<td>2.6 Box-Jenkins ARIMA Process</td>
<td>13</td>
</tr>
<tr>
<td>2.6.1 The Mixed Autoregressive-Moving Average Process</td>
<td>14</td>
</tr>
<tr>
<td>2.6.2 Stationarity</td>
<td>14</td>
</tr>
<tr>
<td>2.6.3 Detecting Stationarity</td>
<td>14</td>
</tr>
<tr>
<td>2.6.4 Invertibility</td>
<td>15</td>
</tr>
<tr>
<td>2.6.5 Random Walk</td>
<td>15</td>
</tr>
<tr>
<td>2.6.6 Seasonality</td>
<td>15</td>
</tr>
<tr>
<td>2.6.7 Detecting Seasonality</td>
<td>16</td>
</tr>
<tr>
<td>2.6.9 First Differencing</td>
<td>16</td>
</tr>
<tr>
<td>2.7 Model Building Strategy</td>
<td>16</td>
</tr>
<tr>
<td>2.7.2 Model Fitting</td>
<td>17</td>
</tr>
<tr>
<td>2.7.3 Maximum likelihood estimation method</td>
<td>17</td>
</tr>
<tr>
<td>2.7.4 Method of moments</td>
<td>18</td>
</tr>
<tr>
<td>2.7.5 Least squares estimation method</td>
<td>18</td>
</tr>
<tr>
<td>2.8 Diagnostic Checking</td>
<td>19</td>
</tr>
<tr>
<td>2.9 Conclusion</td>
<td>19</td>
</tr>
<tr>
<td>CHAPTER 3.</td>
<td>19</td>
</tr>
<tr>
<td>3.0 Research Methodology</td>
<td>19</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>19</td>
</tr>
<tr>
<td>3.2 Research Design</td>
<td>20</td>
</tr>
<tr>
<td>3.3 Instruments used to collect data</td>
<td>20</td>
</tr>
<tr>
<td>3.4 Data presentation and analysis procedures</td>
<td>20</td>
</tr>
<tr>
<td>3.5 Model identification</td>
<td>20</td>
</tr>
<tr>
<td>3.5.1 Time series plot</td>
<td>21</td>
</tr>
<tr>
<td>3.5.2 Stationarity</td>
<td>21</td>
</tr>
<tr>
<td>3.5.3 Checking Stationarity</td>
<td>21</td>
</tr>
<tr>
<td>3.6 Tools used to identify the values p (autoregressive component) and q (moving-average component)</td>
<td>21</td>
</tr>
<tr>
<td>3.6.1 Measuring autocorrelation</td>
<td>21</td>
</tr>
</tbody>
</table>
1.0 Background to the study

Crime occurrence ranging from violent crimes to property crimes continue to be on the rise in many parts of Zimbabwe. In Zimbabwe crimes like house breaking and theft, arson, malicious injury to property, murder, rape, robbery, and fraud are some common cases. The occurrence of such crimes has seen the community loosing property, sums of money and even lives. This research seeks to make an input by developing a statistical ARIMA model which will be used to analyze crime series and their future predictions to the challenges given to members of the Zimbabwe Republic Police and the general public. The evaluation of the patterns and duration of the crime series using the model, will assist the Police administration in reviewing their activities and operations by making informed and intelligent decisions on the basis of such analysis. The main purpose of this research is to develop a statistical model which will be used to analyze and forecast crime rate occurrences. Data used in this paper was collected from the Zimbabwe National Statistics Agency (ZimStat). Our data focused on House Breaking and Theft. Warr (2005) argues that crime rates appear to be the main source of information people use to establish their concern for crime and to determine how safe they feel across time. The researcher noted that not much has been done to apply statistical methods, especially time series analysis in describing and forecasting crime occurrence in Zimbabwe. Bohr, in Montgomery, Jennings and Kulahci (2008) postulated that it is difficult to make predictions, especially about the future. Crime records data may be meaningless unless descriptive statistics are used to arrange numbers in a coherent and more meaningful summary of information. Thus time series analysis is a statistical technique which can be used to analyze data when observations like counts of crime are made repeatedly. In this research it is hoped that predictable patterns create opportunities for crime prevention and detection like deployment of policing resources and designing crime prevention initiatives. This will contribute to peace and stability in Zimbabwe’s communities and also to the government’s national crime prevention strategy.
1.1 Statement of the problem
The research seeks to fit the best time series model (based on the AIC) which describe and forecast crime rate pattern in Zimbabwe. Good forecasts are needed by the home affairs department in order to design crime prevention initiatives and deployment of policing resources.

1.2 Research Questions.
The major aim of this study is to fit a time series model that can be used to forecast the future patterns of crime rate and reveal trends, basing information on the data collected from January 2008 to June 2014. Within the context of this larger objective, I seek to explore the following:

1. Identify a time series model that can be used to forecast the pattern of crime (house breaking and theft) in Zimbabwe, basing on data collected from Zimbabwe National Statistical Agency (ZimStat), from January 2008 to June 2014.

2. Identify and monitor primary patterns and trends for effective deployment of resources.

3. Use the model to forecast crime occurrence in Zimbabwe.

1.3 Assumptions
The following assumptions will be made in this project.

a) data was recorded and collected at uniform time intervals.

b) data collected accurately and is representing the area under study.
1.4 Significance of the study
1. The analysis of the crime rate might help the home affairs department to evaluate crime awareness campaigns in communities to see whether crime prevention programs are effective or not.

2. It will help facilitate decision making with regards to police patrols and crime detection.

3. The Zimbabwe Republic Police will be able to come with a good estimate of the policing resources.

1.5 Limitations
1) The ARIMA model produced in this research will be useful for short term forecasting, because they are simple models but good performing methods

2) Observations are never pure measures of phenomena. Inaccuracies in reporting lead to potential confounding in police reports. Outside influences can exert pressure on crime record keepers so that records reflect numbers of that politician, chief of police and other decision-makers would like to see because the doctored reports are professionally flattering rather than accurate counts of the crime taking place under their watch. External pressure may influence crime record keepers to inflate or deflate actual numbers.

1) Observations close together in time will be more closely related than observations further apart.
2) A major consideration in criminal research is the issue of measuring crime and understanding the characteristics of the data and measurement methodologies, as crime is measured from a variety of perspectives.

1.6 Definition of Terms.
1.6.1 Basic terms: Time series, Time series analysis, Crime and Crime Rate

a) Time series
Holden and Perman (1994) defined time series as a collection of observations or measurements on quantitative variables made sequentially or in a uniform set of time period, usually daily, weekly, monthly, quarterly or annually. Again, Pandit and Wu (1983) views time series as a statistical series which tells us how data has been behaving in the past and gives value of the variable we are considering at various points in time. Examples include total monthly crime for a jurisdiction for a period of ten years, daily stock prices of a firm for a period of one year, monthly electricity consumption for a household for a period of five years, etc.

b) Time Series Analysis

Time series analysis comprises methods or processes that break down a series into components and explainable portions that allow trends to be identified, estimates and forecasts to be made. Granger and Newbold (1974) assert that time series analysis attempts to understand the underlying context of the data points through the use of a model to forecast future values based on known past values. Such time series models include GARCH, EGARCH, TARCH, CGARCH and ARIMA, etc. The main focus of this study is based on ARIMA model by Box-Jenkins (1976).

c) Crime

Crime is a word commonly used to describe the actions that the laws of a community or a state deemed to be wrong. Cater (1990) argued that crime is a conduct which common or statute law prohibits and expressively or impliedly subjects to a punishment which is excusable by the state alone and which the offender cannot avoid by his own acts once he has been convicted. In addition, The Law Commission of Canada (2004) defines crime as something that is against the law.

d) Crime rate

These are statistical measures of crimes committed in society. They are used for both theoretical and practical purposes to analyze the causes of crime and its prevalence in different parts of society, examine success of law enforcement and criminal justice system. In their studies, Layson (1985), Witte (1980) and Grogger (1991) reported evidence that increasing criminal justice sanctions reduce criminal activity.
CHAPTER 2
Review of Related Literature
2.0 Introduction
This chapter of the research primarily deals with the review of empirical related literature of previous authors regarding time series analysis, crime and autoregressive integrated moving average (ARIMA) models. The literature review will basically be based on time series components that are trend, cyclic, seasonal, and irregular component. Generally, time series on crime forecasting is a new area of study in Zimbabwe, since not much literature has been published concerning it.

2.1 Literature Review of Related Research
Girard (2000) used an ARIMA model to analyze and assess the epidemiology situation of whooping-cough in England and Wales for the period 1940 to 1990. The ARIMA modeling of this illness contains intervention variables, such as the introduction of widespread vaccination in 1957 and the fall in the level of vaccination down to 30% in 1978. The results of the study confirmed the role of the intervention variables on the evolution of the morbidity due to whooping-cough, by quantifying their impact on the level of the morbidity, as well as the delay needed before they have an influence on the increase of recorded cases of whooping-cough. Bianchi, Jeffrey and Hanumara (1998) analyzed existing and improved methods for forecasting incoming calls to telemarketing centers for the purposes of planning and budgeting. They also analyzed the use of additive and multiplicative versions of Holt-Winters exponentially weighted moving average models and compare it Box-Jenkins (ARIMA) modeling. They determine the forecasting accuracy of HW and ARIMA models for samples of telemarketing data and concluded that ARIMA models with intervention analysis performed better.

Furthermore, Chung and Chan (2008) applied an autoregressive integrated moving average (ARIMA) to make weekly and daily forecasting of property crime for a city of China. It is shown that the model of AR (1) is suitable for crime sample distributing by week and IMA (1,1) by day. The mean absolute percentage error (MAPE) and magnitude relative error were taken as the error measurements for model fitting and forecasting. The results obtained proved that the model of AR (1) had higher accuracy in fitting and forecasting than BOA (1, 1). This result could be attributed to the crime stochastic difference between day and week. When forecasting for day, the crime
stochastic was strong, so it is hard to pick up the turning points. But for week, the stochastic of the crime was eliminated effectively. So, for short-term crime forecasting, it is better to make predictions for week than for day. Jennifer, Hsien-Hung and Liu (2010) applied Autoregressive Integrated Moving Average (ARIMA) to evaluate the impact of different local, regional and global incidents of a man-made, natural and healthy character, in Taiwan over the last decade. The incidents used in this study are the Asian financial crisis starting in mid 1997, the September 21st earthquake in 1999, the September 11th terrorist attacks in 2001, and the outbreak of Severe Acute Respiratory Syndrome (SARS) in 2003. Empirical results revealed that the SARS illness had a significant impact, whereas the Asian economic crisis, the September 21st earthquake and the September 11th terrorist attacks showed no significant effect on air movements.

Another scholar, Huang (2011) focused on identifying the patterns of the non-gun related crimes in 21 areas of Los Angeles based on exploratory data analysis, principle component analysis, cluster analysis, as well as Pearson’s $X^2$ statistics to discover unusual crimes in areas with the following conclusions: The percentages of total crime in 21 areas are almost the same but the percentages of a specific crime type in 21 areas differ a lot, and each crime type has its own pattern. BTFV is the crime type that happens most. HOM, ARSON and KIDS were rare. The distributions of 13 crime types’ frequency vary according to areas. Most crimes are most frequent in January and less frequent in February. Sanjeev, Suncica and Siem (2003) examined the impact of abolished parole and reformed sentencing for all felony offenders committed on or after January 1, 1995 by the Commonwealth of Virginia, considering structural time series models as an alternative to the Box-Jenkins ARIMA models that form the standard time series approach to intervention analysis. The study revealed limited support for the deterrent impact of parole abolition and sentence reform is obtained using univariate modeling devices, even after including unemployment as an explanatory variable.

In addition, Vaugh (2012) used ARIMA model in examining patterns of crime rate and concern for crime in the United States. The study conducted was to explore the relationship of concern for
The study finds support for the argument that people use violent crime rates to logically determine their concern for crime as opposed to using competing sources of information. The model revealed that crimes peak in the early 1990s except for murder, which spikes in 1980. Minimum rates are reached in 2010. Furthermore, Donohue and Levitt (2011) offered evidence that legalized abortion has contributed significantly to recent crime reductions in his study titled The Impact of Legalized Abortion on Crime using the ARIMA model. Crime began to fall roughly 18 years after abortion legalization. The 5 states that allowed abortion in 1970 experienced declines earlier than the rest of the nation, which legalized in 1973 with Roe V Wade. States with high abortion rates in the 1970s and 1980s experienced greater crime reductions in the 1990s. In high abortion states, only arrests of those born after abortion legalization fall relative to low abortion states. Legalized abortion appears to account for as much as 50 percent of the recent drop in crime.

Appiahene-Gyamfi (1998) studied and discussed the trends and patterns of robbery, and reactions to it in Contemporary Ghana between 1982 and 1993. The study revealed that robbery as a crime of opportunity appears to have been prevalent in pre-colonial times as well as during the subsequent period of slavery. Its trends and patterns however, have changed with the introduction of a monetary economy that has resulted in increased opportunities and targets for robbery. Shittu (2009) used the intervention analysis approach to model exchange rate in Nigeria in the presence of finance and political instability. Monthly exchange rate of Naira vis-à-vis US dollar from 1970 to 2004 was used on some identified intervention variables. The result showed that most of the interventions are pulse function with gradual and linear but significant impact in the Naira-Dollar exchange rates.

Pengetal (2008) used time series model of ARIMA to make short term forecasting of property crime for one city of China. With the given data of property crime for 50 weeks, an ARIMA model is determined and the crime amount of one week ahead is predicted. The model fitted and forecast results were compared with the SES and HES. It showed that the ARIMA model had higher fitting
and forecasting accuracy than exponential smoothing and therefore would be helpful for the local police stations and municipal governments in decision making and crime suppression. Chung, et al (2009) analyzed the impact of financial crisis on the manufacturing industry in China using data collected from March 2005 to November 2008 by the China Statistical databases of the national bureau of statistics in China. The intervention effect of the global financial crisis that began in September 2008 on China’s manufacturing industry, as measured in this study, was temporary and abrupt. The results again indicated that China’s manufacturing industry may have to tolerate a significant negative effect cased by the global financial crisis over a period of time, with its gross industry output value declining throughout 2008 and 2009 before reaching a state of equilibrium.

2.2 Theoretical Review
Based on the above literature reviewed by previous authors it can be concluded that ARIMA models with and without interventions are applied not only for criminal analysis (criminology), but rather to a wide variety of applications ranging from archaeology to zoology. The ARIMA Model developed by Box and Jenkins (1976) is found to be a powerful statistical technique in evaluating or assessing the impact of events, policies or programs on time series data.

2.3 Components of Time Series
A vital step in choosing appropriate modeling and forecasting procedure is to consider the type of data patterns exhibited from the time series graphs of the time plots. The sources of variation in terms of patterns in the time series data are mostly classified into four main components. These components include seasonal variation, trend variation, cyclic changes and the irregular (random) fluctuations.

2.3.1 Trend Variation
The trend is simply the underlying long term behavior or pattern of the data or series. The Australian Bureau of Statistics (ABS, 2008) viewed trend as the long term movement in a time series without calendar related and irregular effects and is a reflection of the underlying level. It is the result of influences such as population growth, price inflation and general economic changes.
Furthermore, Pankratz (1983) defines trend as a long term smooth underlying movement in a time series. Trend analysis can be viewed as a statistical technique used to isolate the underlying long-term movement. The pictorial graph is shown below in figure 2.1

![Economic Time Series Graph](image)

**Figure 2.1 Graphical Representation of Trend Variation (Source:)**

**2.3.2 Seasonal Variation.**
Charemza and Deadman (1997) define seasonal variation as fluctuations that are repeated periodically after a fixed time interval. Usually, seasonal variations are captured in a series of monthly or quarterly data, meaning one single yearly observation hardly reveals variations that occur during the year. In addition, seasonality in a time series can be identified by regularly spaced peaks and troughs which have a consistent direction and approximately the same magnitude every year, relative to the trend. Seasonal adjustment is the process of estimating and then removing from a time series influences that are systematic and calendar related. Observed data needs to be seasonally adjusted as seasonal effects can conceal both the true underlying movement in the series, as well as certain non-seasonal characteristics which may be of interest to analysts.

Other techniques that can be used to detect seasonality include:

1. A seasonal subseries plot is a specialized technique for showing seasonality.
2. Multiple Box plots can be used as an alternative to the seasonal subseries plot to detect seasonality.
3. The autocorrelation plot can help identify seasonality.

2.3.2 Cyclical Variation.
Cyclical variations are the short term fluctuations (rise and falls) that exist in the data that are not of a fixed period. They are usually due to unexpected or unpredictable events such as those associated with the business cycle sharp rise in inflation or stock price, etc. Cycles are medium to long term deviations from the trend. However, Greene (1997) revealed that it is not possible to make predictions using this component and with crime. This component is not applicable because these variations are generally thought to be caused by such factors as national and international market conditions, interest rates, money supply and earthquakes just to mention a few. Again, Hamilton (1994) noted that, cycles are difficult to measure statistically and their use in statistical forecasting is limited. The main difference between the seasonal and cyclical variation is the fact that the former is of a constant length and recurs at regular intervals, while the later varies in length. More so, the length of a cycle is averagely longer than that of seasonality with the magnitude of a cycle usually being more variable than that of seasonal variation.

Figure 2.2 Graphical Representation of Cyclic Variation (Source:)

2.3.3 Irregular variation.
The irregular component (sometimes also known as the residual) is what remains after the seasonal and trend components of a time series have been estimated and removed. Random fluctuations in
a time series are attributed to unpredictable occurrences such as wars, riots, strikes, boycotts or accidents. Since the future is unforeseen it is difficult to forecast. Random variations are naturally erratic rather than regular and can be seen as a residual factor after allowing for other causes of variation.

2.4 Univariate Time Series Models.
These are time series models with only one series of observations. The monthly crime data displayed in Appendix 1 is a typical example of a univariate time series. Univariate time series models usually view their series as a function of its own past, random shocks and time.

2.5 Common Approaches to Univariate Time Series.

2.5.1 Decomposition
It involves decomposing the time series into a trend, seasonal, and residual component. Decomposition refers to separating a time series into trend, cyclical, and irregular effects. Decomposition may be linked to de-trending and de-seasonalizing data so as to leave only irregular effects, which are the main focus of time series analysis. Triple exponential smoothing is an example of this approach. Kathori (2004) asserts that there is some form of random variation inherent in the collection of data taken over time. Smoothing is a method for reducing or cancelling effect due to random variation.

2.5.2 The spectral plot.
Another approach, commonly used in scientific and engineering applications, is to analyze the series in the frequency domain. An example of this approach in modeling a sinusoidal type data set is shown in the beam deflection case study. Jenkins and Watts (1968) and Chatfield (1996) discussed that, the spectral plot is the primary tool for the frequency analysis of time series.

2.5.3 Autoregressive (AR) Models.
Another common approach for modeling univariate time series is the autoregressive (AR) model. Cryer (1986) asserts that Yule was the first person to introduce AR processes. Autoregressive processes are regressions on themselves. An Autoregressive model is simply a linear regression of the current value of the series against one or more prior values of the series AR \( p \). The value of
\( p \) is called the order of AR model. In practice however \( p \) normally takes values 1 and 2 corresponding to AR (1) and AR (2) respectively. A \( p^{th} \) order AR process is given in the form of an equation as:

\[
Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + a_t \quad (1)
\]

Where \( \phi_1 \ldots \ldots \phi_p \) are parameters. A current value of the series \( Z_t \) is a linear combination of the most recent past values of itself plus an “innovation” term, \( a_t \), which incorporates everything new in the series at a time \( t \) that is not explained by the past values.

### 2.5.4 Moving Average (MA) Models

Cryer (1986) stated that present values may be regressed on their residuals forming Moving Average processes, which were first used by Shitzky around 1937. This is a smoothing technique which involves taking successive averages of groups of observations. It removes short term fluctuations in a time series. Each time period’s value is replaced by the average of observations which surround it. A general MA process of order \( q \) is given by the equation:

\[
Z_t = a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2} - \ldots - \Theta_q a_{t-q}. \quad (2)
\]

Moving Average arises from the fact that \( Z_t \) is obtained by applying the weights \( 1, -\Theta_1, -\Theta_2, \ldots - \Theta_q \) to the variables \( a_t, a_{t-1}, a_{t-2}, \ldots, a_{t-q} \) and the moving the same weight 1 unit of time forward and applying them to \( a_{t+1}, a_t, a_{t-1}, \ldots, a_{t-q+1} \) to obtain \( Z_{t+1} \). Just like \( p \) above, \( q \) also normally takes values 1 and 2 corresponding to MA (1) and MA (2) respectively.

A Moving Average model is conceptually a linear regression of the current value of the series against the white noise or random shocks of one or more prior values of the series. The random shocks at each point are assumed to come up from the same distribution, typically a normal distribution, with location at zero and constant scale.

### 2.6 Box-Jenkins ARIMA Process

In Statistics and Econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. The Box-Jenkins methodology named after the statisticians George Box
and Gwilym Jenkins (1976), applies ARIMA models to find the best fit a time series to past values of this time series, in order to make forecasts.

2.6.1 The Mixed Autoregressive-Moving Average Process
This is a combination of both the (MA\(q\)) and the (AR\(p\)) process. The process \(Z_t\) is mixed autoregressive moving average process of order \(p\) and \(q\) respectively. The features of the ARMA process match those of (AR\(p\)) and (MA\(q\)). \(Z_t\) is the crime variable. A general mixed \(Z_t\) ARMA process of orders \(p\) and \(q\), (ARMA\(p, q\)) is given by:

\[
Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2} - \ldots - \Theta_q a_{t-q},
\]  

(2)

2.6.2 Stationarity
A stationary process is one whose statistical properties are the same over time. Stationarity means that there is no growth or decline in the data roughly horizontal along the time axis. The data fluctuate around a constant mean, independent of time, and the variance of the fluctuations is essentially constant over time. A mixed ARIMA process is stationary if the AR part is stationary. Cryer (1986) underscores the importance of stationarity, by saying that if stationarity is not present, it must be incorporated into the time series model.

2.6.3 Detecting Stationarity
One of the very first issues a researcher must confront when analyzing a time series is the question of whether it is stationary. Stationarity can be assessed from a run sequence plot. The run sequence plot should show constant location and scale. It can also be detected from an autocorrelation plot. Specifically, non-stationarity is often indicated by an autocorrelation plot with very low decay. Phillips (1986) noted that when a time series is not stationary, classical statistical theory breaks down and special procedures are needed. Furthermore, Charemza and Deadman (1997) assert that “unit root” tests allow one to determine whether a series is stationary and if it is not, whether it is a random walk, a random walk with drift, or a random walk with drift and trend. These tests consider whether the coefficient \(a\) in an equation \(Y_t = \mu + a Y_{t-1} + \ldots + e_t\) is significantly different from 1.
Finally, unit root tests provide a more formal approach to determining the degree of differencing such as the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and Augmented Dickey-Fuller or Phillips–Peron Unit Root Tests are carried out employing the unit root testing procedures of Hamilton (1994). The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for the null hypothesis of a level stationary against an alternative of unit root. Augmented Dickey-Fuller and Phillips-Peron test for the null hypothesis of a unit root against the alternative of a stationary series.

Decision rule for the KPSS test: if the p-value of the test statistic is greater than the critical value of say 0.05, reject the null hypothesis of having a level stationary series and therefore conclude the alternative hypothesis that it has a unit root. Decision rule for the Phillips-Peron test: if the p-value is less than the critical value chosen, reject the null hypothesis of unit root in favour of an alternative hypothesis of stationarity. Decision rule for the Augmented Dickey-Fuller test: if the test statistic is more negative than the chosen critical value, we reject the null hypothesis of unit root in favour of an alternative hypothesis of no unit root.

2.6.4 Invertibility
Invertibility condition of the general order MA process is best expressed by the use of the backward shift operator B, defined by:

\[ B^jZ_t = Z_{t-j} \text{ for all } j. \]

\[ Z_t = \phi Z_{t-1} + a_t. \]

2.6.5 Random Walk
Let \( \{a_t\} \) be a discrete, purely random process with mean \( \mu \) and variance \( \sigma_a^2 \). The process \( Z_t \) is said to be a random walk if:

\[ Z_t = Z_{t-1} + a_t. \]

2.6.6 Seasonality
Holden and Perman (1994) view seasonality as periodic fluctuation. For example retail price sales tend to peak for Christmas season and then decline after the holiday. Again, Greene (1997) observes that it is useful to project past patterns into the future, basing on seasonal variation.
2.6.7 Detecting Seasonality.

Seasonality (or periodicity) can be assessed from an autocorrelation plot, a seasonal subseries plot or a spectral plot. At the model identification stage, the goal is to detect seasonality, if it exists, and to identify the order for the seasonal autoregressive and seasonal moving average terms. For many series, the period is known and a single seasonality term is sufficient. For example, for monthly data one would typically include either a seasonal AR 12 term or a seasonal MA 12 term. However, it may be helpful to apply a seasonal difference to the data and regenerate the autocorrelation and partial autocorrelation plots. This may help in the model identification of the non-seasonal component of the model. In some cases, the seasonal differencing may remove most or all of the seasonality effect.

2.6.8 Seasonal differencing.

Data series collected quarterly could have seasoned differences computed as follows:

\[ \Delta Z_t = Z_t - Z_{t-4} = (1-B^4)Z_t. \]

The general short hand notation is; ARIMA \((p, d, q)(P, D, Q)s\), Where

- \(d\) is non-seasonal part of the model.
- \(D\) is seasonal part of the model
- \(S\) is number of periods per season

The backward shift operator is convenient for describing the process of differencing, applied to make a non-stationary time series more nearly stationary.

2.6.9 First Differencing

\[ \Delta Z_t = Z_t - Z_{t-1} \]

2.7 Model Building Strategy

In the forecasting process, if we do not know the form of the model or the parameters, we can use certain techniques to obtain a forecast. Wagner (1995) questioned on how we decide on the models to use. Cryer (1986) suggested developing a model building strategy which was developed by Box and Jenkins in 1976. This is a powerful model building approach which is a combination of both the AR and the MA. Although both the moving average and the autoregressive approaches were already known, the contribution of Box and Jenkins was in developing a systematic methodology
for identifying and estimating models that could incorporate both approaches. However, the problem of building a stochastic Box-Jenkins model of a process involves determining the number of terms in the autoregressive and moving average parts of the model. There are three main steps in the Box-Jenkins procedure, namely model specification, model fitting and model diagnostics.

2.7.1 Model Specification/ Identification

In model specification we select classes of time series models that may be appropriate for a given observed series. Cryer (1986) defined model specification as the initial preprocessing of data to make the series stationary and also the identification of suitable orders of p and q for the ARMA components of the model. The Box-Jenkins approach is applicable to stationary time series data. Non-stationary time series data needs to be made stationary first. In this step, we look at the time plot of the data which will make it easier for the forecaster to note that the data is stationary or not. Computations of autocorrelations and partial autocorrelations are done and examine their patterns. Autocorrelations plots expose non-stationarity and measures the correlation between time series values separated by a fixed number of periods called lags. The partial autocorrelation between the original time series and the same series moved forward a fixed number of periods, holding the other lagged time fixed. This is done by using computer software MINITAB 14. From the ACF and the PACF we can identify the form of the possible model. The model chosen must employ the principle of parsimony.

2.7.2 Model Fitting

After identifying the order of the tentative model ARIMA (p, d, q), the parameters of the model are estimated using method of moments, least squares estimation and the maximum likelihood estimation methods.

2.7.3 Maximum likelihood estimation method

Let $a_t$ be independent and normally distributed with mean zero and variance $\sigma_a^2$. The density function of $a_t$ is given by $f(a_t) = (2\pi\sigma_a^2)^{-0.5}\exp[-a_t^2/2\sigma_a^2]$
The corresponding likelihood function is give by

\[
L = L(\Phi, \mu, \Theta, \sigma_a^2) = (2\pi \sigma_a^2)^{n/2} \exp[-1/2 \sigma_a^2 \sum a_t^2]
\]

The log likelihood function is given by

\[
\ln L = -n/2 \ln(2\pi \sigma_a^2) - 1/2 \sigma_a^2 \sum a_t^2 \quad \text{where } t=1,2,\ldots,n.
\]

The conditional log likelihood function is given by

\[
\ln L^*(\Phi, \mu, \Theta, \sigma_a^2) = -n/2 \ln(2\pi \sigma_a^2) - 1/2 \sigma_a^2 \sum a_t^2 (\Phi, \mu, \Theta|Z^*, a^*, Z) \quad \text{where } t=1,2,\ldots,n
\]

\[
= -n/2 \ln(2\pi \sigma_a^2) - 1/2 \sigma_a^2 [S^*(\Phi, \mu, \Theta)] \quad \text{where } S^*(\Phi, \mu, \Theta) = \sum a_t^2 (\Phi, \mu, \Theta|Z^*, a^*, Z)
\]

where \( t=1,2,\ldots,n \).

2.7.4 Method of moments
This method consists of equating sample moments such as the sample mean, sample variance and sample autocorrelation function to theoretical counterparts and solving the resultant equations. Wei (1990) observes that moments estimators are very sensitive to rounding errors, meaning they are not particularly good. Again, Abraham and Ledolter (1983) highlighted that, this method is not very useful when dealing with MA or mixed ARIMA processes, but useful when dealing with AR processes through the use of Yule-Walker equations.

2.7.5 Least squares estimation method
This is an estimation procedure developed for standard regression models and can only be used provided the assumptions on the error term \( a_t \) hold, that is, zero mean and constant variance, non-autocorrelation and uncorrelated with explanatory variable \( X_t \).

The backbone of the least squares error method (LSE) is minimizing error sum of squares. This method can be used for both AR and MA processes.

Remark: In this research, we will use the statistical package called MINITAB 14 to estimate the model parameters. A goodness of fit of the model test is done once model parameters have been estimated. A good model should obey the principle of parsimony.
2.8 Diagnostic Checking.
The diagnostic stage of the Box-Jenkins ARIMA process is to examine whether the fitted model follows a white noise process. Model diagnostics is primarily concerned with testing the goodness of fit of a tentative model. Basically, we should consider two complementary approaches namely analysis of residuals from fitted models and analysis of over-parameterized models. If both the residuals and the estimated parameters behave as expected under the presumed, then the model appears validated. If they do not, then the model should be modified and the procedure repeated until a model is validated. When the model passes the diagnostic tests, it can then be expanded in a more conventional regression equation in order for it to be used for calculating step-ahead forecasts. Forecasts will be generated by the Statistical software package, MINITAB 14.

2.9 Conclusion
A stochastic, time series ARIMA model will be used to analyze the dynamics of changes, variations and interruptions in the crime situation in Zimbabwe through time series data.

CHAPTER 3.

3.0 Research Methodology

3.1 Introduction
This chapter focuses on how the study is going to be conducted, the statistical tools and computer packages to be used. The researcher shall use the Box and Jenkins approach. This ARMA model has three phases namely Model Identification, Estimation, Model Testing and Model Application.
3.2 Research Design
This research is based on crime statistics of House breaking and Theft obtained from the Zimbabwe National Statistics Agency (ZIMSTAT) Quarterly digest of statistics 2014 and the compendium of statistics 2012. The ordered monthly data, coupled with the fact that it was collected over a fairly long period of time from January 2008 to June 2014 necessitates the use of Box – Jenkins ARMA model. However, the analysis will only consider data to be used in updating forecasts and testing the forecasting power of the fitted tentative model.

3.3 Instruments used to collect data
There are many statistical computer packages for analyzing a time series data. Due to limitations of such computer packages, the researcher will use the MINITAB 14 because it is user friendly and is highly efficient.

3.4 Data presentation and analysis procedures
The researcher shall use the Box and Jenkins approach (a simple version of the ARMA model) postulated by George Box and Gwilym Jenkins (1976). The approach has three phases namely:

3.5 Model identification
This phase has three main activities namely problem definition, data collection and data analysis. Problem definition involves developing understanding of how the forecast will be used along with the expectations of the “customer” (user of the forecast). Much of the ultimate success of the forecasting model in meeting the customer expectations is determined in the problem definition phase. Data collection consists of obtaining the relevant history for the variables that are to be forecast, including historical information on potential predictor variables. During this phase, planning on how data collection and storage issues in the future will be handled so that the reliability and integrity of the data will be preserved. Lastly, data analysis is an important preliminary step to selection of the forecasting model to be used. Time series plots of the data should be constructed and visually inspected for recognizable patterns, such as trends and seasonal or other cyclical components. This information will usually suggest the initial types of quantitative
forecasting methods and models to explore. For the ARMA model (p, d, q), the middle element d is investigated first. The goal is to determine if the data is stationary.

### 3.5.1 Time series plot
Firstly have a time series plot of original data. Time series data will be plotted graphically with crime monthly totals on vertical axis and time in months on horizontal axis. Inspection of the plot helps to give initial impression about the presence of seasonal pattern or trend. Time series plot helps in deciding whether the mean and variance of the realization are stationary i.e. N (0, σ²).

### 3.5.2 Stationarity
Stationarity refers to a constant mean and variance. If the mean is changing (that the trend is going up or down), the trend can be removed by differencing. If the variability is changing, the data may be made stationary by logarithmic transformation (the LOG formula in excel) Ascertained variance stability after transformation by plotting it again.

### 3.5.3 Checking Stationarity
Once the mean and variance are as constant as possible, assign d a value, for example if d=0, the data are stationary.

If d=1 the data needs to be differenced once so that the linear trend is removed.

If d = 2 the data need to be differenced twice so that the linear and quadratic trends are removed. Once the series accepted is stationary, identify the values of p and q.

### 3.6 Tools used to identify the values p (autoregressive component) and q (moving-average component)
Main tools used are the ACF and PACF plots. (Autocorrelation coefficient function of the data and the partial autocorrelation of the data). A correlogram may be used to assess whether autocorrelation are present in the time series data.

#### 3.6.1 Measuring autocorrelation
If adjacent observations are highly correlated, then so too will those that are say two or three observations apart. We use an autocorrelation function (ACF) in conjunction with partial autocorrelation function (PACF) as a control measure.
3.6.2 Removing Non-Stationarity
This is done by differencing and logarithmic transformation. Differencing takes care of non-stationarity in the mean while logarithmic transformation accounts for non-stationary in the variance. Differencing is done by calculating the difference among pairs of observations at time interval. Letting $W_t = Z_t - Z_{t-1}$, for $t=2, 3, 4...$. Normally the first and second difference are sufficient to transform the data to stationary. However, logarithmic transformation was not used in this research because it uses the Log formula in Excel while the researcher is using MINITAB.

3.6.3 Seasonality
Seasonality mean periodic fluctuations for example, crime rate tend to peak during festive periods and decline after it. Seasonality can only be determined when the data is stationary. Differencing the time series will not most likely expose seasonality since no autocorrelation coefficients except the seasonal ones are significantly different from zero.

3.7 Model Fitting and Validation

3.7.1 Model Fitting.
Models are estimated through patterns in their AFCs and PACFs. Estimates of the unknown parameters will be determined. Several methods to estimate models like parameter maximum likelihood estimation, methods of moments and least squares method can be used. In this case a statistical computer package MINITAB will be used. Basically it involves determining the number of terms in the autoregressive and moving average parts of the model.

3.7.2 First Order Regressive Process: AR(p)
Large value for lag 1 with the ACF decaying over greater lags and the PACF only being significant for lag 1.

1. When $p = 0$ there is no relationship between adjacent observations
2. When $p = 1$ there is relationship between observations at lag 1
3. When \( p = 2 \) there is a relationship between observation at lag 2

### 3.7.3 A Moving Average Process: MA (q)

Large value for lag 1, and the ACF will only be significant for this lag whilst the PACF decays over greater lags.

1. When q=0, there are no moving average components.
2. When q=1, there is a relationship between the current score and the random shock at lag 1.
3. When q=2, there is a relationship between the current score and the random shock at lag 2.

### 3.8 Residual Analysis

After a tentative ARMA model has been fitted various diagnostic checks to test its adequacy as a stochastic representation of the process under study must be noted. A good model identified and chosen should satisfy the following:

1. Principle of parsimony i.e. should have smallest number of parameters.
2. \( Z_t = \phi Z_{t-1} - \Theta_1 A_{t-1} + A_t \) for ARMA(1,1) i.e. free from any mis-specification errors.
3. ARMA assumptions of stationarity are invertibility and the residuals should be normally, independent and identically distributed i.e. \( \sim \mathcal{N}(0, \sigma^2) \)

### 3.8.1 Residual plot

A residual plot is constructed to assist us in whether to accept or reject the model. If plot of residuals against the fitted values show no pattern observed, there may be “heteroscedasticity” in the errors. We reject the model and design an appropriate one. A logarithmic or square root transformation may be required to overcome this. The model is accepted when the residuals show a pattern. Autocorrelation in the ACF plot of residuals suggest that there is information that has not been accounted for in the model. We can also test autocorrelation using the Durbin-Watson test of significant lag 1 or the Breusch-Godfrey test designed to look for significant higher-lag autocorrelation.
3.8.2 Normality test.
Check the histogram plots of residuals frequencies against corresponding residuals. The Normality test will be passed if the histogram is almost symmetric and bell-shaped otherwise the test maybe rejected. (MINITAB)

3.8.3 The Box-pierce Q-statistic / Portmanteau Test
This is a test for residuals in the time series.

H₀: the residuals are uncorrelated.
H₁: the residuals are correlated

Test statistic: \( Q = n(n+2)\sum_{n=1}^{k} r_n^2/n-k \)

Critical value: \( X^2_{k-p-q, \alpha} \)

Where \( k \) is the maximum lag considered, \( n \) the number of observation. \( Q \) has the asymptotic \( X^2 \) distribution with \( k-p-q \) degrees of freedom and \( p \) is the order of moving average process. If \( Q > X^2_{k-p-q, \alpha} \), the model is rejected as it will not be adequate. This means that the test is significant and its corresponding null hypothesis (H₀) is rejected if the p-value is less than the chosen critical values of 0.05

3.8.3 Model application
If the identified model passes all tests it can now be expanded in a more conventional regression equation in order for it to be used for calculating step forecasts. The computer package MINITAB used for identifying and estimating an ARIMA model in this study has an option to generate forecasts from the estimated model. Hence the forecasts will be generated by the computer package MINITAB.
CHAPTER 4

4.0 Data Presentation, Analysis and Interpretation.

4.1 Introduction

This chapter deals with the analysis of the crime statistics data obtained from Zimbabwe statistical agency, (ZimStat), Quarterly digest and Compendium of statistics, for the crime House Breaking and Theft (HBT) from 2008 to 2014. Table 4.1 comprise of the monthly crime statistics totals as shown below.

Table 4.1 monthly crime HBT statistics in Zimbabwe from 2008 to 2014.
<table>
<thead>
<tr>
<th>MONTH</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>381</td>
<td>3249</td>
<td>3127</td>
<td>2601</td>
<td>2985</td>
<td>2650</td>
<td>2930</td>
</tr>
<tr>
<td>February</td>
<td>273</td>
<td>3694</td>
<td>1530</td>
<td>2665</td>
<td>2952</td>
<td>2740</td>
<td>2883</td>
</tr>
<tr>
<td>March</td>
<td>610</td>
<td>3402</td>
<td>2022</td>
<td>2820</td>
<td>2898</td>
<td>3211</td>
<td>2887</td>
</tr>
<tr>
<td>April</td>
<td>497</td>
<td>3322</td>
<td>1648</td>
<td>2660</td>
<td>2603</td>
<td>2623</td>
<td>2818</td>
</tr>
<tr>
<td>May</td>
<td>703</td>
<td>2873</td>
<td>1866</td>
<td>2885</td>
<td>2590</td>
<td>2547</td>
<td>2625</td>
</tr>
<tr>
<td>June</td>
<td>875</td>
<td>2718</td>
<td>2111</td>
<td>2653</td>
<td>2355</td>
<td>2420</td>
<td>2426</td>
</tr>
<tr>
<td>July</td>
<td>1095</td>
<td>2593</td>
<td>1520</td>
<td>2601</td>
<td>2377</td>
<td>2393</td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>1340</td>
<td>2538</td>
<td>2201</td>
<td>2730</td>
<td>2324</td>
<td>2563</td>
<td></td>
</tr>
<tr>
<td>September</td>
<td>1443</td>
<td>2613</td>
<td>1913</td>
<td>2615</td>
<td>2501</td>
<td>2529</td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>1924</td>
<td>3057</td>
<td>1976</td>
<td>2661</td>
<td>2573</td>
<td>2534</td>
<td></td>
</tr>
<tr>
<td>November</td>
<td>2500</td>
<td>2709</td>
<td>2809</td>
<td>2812</td>
<td>2541</td>
<td>2607</td>
<td></td>
</tr>
<tr>
<td>December</td>
<td>2508</td>
<td>3029</td>
<td>2291</td>
<td>2852</td>
<td>2638</td>
<td>2842</td>
<td></td>
</tr>
</tbody>
</table>

In model identification, the first step is to have a time series plot of original data. Inspection of the plot helps to give initial impression about the presents of seasonal pattern or trend.
The figure 4.1 above clearly indicates that House breaking and theft series in Zimbabwe were not constant but rather varied from one month to the other as well from one year to the other with no systematic visible pattern, structural breaks, outliers and no identifiable trend components in the time series data. Visible inspection of the graph seems reveal an ascending mean and a roughly constant variance. The histogram of the time series data is shown in figure 4.2 below.

![Trend analysis of HBT crime series](image)

Figure 4.1 Trend analysis of HBT crime series.

![Histogram of HBT crime series data](image)

Figure 4.2 Histogram of HBT crime series data
The histogram plot of residuals is not symmetrical and does not portray a bell-shaped appearance. This means the normality assumption of the error term is violated, hence the data is non-stationary. Further analysis of stationarity is investigated by examining the ACF and PACF functions shown below in figure 4.3

Figure 4.3 ACF plot of HBT crime series data.

The computed ACF in figure 4.3 drops off slowly towards zero, meaning non-stationary. The ACFs are statistically significant up to lag 3 and so are not cutting off to zero. This tells us that we cannot fit an autoregressive model at this stage. They are becoming statistically insignificant at lag 4. This may suggest a non-stationary mean. The data needs to be transformed by differencing, to make it stationary.

**Exhibit 4.1: Autocorrelation Function: data**

<table>
<thead>
<tr>
<th>Lag</th>
<th>ACF</th>
<th>T</th>
<th>LBQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.835445</td>
<td>7.38</td>
<td>56.56</td>
</tr>
<tr>
<td>2</td>
<td>0.719077</td>
<td>4.10</td>
<td>99.02</td>
</tr>
<tr>
<td>3</td>
<td>0.586389</td>
<td>2.80</td>
<td>127.63</td>
</tr>
<tr>
<td>4</td>
<td>0.392505</td>
<td>1.71</td>
<td>140.62</td>
</tr>
<tr>
<td>5</td>
<td>0.286389</td>
<td>1.20</td>
<td>147.61</td>
</tr>
<tr>
<td>6</td>
<td>0.135659</td>
<td>0.56</td>
<td>149.21</td>
</tr>
<tr>
<td>7</td>
<td>-0.044314</td>
<td>-0.18</td>
<td>149.48</td>
</tr>
<tr>
<td>8</td>
<td>-0.147221</td>
<td>-0.60</td>
<td>151.44</td>
</tr>
<tr>
<td>9</td>
<td>-0.189209</td>
<td>-0.77</td>
<td>154.73</td>
</tr>
<tr>
<td>10</td>
<td>-0.229802</td>
<td>-0.93</td>
<td>159.65</td>
</tr>
<tr>
<td>11</td>
<td>-0.286800</td>
<td>-1.15</td>
<td>167.42</td>
</tr>
</tbody>
</table>

The ACFs for House breaking and theft series of figure 4.3 above, displayed in exhibit 4.1 above shows a large positive significant spike at several lag 1 with ACF of 0.835445 and a corresponding
T statistic of 7.38 and lag 3 with ACF of 0.586389 and a corresponding T statistic of 2.80. Again, in checking for stationary series using ACF, one commonly used rule is that the t-statistic greater in absolute value than 2 indicates that the corresponding ACF is not equal to zero, or significantly different from zero. Since the T-statistic for lag 1 and 3 are greater than 2 in absolute value it follows that the corresponding ACF of 0.835445 for lag 1 and 0.586389 are significantly different from zero. Furthermore, the ACFs for the highest lags that is 5,6,8,9,10 and 12 do not tend or approximate to zero indicating a typical case of a non-stationary series since, for a stationary process Harvey (1993) revealed that the main feature of the correlogram is that the autocorrelations tend towards zero as the lag increases.

Figure 4.4 PACFplot of HBT crime series data.

The PACF plot of HBT series in figure 4.4 is not different from that of its ACF on figure 4.3 above, thereby depicting a series that is either non-stationary in the mean and perhaps its variance. Therefore, requires some form of differencing to make it stationary.

<table>
<thead>
<tr>
<th>Lag</th>
<th>PACF</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.835445</td>
<td>7.38</td>
</tr>
<tr>
<td>2</td>
<td>0.069888</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>-0.102356</td>
<td>-0.90</td>
</tr>
<tr>
<td>4</td>
<td>-0.300203</td>
<td>-2.65</td>
</tr>
<tr>
<td>5</td>
<td>0.116650</td>
<td>1.03</td>
</tr>
<tr>
<td>6</td>
<td>-0.161969</td>
<td>-1.43</td>
</tr>
<tr>
<td>7</td>
<td>0.040868</td>
<td>0.36</td>
</tr>
<tr>
<td>8</td>
<td>-0.070403</td>
<td>-0.62</td>
</tr>
<tr>
<td>9</td>
<td>-0.093038</td>
<td>-0.82</td>
</tr>
<tr>
<td>10</td>
<td>-0.010231</td>
<td>-0.09</td>
</tr>
<tr>
<td>11</td>
<td>0.005105</td>
<td>0.05</td>
</tr>
<tr>
<td>12</td>
<td>-0.147252</td>
<td>-1.30</td>
</tr>
</tbody>
</table>

The PACF of HBT series displayed in exhibit 4.2 above shows large positive significant spikes at several lag 1 with PACF of 0.835445 and a corresponding T-statistic of 7.36 and lag 4 with PACF
of -0.300203 and a corresponding T-statistic of -2.65. Since the T-statistic for lag 1 and 4 are greater than 2 in absolute value, it follows that the corresponding PACF of -3.300203 for lag 1 and -2.65 for lag 4 are significantly different from zero, likewise the PACF for higher lags indicating a typical case of a non-stationary series. This suggests that we cannot fit an autoregressive process at this stage. A transformation by differencing may be required to induce stationary mean. The graph of differenced data versus period in month is shown in figure 4.5.

### 4.2 Unit Root Tests for checking stationarity.

#### 4.2.1 Original data

<table>
<thead>
<tr>
<th>Type</th>
<th>Test statistic</th>
<th>Lag</th>
<th>Critical value(5%)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPS</td>
<td>-3.193205</td>
<td>5</td>
<td>-2.899619</td>
<td>0.0242</td>
</tr>
<tr>
<td>ADF</td>
<td>-2.145450</td>
<td>3</td>
<td>-2.901217</td>
<td>0.1615</td>
</tr>
</tbody>
</table>

Table 4.3 above presents the adf test for the null hypothesis of a unit root against an alternative hypothesis of a stationary series together with the Phillips-Perron test for the null hypothesis of a unit root against the alternative of a stationary series. The KPSS test statistic of 0.089 with critical value of 0.05 as presented in table 4.3 reject the null hypothesis of having a level stationary series and therefore conclude the alternative hypothesis that it has a unit root. Phillips-Perron test on the other test statistic and its p value fails to reject the null hypothesis of a unit root at 5% significance level since its p value of 0.0242 was greater than -2.899619

In conclusion it is clear from the time series plot of crime series data the ACF and PASF with their graphical displays and the objective test, the series has to be differenced to stationarize the data.

#### 4.2.2 First order differenced data

<table>
<thead>
<tr>
<th>Type</th>
<th>Test statistic</th>
<th>Lag</th>
<th>Critical value(1%)</th>
<th>Critical value(5%)</th>
<th>Critical value(10%)</th>
<th>P-value</th>
</tr>
</thead>
</table>


Table 4.4 above represents the Kwiatkowski-Phillips-Schmidt-Shin(KPSS) test for the null hypothesis of a level stationary against an alternative of unit root together with the Phillips-peron test for the null hypothesis of a unit root against the alternative of a stationary series. The KPSS test statistic of 0.664308 which is greater than the critical value of 0.146000 as presented in above table do not reject the null hypothesis of having a level stationary series. Phillips-Peron test on the other hand test statistic of -5.08581 and its p value of 0.0083 reject the null hypothesis of a unit root at 5% significance level since its p value is less than the test statistic -5.08581. It is therefore concluded that the time series plot of the differenced crime series data, and the objective test indicate that the series is stationary.

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p-value</th>
<th>Critical Value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPS</td>
<td>-5.08581</td>
<td>0.0083</td>
<td></td>
<td>Reject</td>
</tr>
<tr>
<td>ADF</td>
<td>-3.593942</td>
<td>0.0093</td>
<td></td>
<td>Reject</td>
</tr>
<tr>
<td>KPSS</td>
<td>0.664308</td>
<td></td>
<td>0.146000</td>
<td>Not Reject</td>
</tr>
</tbody>
</table>

The Null hypothesis which state that there is a unit root against the alternative for no unit root was tested. Test Statistic: -3.193205 with p-value = 0.0242. Critical value at 1% significance level = -3.517847. Decision rule: we fail to reject H₀ at 1% significance level since the test statistic -3.193205 is greater than the critical value -3.517847 (i.e p-value of 0.0242 is greater than 0.01the critical value). Conclusion: we conclude that the data has a unit root hence non-stationary.

Figure 4.5 Time series plot of differenced data

In model identification, initial step is graphical analysis of the data that would suggest whether a series is likely to be stationary or non-stationary. The saw-toothed feature on the time series plot in figure 4.5 confirms that the data is indeed a time series data. From the trend it is clear that the series is reasonably stationary and there is no need for second differencing. The crime data were plotted against time to determine trend of HBT levels over the six and half year period. Visual inspection of the time series plot reveals that the differenced series fluctuates around zero and so
the mean is largely stationary and the variance is roughly constant hence differencing was necessary. The ACF of differenced data is shown in figure 4.6.

Figure 4.6. ACF plot of differenced data

The ACF plot in figure 4.6 shows three significant spikes, highest spike at the fourth lag suggesting that an Autoregressive model can be fitted on the crime data since there is clear evidence that the ACFs cuts off to statistical insignificance at lag 4 (decaying to zero) as the lag increases. The PACF of differenced data is shown in figure 4.7.
Figure 4.7. PACF plot of differenced HBT Data

The PACF plot in figure 4.7 shows that there are 3 significant spikes at lag1, lag 3 and lag 4 with the highest significant spike at lag 3. This means that we can fit a tentative model of AR (3). The spike in the PACF tells the researcher that the order of the model AR (p) is AR (3). The first, third and fourth lags are statistically significant since they are outside the confidence limits. The model was discovered to be good as the plot is decaying to zero. The final fitted model is ARIMA (2,1,0) because the data was differenced once to make it stationary.

4.2.3 Model Selection.
Firstly, i selected an ARIMA (3,1,2), ARIMA(3,1,1),and ARIMA(3,1,0) models basing my selection on the ACF and PACF plots on Appendix 2 respectively. From the study of the sampling statistics and residual analysis, the models were rejected on the basis of unsuitable optimum coefficient values since the standard errors were large and the correlation coefficient between the optimum coefficient values were also large.(see appendix 2). Again, other models were also tested and dropped for example ARIMA(2,1,1), the ACFs were statistically significant with 2 spikes at lag 1 and lag 4. L.ag 1 with an ACF of -0.387430 and lag 4 with ACF of -0.352024. These ACFs and PACFs lies outside the confidence limits of $\rho_k= (-0.2, +0.2)$, so were dropped on the basis of sampling statistics and residual analysis (Appendix 3).Finally, I tentatively chose an ARIMA (2,1,0) model to represent the data basing my proceedings on the principle of parsimony.
4.2.4 ACF and PACF of the Tentative model ARIMA(2,1,0).

Figure 4.8. Residuals ACF plot of tentative model (2,1,0)

Figure 4.9. PACF plot of tentative model ARIMA (2,1,0).

4.3 Parameter Estimation
Parameters were estimated using MINITAB 14 software package shown in exhibit 4.3: model ARIMA (2,1,0)
Exhibit 4.8: Final Estimates of parameters

* WARNING * Back forecasts not dying out rapidly

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SECoef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAR 12</td>
<td>-1.0895</td>
<td>0.0844</td>
<td>-12.91</td>
<td>0.000</td>
</tr>
<tr>
<td>SAR 24</td>
<td>-0.9227</td>
<td>0.0734</td>
<td>-12.58</td>
<td>0.000</td>
</tr>
<tr>
<td>Constant</td>
<td>53.61</td>
<td>38.06</td>
<td>1.41</td>
<td>0.164</td>
</tr>
</tbody>
</table>

Differencing: 0 regular, 1 seasonal of order 12
Number of observations: Original series 78, after differencing 66
Residuals: SS = 6010821 (backforecasts excluded)
           MS = 95410  DF = 63

4.4 Model Diagnosis: (residual analysis)

This is to check how well the model fits, by analyzing the residuals. After I have chosen ARIMA (2,1,0) model as my tentative model as opposed to other models, the model adequacy is further checked by performing tests using the Ljung-Box Test coupled with analysis of the residuals as reported in tables and figures. The model adequacy is checked to draw empirical conclusions regarding the model as good fit and for that matter its usage in estimation and forecasting.

4.4.1 Hypothesis testing using the Ljung-Box Test. (Portmanteau test).

Exhibit 4.9: Ljung-Box test

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>16.2</td>
<td>33</td>
<td>0.063</td>
</tr>
<tr>
<td>24</td>
<td>24.5</td>
<td>45</td>
<td>0.271</td>
</tr>
<tr>
<td>36</td>
<td>29.1</td>
<td>63</td>
<td>0.663</td>
</tr>
<tr>
<td>48</td>
<td>33.4</td>
<td>95</td>
<td>0.899</td>
</tr>
</tbody>
</table>

H0: p-value > 0.05 (The residuals are uncorrelated)

H1: p-value < 0.05 (The residuals are correlated)

Test statistic: Q = 16.2 with p-value = 0.063 (MINITAB table)
Decision rule: We fail to reject $H_0$ since $Q = 16.5$ with $p$-value = 0.063 greater than 0.05 (chosen value).

Conclusion: We conclude that the residuals appear to be uncorrelated. This indicates that the residuals of the fitted AR(2) model are a white noise, and for that matter, the model fits the series well, so you can use this model to make forecasts.

Figure 4.10 Residuals versus observation order.

From the residual plot in figure 4.10 we observe that the series is stationary. This means that there is no need to take any transformation to make it stationary. The fitted regression line is horizontal at zero, meaning the gradient of the trend line is equal to zero.
Figure 4.11 Histogram of residuals.

The histogram of the residuals in figure 4.11 is reasonably symmetrical. This means that the normality assumption of the error term is not violated.

Figure 4.12 Normal score plot

The normal score plot in figure 4.12 is reasonably straight, again showing minor bumps at the extremes. Overall, the model given is adequate since the residuals are independent and normally distributed.
The model fitted is accepted as good since the plot of the residuals against fitted values in figure 4.13 shows a pattern observed. This means that there is no heteroscedasticity in errors.

Visual inspection from figure 4.14 reveals that the differenced series fluctuates around zero and so the mean, the variance and the covariance is roughly constant, which confers a stationary series.
4.5 Tentative Model: ARIMA (2,1,0)

4.5.1 Forecasting using a Tentative Model ARIMA(2,1,0).
It is possible to forecast future values of the series. Exhibit 4.10 shows the forecasted values and
the 95% confidence intervals of the estimated values and the actual values. Forecasts from period
79 to 85 are shown in exhibit 4.10. The actual values of the forecasts are within the 95% confidence
interval.

Exhibit 4.10: Forecasting using model ARIMA (2,1,0)

Correlation matrix of the estimated parameters

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.563</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.043</td>
<td>-0.023</td>
</tr>
</tbody>
</table>

Forecasts from period 79

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>41.346</td>
<td>-564.190</td>
<td>646.883</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>-127.140</td>
<td>-732.677</td>
<td>478.396</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>0.397</td>
<td>-605.139</td>
<td>605.934</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>73.835</td>
<td>-531.701</td>
<td>679.371</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>-21.562</td>
<td>-627.098</td>
<td>583.974</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>63.226</td>
<td>-542.310</td>
<td>668.762</td>
<td></td>
</tr>
</tbody>
</table>

SUMMARY
From the graphical displays and objectives tests, the model has proved to be tentative. It is possible
to forecast future values of the series from the above forecasting analysis. The Box-Jenkins
ARIMA models basically produce better models when leading time is short. The actual values of
the forecasts are within 95% confidence interval. Above all, since the data have been differenced,
forecasts are based in part on the past values of crime series data. It can however be safely
concluded that the model tries to adequately represent the data.
Chapter 5
Summary, Conclusions and Recommendations

5.0 Introduction
This chapter of the study begins with summary of the main findings whose purpose is to reveal the relationships among the observed crime series data. This chapter deals with the conclusions drawn from the evidence provided in the study and recommendations made from aspects that emerged and appropriate suggestions concerning further research in such areas.

5.1 Summary
The time series plot of the original data reveals an ascending mean and a roughly constant variance, depicting a non-stationarity in the data hence calling for data differencing. A logarithmic transformation is not necessary since the variance is stable. The histogram of the original data is not symmetrical and is not portraying a bell-shaped appearance. This violated the normality assumption of the error terms. Again, further investigations of the computed ACF and PACF drops off slowly towards zero, suggesting a non-stationary mean. Furthermore, time series plot of differenced data portrays a horizontal appearance about zero showing that the mean has been made stationary and the variance as well. The histogram of the residuals in the differenced data is symmetrical. Again, the normal score plot is reasonably straight, depicting that the model given is adequate since the residuals are independent and normally distributed. The plot of residuals versus fitted values show a pattern observed, meaning that there is no heteroscedasticity in the errors. The forecasted values are also found to be within the 95% confidence interval, confirming a good model.

In addition, the ACF and the PACF of the differenced data series suggest an AR(2) process, and can be expressed as: $Z_t = 53.61 - 1.0895Z_{t-1}$. Above all, the study of the sampling statistics and analysis of the residuals gives an impression that the model is valid for the data.

5.2 Conclusion
Based on the literature reviewed, as well as the summary of the findings made from the analysis, a conclusion can be made that the model represented by the equation $Y_t = 0.0844Z_{t-1} + 38.06$ provides a good representation of the crime series (HBT) occurrences in Zimbabwe. We noted from the model that statistical methods can be applied, for example, time series can be used to analyze crime rate data and have meaning from statistics recorded. The ARIMA (2,1,0) model is worth to be used to give very short-term forecasting because all the assumptions of normality, independence, linear relationship between the response and the predictors and homoscedasticity were not violated. Crime analysis information can be used as an aid to crime prevention and programs, like deployment of policing resources. At national level, it is of great advantage as budgets can be altered to cater for the forecasts. Evaluations of different programs can be done because predictable patterns can create chances for designing crime prevention initiatives and
detection. The forecasts display all the actual values of forecasts, lie within the 95% confidence interval hence the model is appropriate.

5.3 Recommendations
Based on the discussion of the results and conclusions drawn from the study, the researcher therefore recommends that more researches must be done on crime forecasting using time series analysis as there is little research done on the topic. Secondly, through crime prevention initiatives it is recommended that the size of the Community Police force should be increased in order to help meet the police–citizen ratio. Again, the researcher proposes that the computer statistical package MINITAB be installed at every police station in Zimbabwe, so that it can be used to analyze data using the time series technique of ARIMA models. Furthermore, crime series data collected must have an accurate weekly feedback on crime series programs they will be running as the time will be short rather than monthly reports. Crime and intervention analysis is however, some other area of interesting to study further.
REFERENCES


Appendix B: ACF and PACF plots of tested different ARIMA models

4.6 Model Testing
The following models were also checked in order to find the best or tentative model that suits our data. (principle of parsimony)

ARIMA (4,1,1).

Figure 4.15 ACF Plot ARIMA(4,1,1)

Figure 4.16 PACF plot ARIMA (4,1,1)
ARIMA (3,1,2)

Figure 4.17 ACF plot of ARIMA (3,1,2)

Figure 4.18 PACF Plot of ARIMA (3,1,2)

ARIMA(3,1,1)
Figure 4.19 ACF Plot of ARIMA (3,1,1)

Figure 4.20 PACF plot ARIMA (3,1,1)

ACF AND PACF OF ARIMA (3,1,0)
Figure 4.21 PACF plot ARIMA (3,1,0)

Exhibit 4.3: Phillips-perron test

Null Hypothesis: SERIES01 has a unit root
Exogenous: Constant, linear trend.
Bandwidth: 5 (Newey-West using Bartlett kernel)

<table>
<thead>
<tr>
<th>Adj. t-Stat</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phillips-Perron test statistic</td>
<td>-3.193205</td>
</tr>
<tr>
<td>Test critical values: 1% level</td>
<td>-3.517847</td>
</tr>
<tr>
<td>5% level</td>
<td>-2.899619</td>
</tr>
<tr>
<td>10% level</td>
<td>-2.587134</td>
</tr>
</tbody>
</table>

Appendix A: Minitab Exhibit and Raw data

Retrieving project from file: 'C:\Program Files\MINITAB 14\Data\original data\minitab 2.MPJ'

Exhibit 4

Moving Average for C1

Data C1
Length 78
NMissing 0

Moving Average
Length 1

Accuracy Measures
MAPE 11
MAD 224
<table>
<thead>
<tr>
<th>Time</th>
<th>C1</th>
<th>MA</th>
<th>Predict</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>381</td>
<td>381</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>273</td>
<td>273</td>
<td>381</td>
<td>-108</td>
</tr>
<tr>
<td>3</td>
<td>610</td>
<td>610</td>
<td>273</td>
<td>337</td>
</tr>
<tr>
<td>4</td>
<td>497</td>
<td>497</td>
<td>610</td>
<td>-113</td>
</tr>
<tr>
<td>5</td>
<td>703</td>
<td>703</td>
<td>497</td>
<td>206</td>
</tr>
<tr>
<td>6</td>
<td>875</td>
<td>875</td>
<td>703</td>
<td>172</td>
</tr>
<tr>
<td>7</td>
<td>1095</td>
<td>1095</td>
<td>875</td>
<td>220</td>
</tr>
<tr>
<td>8</td>
<td>1340</td>
<td>1340</td>
<td>1095</td>
<td>245</td>
</tr>
<tr>
<td>9</td>
<td>1443</td>
<td>1443</td>
<td>1340</td>
<td>103</td>
</tr>
<tr>
<td>10</td>
<td>1924</td>
<td>1924</td>
<td>1443</td>
<td>481</td>
</tr>
<tr>
<td>11</td>
<td>2500</td>
<td>2500</td>
<td>1924</td>
<td>576</td>
</tr>
<tr>
<td>12</td>
<td>2508</td>
<td>2508</td>
<td>2500</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>3249</td>
<td>3249</td>
<td>2508</td>
<td>741</td>
</tr>
<tr>
<td>14</td>
<td>3694</td>
<td>3694</td>
<td>3249</td>
<td>445</td>
</tr>
<tr>
<td>15</td>
<td>3402</td>
<td>3402</td>
<td>3694</td>
<td>-292</td>
</tr>
<tr>
<td>16</td>
<td>3322</td>
<td>3322</td>
<td>3402</td>
<td>-80</td>
</tr>
<tr>
<td>17</td>
<td>2873</td>
<td>2873</td>
<td>3322</td>
<td>-449</td>
</tr>
<tr>
<td>18</td>
<td>2718</td>
<td>2718</td>
<td>2873</td>
<td>-155</td>
</tr>
<tr>
<td>19</td>
<td>2593</td>
<td>2593</td>
<td>2718</td>
<td>-125</td>
</tr>
<tr>
<td>20</td>
<td>2538</td>
<td>2538</td>
<td>2593</td>
<td>-55</td>
</tr>
<tr>
<td>21</td>
<td>2613</td>
<td>2613</td>
<td>2538</td>
<td>75</td>
</tr>
<tr>
<td>22</td>
<td>3057</td>
<td>3057</td>
<td>2613</td>
<td>444</td>
</tr>
<tr>
<td>23</td>
<td>2709</td>
<td>2709</td>
<td>3057</td>
<td>-348</td>
</tr>
<tr>
<td>24</td>
<td>3029</td>
<td>3029</td>
<td>2709</td>
<td>320</td>
</tr>
<tr>
<td>25</td>
<td>3127</td>
<td>3127</td>
<td>3029</td>
<td>98</td>
</tr>
<tr>
<td>26</td>
<td>1530</td>
<td>1530</td>
<td>3127</td>
<td>-1597</td>
</tr>
<tr>
<td>27</td>
<td>2022</td>
<td>2022</td>
<td>1530</td>
<td>492</td>
</tr>
<tr>
<td>28</td>
<td>1648</td>
<td>1648</td>
<td>2022</td>
<td>-374</td>
</tr>
<tr>
<td>29</td>
<td>1866</td>
<td>1866</td>
<td>1648</td>
<td>218</td>
</tr>
<tr>
<td>30</td>
<td>2111</td>
<td>2111</td>
<td>1866</td>
<td>245</td>
</tr>
<tr>
<td>31</td>
<td>1520</td>
<td>1520</td>
<td>2111</td>
<td>-591</td>
</tr>
<tr>
<td>32</td>
<td>2201</td>
<td>2201</td>
<td>1520</td>
<td>681</td>
</tr>
<tr>
<td>33</td>
<td>1913</td>
<td>1913</td>
<td>2201</td>
<td>-288</td>
</tr>
<tr>
<td>34</td>
<td>1976</td>
<td>1976</td>
<td>1913</td>
<td>63</td>
</tr>
<tr>
<td>35</td>
<td>2809</td>
<td>2809</td>
<td>1976</td>
<td>833</td>
</tr>
<tr>
<td>36</td>
<td>2291</td>
<td>2291</td>
<td>2809</td>
<td>-518</td>
</tr>
<tr>
<td>37</td>
<td>2601</td>
<td>2601</td>
<td>2291</td>
<td>310</td>
</tr>
<tr>
<td>38</td>
<td>2665</td>
<td>2665</td>
<td>2601</td>
<td>64</td>
</tr>
<tr>
<td>39</td>
<td>2820</td>
<td>2820</td>
<td>2665</td>
<td>155</td>
</tr>
<tr>
<td>40</td>
<td>2660</td>
<td>2660</td>
<td>2820</td>
<td>-160</td>
</tr>
<tr>
<td>41</td>
<td>2885</td>
<td>2885</td>
<td>2660</td>
<td>225</td>
</tr>
<tr>
<td>42</td>
<td>2653</td>
<td>2653</td>
<td>2885</td>
<td>-232</td>
</tr>
<tr>
<td>43</td>
<td>2601</td>
<td>2601</td>
<td>2653</td>
<td>-52</td>
</tr>
<tr>
<td>44</td>
<td>2730</td>
<td>2730</td>
<td>2601</td>
<td>129</td>
</tr>
<tr>
<td>45</td>
<td>2615</td>
<td>2615</td>
<td>2730</td>
<td>-115</td>
</tr>
<tr>
<td>46</td>
<td>2661</td>
<td>2661</td>
<td>2615</td>
<td>46</td>
</tr>
<tr>
<td>47</td>
<td>2812</td>
<td>2812</td>
<td>2661</td>
<td>151</td>
</tr>
<tr>
<td>48</td>
<td>2852</td>
<td>2852</td>
<td>2812</td>
<td>40</td>
</tr>
<tr>
<td>49</td>
<td>2985</td>
<td>2985</td>
<td>2852</td>
<td>133</td>
</tr>
<tr>
<td>50</td>
<td>2952</td>
<td>2952</td>
<td>2985</td>
<td>-33</td>
</tr>
<tr>
<td>51</td>
<td>2898</td>
<td>2898</td>
<td>2952</td>
<td>-54</td>
</tr>
<tr>
<td>52</td>
<td>2603</td>
<td>2603</td>
<td>2898</td>
<td>-295</td>
</tr>
<tr>
<td>53</td>
<td>2590</td>
<td>2590</td>
<td>2603</td>
<td>-13</td>
</tr>
<tr>
<td>54</td>
<td>2355</td>
<td>2355</td>
<td>2590</td>
<td>-235</td>
</tr>
<tr>
<td>55</td>
<td>2377</td>
<td>2377</td>
<td>2355</td>
<td>22</td>
</tr>
<tr>
<td>56</td>
<td>2324</td>
<td>2324</td>
<td>2377</td>
<td>-53</td>
</tr>
<tr>
<td>57</td>
<td>2501</td>
<td>2501</td>
<td>2324</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>58</td>
<td>2573</td>
<td>2573</td>
<td>2501</td>
<td>72</td>
</tr>
<tr>
<td>59</td>
<td>2541</td>
<td>2541</td>
<td>2573</td>
<td>-32</td>
</tr>
<tr>
<td>60</td>
<td>2638</td>
<td>2638</td>
<td>2541</td>
<td>97</td>
</tr>
<tr>
<td>61</td>
<td>2650</td>
<td>2650</td>
<td>2638</td>
<td>12</td>
</tr>
<tr>
<td>62</td>
<td>2740</td>
<td>2740</td>
<td>2650</td>
<td>90</td>
</tr>
<tr>
<td>63</td>
<td>3211</td>
<td>3211</td>
<td>2740</td>
<td>471</td>
</tr>
<tr>
<td>64</td>
<td>2623</td>
<td>2623</td>
<td>3211</td>
<td>-588</td>
</tr>
<tr>
<td>65</td>
<td>2547</td>
<td>2547</td>
<td>2623</td>
<td>-76</td>
</tr>
<tr>
<td>66</td>
<td>2420</td>
<td>2420</td>
<td>2547</td>
<td>-127</td>
</tr>
<tr>
<td>67</td>
<td>2393</td>
<td>2393</td>
<td>2420</td>
<td>-27</td>
</tr>
<tr>
<td>68</td>
<td>2563</td>
<td>2563</td>
<td>2393</td>
<td>170</td>
</tr>
<tr>
<td>69</td>
<td>2529</td>
<td>2529</td>
<td>2563</td>
<td>-34</td>
</tr>
<tr>
<td>70</td>
<td>2534</td>
<td>2534</td>
<td>2529</td>
<td>5</td>
</tr>
<tr>
<td>71</td>
<td>2607</td>
<td>2607</td>
<td>2534</td>
<td>73</td>
</tr>
<tr>
<td>72</td>
<td>2842</td>
<td>2842</td>
<td>2607</td>
<td>235</td>
</tr>
<tr>
<td>73</td>
<td>2930</td>
<td>2930</td>
<td>2842</td>
<td>88</td>
</tr>
<tr>
<td>74</td>
<td>2883</td>
<td>2883</td>
<td>2930</td>
<td>-47</td>
</tr>
<tr>
<td>75</td>
<td>2887</td>
<td>2887</td>
<td>2883</td>
<td>4</td>
</tr>
<tr>
<td>76</td>
<td>2818</td>
<td>2818</td>
<td>2887</td>
<td>-69</td>
</tr>
<tr>
<td>77</td>
<td>2625</td>
<td>2625</td>
<td>2818</td>
<td>-193</td>
</tr>
<tr>
<td>78</td>
<td>2426</td>
<td>2426</td>
<td>2625</td>
<td>-199</td>
</tr>
</tbody>
</table>