A COMPARISON OF DISCRIMINANT FUNCTIONAL ANALYSIS AND LOGISTIC REGRESSION BY CATEGORIZING THE INCARCERATED MENTALLY ILL

by

MONA KING

Dissertation

Submitted to the Graduate School

of Wayne State University

Detroit, Michigan

In partial fulfillment of the requirements

for the degree of

DOCTOR OF EDUCATION

2018

MAJOR: EVALUATION AND RESEARCH

Approved By:

_____

Advisor Date

_____

_____

_____

DEDICATION

To my Lord and Savior Jesus Christ who is always with me. To my son Devin, who was understanding during class and study time. Your encouraging words of ‘you can do it, Mom’ said how much you believed in me, and they will never be forgotten. I love you to the moon and back. To my mother and father for your love and always being in my corner. To my sister Andrea, who not only went through the evaluation and research program with me, but who helped me find a balance in life and for telling me to ‘celebrate the small accomplishments’. Yes, I am doing the happy dance! To all of my siblings, nieces and nephews who always cheered for me. I am thankful for and love all of you.

ACKNOWLEDGMENTS

To Professor Shlomo Sawalowsky, my advisor, for all of the patience, encouragement, and guidance you provided during the pursuit of my doctoral degree. The knowledge and insight you shared was invaluable, especially during the writing of the dissertation. I would also like to acknowledge my other committee members, Dr. Barry Markman and Dr. Monte Piliawsky for your advisement, instruction, and time. To everyone at Wayne State University, who assisted directly and/or indirectly, and helped to facilitate this journey. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

In 2015, there were more than 7.3 billion people in the world, and it was estimated 10.3 million were being held in a penal institution (Walmsley, 2016). This was likely an underestimate, because some areas of the world were not included in the survey.

Although the population of the US is approximately 4.4% (.322/7.3) of the world's population, it represents approximately 21% of the world's incarcerated population (United Nations, Department of Economic and Social Affairs, Population Division, 2015), which is an incarceration rate of approximately 683 per 100,000 people, the highest incarcerated rate in the world. According to Tsai and Scommegna (2012), it has held this position since 2002.

The primary purpose of the US prison system is to house criminals. During their incarceration, the prisoner is provided food and drink, clothing, and a bed. Other amenities, such as, educational programs, and televisions in the cells, vary depending on the facility. A warden and guards manage prisoner activities, and these officials must not violate the rights of its occupants. The Federal Bureau of Prisoners (n.d.) stated, "offenders are to be confined in the controlled environments of prisons and community-based facilities that are safe, humane, cost-efficient, and appropriately secure... (About our Agency, Mission)." In an effort to meet these requirements, the U.S. penal system is expensive and overcrowded (Haney, 2012).

Funding of the penal system is a major concern in the United States. There has been growth in the Federal Budget for the Department of Corrections for more than

thirty years, requiring concomitantly increased funding. In 1976, there was a budgeted amount of $US 2,066 million, whereas $US 32,753 million was budgeted in 2014. This represented a 1,485% increase in funding (Administration Office of Management and Budget, n.d.). Drucker (2011) and Stephan (2004) indicated the U.S. prison system budgetary requirements often exceed the budgets of national healthcare or education.

The rising incarceration rate has contributed to overcrowding of the facilities. According to Carson (2015), "in 2014, 19 jurisdictions were operating their prison facilities at more than 100% maximum capacity" (p. 11), and 18 jurisdictions were 90% full (p. 12). This accounts for approximately 72% of the jurisdictions in America.

The expensiveness and overcrowding of prisons has also been attributed, in part, with the deinstitutionalization of mental hospitals (Allen, 2008; Daniel, 2007). Fuller, Sinclair, Geller, Quanbeck, and Snook (2016) indicated in "1955 there were approximately 337 state psychiatric hospital beds for every 100,000 people; however, in 2016 there was only 11.7" (p. 1). This represented a decrease of approximately 97% in the quantity of beds available. Furthermore, Harcourt (2011) cited Penrose (1939) as "finding an inverse relationship between the number of persons in prison and the number of mental hospital beds" (p. 45). This finding suggested a statistical analysis of the process used to predict if the mentally ill will be housed in prison is warranted because it will identify the predictors and the outcome would show if a disproportionate number of prisoners shows signs of mental illness.

Statistical Prediction

There are many methods of prediction in the statistical repertoire. When the outcome variable is dichotomous, such as pass/fail, succeed/failure, lived/died, the

Discriminant Function Analysis (DFA) and the Logistic Regression (LR) are appropriate (Pohar, Blas & Turk, 2004). The outcome of incarceration may be dichotomous, such as signs of mental illness (yes/no). Although the two procedures are generally related, there is no clear advice in the statistical literature on when to use DFA vs. LR, although LR appears to be preferred due to the claim that its underlying assumptions are easier met (Liong & Foo, 2013).

Discriminant Function Analysis

A discriminant function analysis (DFA) can be used to predict group membership, and accommodate two or more groups/classifications. "It identifies which variables differentiate between the naturally occurring groups and classifies the observations into one of the groups" (Antonogeorgos, Panagiotakos, Priftis, & Tzonou., 2009, p. 1). Because the group is the outcome in this method, it is the dependent variable (DV), which is categorical. The predictors of the groups are the independent variables (IV), which are continuous. This statistically method is similar to the MANOVA except in reverse, for the IV of the MANOVA is the group membership, which is categorical, and the discriminating variables are the outcomes (DV), which is continuous (Hair, Black, Babin, & Anderson, 2010). The general form and type of measurement of scale are as follows.

$$Y_1 = X_1 + X_2 + X_3 ... + X_n$$

The measurement scale for Y_1 (dependent variable) is nonmetric, which is qualitative data. This type of data can be either nominal (the number serves as a label) or ordinal (the number serves as a rank). It does not have a mathematical value; thus, mathematical operations cannot be applied to the type of data. However, the

measurement scale of the X_1, X_2, X_3...X_n (independent variable) is metric. Metric data are quantitative; the categories/intervals are the same size. This type of data does have mathematical value; mathematical operations can be applied. The data can be either interval (zero can be anywhere on the scale) or ratio (zero has only one location) (Hair et al., 2010; Gravetter & Wallnau, 2009).

The DFA is not limited by the quantity of IV; there are two methods used to select which IV to include in the model. The first is the simultaneous method. The second method is the stepwise method, where variables are included in the analysis one at a time based on their weight (Hair et al., 2010).

The DFA determines the weight (discriminant coefficient) of the IV; this weight measures the predictive power of the IV. The product of the weight and the IV creates a function (linear effect of X), which can be written as b_iX_i (Hair et al., 2010). The order of the functions is determined by the size of the weights. The function with the largest weight is first, the second largest weight is second, and so on. Each function is orthogonal. When the constant and all of the functions are summed, a variate called a discriminant function is created. A discriminant function is the summation of all of the linear combinations, which maximizes the distance of the between group means and minimum the distance within the groups. It can be expressed in the following linear equation (Antonogeorgos et al., 2009).

$$Y = b_0 + b_1X_{i1} + b_2X_{i2} + b_3X_{i3} +\ldots b_iX_{ik}$$

The Y is the discriminant function (variate), b_0 represents the constant, b_1, b_2, b_3... b_{ik} represents the weight of the corresponding IV (denoted by the subscripts), and X

represents the IV. The equation can also be expressed in the following standardized form (Antonogeorgos et al., 2009).

$$Z_{jk} = \alpha + W_1X_{1k} + W_2X_{2k} + W_3X_{3k} +... W_nX_{nk} \ ,$$

"where Z_{jk} represents the discriminant Z Score of the discriminant function j for object k, α is the intercept, W_i identifies the discriminant weight for the IV and X_{ik} are independent variable i for the object k" (Hair et al., 2010, p. 239).

In DFA, the group mean is called a centroid. The greater the distance between the centroids, the better the model's fit (Hair et al., 2010). However, in situations where there is overlapping of the groups, a cut score is calculated. This score marks the point where the groups are mathematically separated. All cases falling below this score are classified into one group, and all cases above this score are classified in another. Then, the level of significance is determined. The measure for the simultaneous and stepwise methods is Wilks' Lambda and Mahalanobis D^2 respectively (Hair et al., 2010). The statistical significance of the analysis confirms if the results are due to random chance.

Even though the discriminating power of each discriminating function has been explained, it does not explain its predictive ability. Therefore, a classification matrix is created, and a hit ratio is determined. This ratio identifies the overall percentage and quantity of cases correctly classifies, which simplifies the practical application of the results (Hair et al., 2010).

Logistic Regression

Logistic regression (LR) is a form of regression used to explain or predicts the relationship between a DV and one or more IV when the DV is dichotomous. Similar to

other forms of regression, the IV can be either continuous or categorical (Hair et al., 2010). The general form and type of measurement of scale are as follows.

$$Y_1 = X_1 + X_2 + X_3 ... + X_n$$

The measurement scale for the Y_1 (DV) is nonmetric. It is qualitative data and either nominal (the number serves as a label) or ordinal (the number serves as a rank). However, the measurement scale for the X_1, X_2, X_3...X_n (independent variable) can be metric, nonmetric, or a combination of both. It is quantitative, and the IV can be either interval (zero can be anywhere on the scale), ratio (zero has only one location) or a combination of both (Hair et al., 2010; Gravetter & Wallnau, 2009).

Similar to the DFA, logistic regression also uses the simultaneous or stepwise method to introduce variables into the model. LR also determines the weight of each IV; this weight (logistic coefficient) represents the amount of influence the IV has on the DV. The product of the weight and IV is the linear effect of X; it can be expressed as b_iX_i (Hair et al., 2010). Then the linear combinations of X are summed to obtain the best predictor of the dependent variable (Tabachnick & Fidell, 2007). It can be expressed in the following linear equation.

$$Y = b_0 + b_1X_{i1} + b_2X_{i2} + b_3X_{i3} +... b_iX_{ik}$$

The Y is the variate, b_0 is the constant (intercept), b_1, b_2, b_3...bi represents the weight of the corresponding variable (which is denoted by the number 1, 2, 3 ...i), and X represents the independent variable.

In the linear equation, the variate can be either positive or negative, and it is infinite. In addition, it is based on the amount of influence and the scale of the IV (Hair et al., 2010). However, the outcome (DV) of the LR can only be one of two values, which

means it is non-linear. Therefore, the equation is changed, to a logit, which transforms a non-linear relationship into a linear one. The transformed equation is expressed as follows.

$$\text{Logit } p = \ln\left(\frac{p}{1-p}\right) = \alpha + b_1X_1 + b_2X_2 + b_3X_3 \ldots b_iX_i \,,$$

where p is the probability of an event occurring, $1 - p$ is the probability of an event not occurring, and b is the log odds ratio associated with the predictors. This equation uses the maximum likelihood estimate. This process selects the IV, which maximize the likelihood of the observed data (Agresti, 2002; Stoltzfus, 2011). When evaluating the results of LR, the Wald statistic is used. This test gives the statistical significance for each coefficient.

Purpose of the Study

An evaluation of how various factors lead to increasing the number of those incarcerated in the U. S. prison system is necessary, particularly for dichotomous variables such as gender (male/female) and served in the military (yes/no). Also, an important step in the statistical analysis process is the selection of the model. However, there must be a clear understanding of the statistical method to ensure the selection will meet the objective of the analysis and accommodate the type of data. This will result in improved accuracy of the output and its interpretation. Therefore, the purpose of this evaluation is to compare the operating characteristics of DFA and LR when using dichotomous predictors, with a particular application of prison data.

Research Questions

"DFA and LR can be used to assess the same research questions (Antonogeorgos et al., 2009)." This will permit a practical comparison of the outcome of DFA and LR to identify the predictors of prisoners showing signs of mental illness. Therefore, the research questions are:

- H_1: Do the results of the DFA and/or LR show a disproportionate number of mentally ill people in the prison system?
- H_2: According to the DFA and LR, what are the best predictors of mental illness in the US prison system?
- H_3: What is the classification accuracy for DFA vs. LR?
- H_4: What are the similarities and/or differences between the DFA and LR?

Independent Variables

The independent variables were obtained based on prisoner interviews from October 2003 through May 2004 (United States Department of Justice, 2016). They include gender, military service, and a series of questions pertaining to psycho-social variables, such as losing temper easily, not feeling close to friends, and if a suicide was attempted.

Dependent Variable

The dependent variable is whether a candidate show signs of mental illness, coded as (0,1), where zero = no and 1 = yes. The response was obtained from prisoners interviewed from October 2003 through May 2004 (United States Department of Justice, 2016). The survey contained a section of questions, which focused on the mental health of the participants. However, because the responses were the result of

self-evaluation, the description of the variable is shows signs of mental illness verses diagnosed with a mental illness. Furthermore, not all of the prisoners in the system were evaluated for a mental illness. The response will be either no or yes. Therefore, the DV is dichotomous.

Assumptions

The dataset contains second hand data, which are data gathered by another source. It was gathered by the Bureau of the Census for the Bureau of Justice Statistics, (United States Department of Justice, 2016). Therefore, it is assumed, all data were entered in accurately and completely, and the participants were randomly selected. However, if any data were omitted or altered in any way, it is assumed it was notated in the information area of the data set. In addition, it is assumed the interviewer was not biased in the tone or delivery of the questions in any manner, and the methods in which the questions were asked are identical. It is also assumed no additional explanations, regarding the survey questions, were provided to the participants during the interview, and the participants did not discuss the survey with other members of the sample prior to their interview.

Limitations

A limitation of this study is related to the availability of data for prisoners in the United States prison systems. In addition, because of the characteristics of the prisoners included in the sample, the study results may not be generalizable.

Furthermore, the data obtained were the result of a self-evaluation, at the time of the interview, and based on the prisoner's interpretation of the question. Therefore, there is no way to verify the accuracy of their answers. In addition, the data are a

sample of the inmates incarcerated during October 2003 through May 2004, and there is no way to determine if this affects their response to and/or understanding of the questions.

Importance of the Evaluation

By comparing the Discriminant Function Analysis and the Logistic Regression, it is the expectation of this evaluation to identify the differences and/or similarities of the two statistical models when determining the predictors of the mentally ill in prison. Substantively, the application will indicate, based on the survey sample, if mentally ill are housed in the US prison system

Definition of Terms

Classification Matrix: a chart or table that shows the accuracy of model's classification ability

Deinstitutionalization: the systematic dismantling of an institution

Dichotomous Variable: a variable, which only has two values

Discriminant Function Analysis: a statistical method used to estimate the relationship between a dichotomous nonmetric (categorical) dependent variable and a set of metric independent variables.

Hit Ratio: is the percentage of cases correctly classified

Incarcerated Population: estimated number of inmates under the jurisdiction of state or federal prisons or held in local jails.

Incarceration Rate: estimated number of inmates under the jurisdiction of state or federal prisons or held in local jails per 100,000 U.S. residents of all ages.

Jurisdiction: The legal authority of state or federal correctional officials over a prisoner, regardless of where the prisoner is held.

Logistic Regression: a form of regression which predicts and explain the relationship between a dichotomous nonmetric (categorical) dependent variable and a set of metric and nonmetric independent variables.

Maximum Likelihood Estimate: "a procedure that iteratively improves parameter estimates to minimize a specified fit function" (Hair et al., 2010, p. 614)

Mental Illness: refer to disorders generally characterized by dysregulation of mood, thought, and/or behavior.

Measures of scale: describes the data type as nominal, ordinal, interval, or ratio; each describes the limitation of the data

Metric: describes data with mathematical value and mathematical operations (i.e. subtraction and multiplication) can be applied to it

Nonmetric: describes data with no mathematical value and mathematical operations (i.e. subtraction and multiplication) cannot be applied to it because the numbers serve as labels or rank

Prison: A long-term confinement facility, run by a state or the federal government that typically holds felons and offenders with sentences of more than 1 year; however, sentence length may vary by state. Alaska, Connecticut, Delaware, Hawaii, Rhode Island, and Vermont operate integrated systems, which combine prisons and jails.

Prison Population: estimated number of inmates incarcerated in a long-term confinement facility, run by a state or the federal government, which typically holds

felons and offenders with sentences of more than 1 year, although sentence length may vary by jurisdiction.

Prisoner: An individual confined in a correctional facility under the legal authority (jurisdiction) of state or federal correctional officials.

Chapter 2

Literature Review

Both discriminant functional analysis and logistic regression are used to classify subjects into a category/group based upon several explanatory variables (Liong & Foo, 2013). Although their functional form is the same, the method used to accomplish this objective is different (Pohar et al., 2004; Antonogeorgos et al., 2009).

LR uses probability to predict group membership; this is done by determining the odds of the outcome (Antonogeorgos et al., 2009). "The coefficients are measured as the changes in the ratio of the probabilities, which are the odds" (Hair et al., 2010, p. 329). It highlights the relationship between the DV and IV (Pohar et al., 2004; Antonogeorgos et al., 2009). The objective of LR is to identify the likelihood of a case belonging to a group.

However, DFA determines the unique characteristics of a group, and assigns the cases accordingly (Antonogeorgos et al., 2009). The DFA coefficients show the impact the corresponding IV has on the DV (Hair et al., 2010). It highlights the outcome (Pohar et al., 2004; Antonogeorgos et al., 2009). The objective of DFA is to identify the groups of the cases.

Classification

Because DFA and LR classify cases into a group, the accuracy of the classification is essential. Liong and Foo (2013) provided the following formula to calculate the classification accuracy. (p. 1160).

$$\text{Percentage of correct classification} = \frac{\text{Number of observations being classified correctly in a particular group}}{\text{Total number of observations in a particular group}} \times 100$$

However, both methods provide a classification matrix; this is where the accuracy of the model is assessed. This matrix contains the percentage of cases correctly classified, which is known as the hit ratio (Hair et al, 2010).

Furthermore, the comparison of the classification accuracy of DFA and LR, in literature, is not a new activity, for it has been done on many occasions (Press & Wilson, 1978; Harrell & Lee, 1985; Lei & Koehly, 2003, Fall; Antonogeorgos et al., 2009; Liong & Foo, 2013; Alrasheedi & Alggandu, 2014; Balogun, Akingbade & Oguntunde, 2015). However, the views of this type of comparison are varied. For example, Liong and Foo (2013) indicated classification accuracy is the easiest way to evaluate the results; Harrell and Lee (1985) viewed it as an ineffective method. Pohar, Blas and Turk (2004) agreed with Harrell and Lee (1985) by indicating the errors are generally the same for both methods; Fan and Wang (1999) concluded either model could be used because of their comparable classification accuracy.

Estimation Techniques

Both DFA and LR use estimation techniques to obtain a best fit, but the processes are different. DFA can use the ordinary least squares solution (Lei & Koehly, 2003) However LR uses an iterative process; it is called maximum likelihood estimation (MLE). Ordinary least squares (OLS) predicts the line closes to the data, where as MLE estimates the likelihood of an estimated line to give the data.

The goal of OLS is to minimize the vertical distance between the observed and the predicted data points. The process involves measuring the distances for each data point to the estimated line. This will provide the residual, which is the difference between the observed and predicted data point. They can be either positive or negative (Gravetter & Wallnau, 2009). The residuals are squared, and then summed. This total equals the residual sum of the squares; the smallest value is the best-fit line.

The goal of MLE is to choose the best estimator for the data. It maximizes the likelihood of the event (Coughlin, 2016; Hair et al, 2010). MLE chooses the coefficients, and evaluates the selection. The process is repeated until there is very little change. This process results in selecting the coefficients that maximizes the probability of finding the sample data (Tabachnick & Fidell, 2007).

Sample Size

Generally, sample size is a consideration when conducting an analysis for two reasons. The first is economic limitation of the study. Larger samples are generally more expense; budgets may dictate the sample size (Lenth, 2001). Therefore, although the original sample size is large, reducing it may be an economical requirement. The second reason is it affects the test's ability to detect an effect. The larger the sample size, the more powerful the analysis. The more powerful the analysis the greater its ability to detect an effect; it decreases the risk of a Type II error (Biau, Kernais & Porcher, 2008).

However, if the sample is too large, the model may be too powerful. This may result in detecting an effect, which has no practical significance and cause a Type 1 error (Tabachnick & Fidell, 2007; Hair et al., 2010). There has to be a balance between

power and the ability to detect an effect; the impact of sample size can cause either insensitivity or overly sensitivity (Hair et al., 2010). According to, Sawilowsky (1990), "the smaller the sample necessary to detect a treatment, the more efficient, or powerful, is the statistic (p. 93)."

Most statistical methods have a minimum sample size requirement. DFA have a minimum sample size requirement for each group and an overall requirement. It requires at least 20 observations for each IV, and 20 for each group (Antonogeorgos et al., 2009). Conversely, a much larger sample size is recommended for LR. The suggested size is 400 (Hosmer & Lemeshow, 2000). The larger sample size is due to LR using the maximum likelihood estimate. However, research on smaller sample size has not been clear (Lei & Koehly, 2003).

Recognizing the disparity of DFA and LR sample size requirements and understanding the effects the size will have on this study is necessary. However, having an adequate sample size, which correctly represent the population is vital. Therefore, an internet based sample size calculator will be used to determine the size needed to complete this analysis. This calculator is a tool provided by Raosoft, Inc. (http://www.raosoft.com/samplesize.html). Because the sample size identified is the result of the parameter for this study, it will be specific to this analysis, and representative of the population, thereby ensuring size adequacy for this study is achieved.

Assumptions of the Competitors

Both methods are parametric, which means they make assumptions about the distribution of data (Atman & Bland, 2009). These assumptions are normal distribution,

homogeneity of variance/covariance, linearity, and absence of correlated errors. If they are not satisfied, the test results or interpretation may be inaccurate (Erceg-Hurn & Mirosevich, 2008). Therefore, assumptions are frequently the determining factor in selecting a method for analysis.

The assumptions of DFA are often viewed as restrictive (Liong & Foo, 2013; Lei & Koehly, 2003); they are generally violated (Pohar et al., 2004; Press & Wilson, 1978). Furthermore, when the DFA assumptions are violated, LR is often recommended because its assumptions are not as restrictive (Liong & Foo, 2013; Press & Wilson 1978). However, Lei & Koehly (2003) indicated LR does not necessarily support this recommendation. This is due to the large sample size requirement of LR. Furthermore, Pohar et. al. (2004) noted small samples affected LR negatively because of the variation in scores.

The objective of DFA and LR is to classify items into groups; the accuracy of the classification is important. Therefore, if violating the assumptions of DFA reduces the methods classification ability, then LR should be used. However, according to Sever, Lajovic, and Rajar (2005), classification is impacted by sample size. Therefore, increasing the sample size will assist in minimizing the effect of the violation of assumptions. Ghasemi & Zahediasl (2012) concurs referencing the central limit theorem, which proposes as the sample size increase non-normal distribution approaches normality (Tabachnick & Fidell, 2007). Furthermore, as the sample approaches normality, the sampling variance will be reduced because the sample will become more representative of the population (Fiske, Bruna, & Bolker, 2008). Pohar et al. (2004) also indicated DFA and LR yielded similar results with larger samples size.

Therefore, DFA may prove to be robust to violations of its assumptions. Nevertheless, a review of the assumptions is warranted.

Normal Distribution

The first assumption, normal distribution refers to how the z-scores are distributed around the mean. Normally distributed data forms a bell shaped curve, which is symmetrical, and most of the scores are in the middle, which is the mean. The range of its values is from positive infinity to negative infinity; the curve never touches the horizontal axis (Bettany-Saltikov & Whittaker, 2013).

When the data are normally distributed the mean, median, and mode are equal (Gravetter & Wallnau, 2009). However, when the data are not normally distributed, it is usually a result of kurtosis and skewness (Tabachnick & Fidell, 2007). Kurtosis refers to the peak of the curve or lack thereof, whereas skewness describes if the curve is symmetrical on both sides of the mean (Gravetter & Wallnau, 2009).

Normality is an underlying assumption of DFA, but it has been suggested that DFA is not as sensitive to violation (Haggstrom, 1983). However, normality is not an underlying assumption of LR. In addition, LR is considered robust to non-normality, and is viewed as an alternative to DFA (Liong & Foo, 2013). LR violates normality because its curve is an S-curve. This type of curve is called a sigmoid curve. The curve has this shape because the DV can only have two outcomes. It represents a Bernoulli trail, where p = probability of success and q or (1-p) = probability of failure, which follows the binomial distribution (Hair et al., 2010).

Homoscedasticity/Homogeneity of Variance/Covariance

Homoscedasticity occurs when the variance of the scores are constant over a range of predictor variables; heteroscedasticity occurs when the variance is not the same (Hair et al, 2010; Nimon, 2012). In addition, the term homoscedasticity is used when the data are ungrouped, whereas homogeneity is used when the data are grouped (Balogun et al., 2015; Nimon, 2012). It is related to normality because when the assumption of normality is met the variance of the scores is constant (Tabachnick & Fidell, 2007).

Homoscedasticity is an underlining assumption of DFA; it is not a requirement of LR. If this assumption is violated, it may be an indication of outliers. DFA is sensitivity them (Antonogeorgos et al., 2009). Hair et al. (2010) indicated violation of this assumption could be associated with sample size. Regardless of the reason, violation of this assumption could result in an unreliable significance test (Antonogeorgos et al. 2009). In addition, violation could also have a negative effect on the classification accuracy (Liong & Foo, 2013).

Linearity

Linearity is the relationship of two variables when it forms a straight line (Tabachnick & Fidell, 2007). Both DFA and LR create a linear classification model, but the methods are different (Pohar at al., 2004; Liong & Foo, 2013). However, linearity is an assumption for DFA, but not for LR.

DFA forms a linear relationship between the variables in its equation; it uses the general linear equation. This is not the case with LR. Although LR begins with the general linear equation, the dependent variable is dichotomous. Thus, the relationship

of Logistic regression's DV and IV is not linear (Tabachnick & Fidel, 2007; Hair et al, 2010). Therefore, the equation is transformed to a logit, which creates a linear relationship between the logit and the IV (Hair et al, 2010).

Multicollinearity

Both DFA and LR are sensitive to multicollinearity (Tabachnick & Fidell, 2007; and Hair et al., 2010). It is considered problematic, for it causes difficulty in determining the effect of a variable independent of another (Tu, Kellett, Clerehugh, & Gilthorpe, 2005). Because coefficients explain the amount of unique variance of the predictor, multicollinearity can affect the estimation of the coefficient and its statistical significance. This can cause the analysis to be unreliable. Because the variables are measuring the same variation, the coefficients are reduced and the standard error is inflated (Tu, Kellett, Clerehugh, & Gilthorpe, 2005; Tabachnick & Fidell, 2007).

Although multicollinearity causes difficulty in evaluating the effect of an IV, the predicted power of the model is not reduced. This is because predictive power is related to the combination of all of the IV, and not just one variable (Midi, Sarkar & Rana, 2010). Therefore, collinearity is not an issue if the purpose of the analysis is prediction (Lia & Valliant, 2012)

Prison and Mental Institutions Prior to the 1900's

The American colonist did not have prisons. The communities were small, and each developed their own form of criminal law (Meskell, 1999). If someone was accused of a crime and found guilty, they were either publically punished or executed (Kirchhoff, 2010; Meskell, 1999; DePuy, 1951). However, in 1773 the Walnut Street Jail located in Pennsylvania was established. This facility was a workhouse. During the prisoner's stay,

they would work and contribute to the earning of funds needed to run the facility (Kirchhoff, 2010). However, in 1790, the Pennsylvania government ordered the placement of additional security measures, within the facility, for those convicted of heinous crimes, thereby converting the facility into the first United States penitentiary (Hirsch, 1982). This facility became the blueprint for today's prison.

Furthermore, the American Colonist did not have mental hospitals. The responsibility of the mentally ill belonged to the individual's family, and because settlements were small, families relied on each other (Osborn 2009). It was not until 1756 when an area in the Pennsylvania Hospital was created for the mentally ill; it was funded by the wealthy and was the result of lobbying efforts of Benjamin Franklin (Penn Medicine, In the Beginning, 2017). However, in 1773, the first public mental hospital named the Eastern Lunatic Asylum of Virginia was established in Williamsburg, Virginia (Harcourt, 2011). Nevertheless, neither one of these facilities actually provided much in the way of treatment. The patients were locked in basements, and/or restrained or chained (Osborn, 2009).

By the early 1800's, there was an increase in building prison facilities because many states began limiting its use of capital punishment and replacing it with incarceration (Kirchhoff, 2010; Pillsbury, 1989). Conversely, it was not until the mid-1800's when the building of mental hospitals throughout the United States began. In 1841, the Pennsylvania Hospital for the Insane began accepting patients (Osborn, 2009). However, as the population of the United States grew, so did its quantity of mentally ill. In addition, the reliance upon the family and community to care for them

was not always possible because the migration to larger cities to find employment frequently divided families (Osborn, 2009).

Similar to criminals, the behavior of the mentally ill was viewed as deviant. Medical diagnosis and/or treatment were performed in the private facilities for the elite not the poor, and psychiatry was in its infancy stage (Osborn, 2009). Therefore, the mentally ill were frequently not serviced and incarcerated (Malsin, 2015). During this time period, the prisons and mental hospitals served the same purpose. According to the Colony of Williamsbury Foundation (2017) mental hospitals were facilities "designed for security and isolation of its occupants" (History, Public Hospital).

It was during this time, Dorothea Dix began lobbying for the mentally ill and secured funding for the expansion of the Worcester Hospital from Massachusetts legislature (Osborn, 2009). However, in 1854, Dix went before Congress to secure support at the federal level. Although the bill passed the House and the Senate, President Franklin Pierce vetoed it because the responsibility was considered to belong to the states (Currie, 2003). Nevertheless, by the end of the 1800s, approximately 140 mental hospitals were built and 70 renovated (Osborn, 2009). By June 1890, there were more than 74,000 persons listed in mental hospitals. This population contained almost 30,000 people more than the incarcerated population. The prison census for 1890 only showed 45,233 prisoners (Cahalan & Parsons, 1986).

Prison and Mental Institutions after the 1900's

During the early to mid 1900's, the institutionalization of the mentally ill continued. By 1950, the population of mental hospitals was at an all time high. According to the United States Bureau of the Census (1985, p. 79), there were approximately 609,950

patients in hospitals for mental disease in 1953, which is approximately 391 people per 100,000. Conversely, the Bureau of Justice Statistics (BJS) (1982) reported approximately 173,579 prisoners in 1953, which represents approximately 108 people per 100,000. In addition, BJS (1982) further reported from 1931 to 1953 the actual quantity incarcerated only increased by approximately 36,000, whereas the mental hospitals showed a substantial increase of 254,693 patients which represents a percentage increase of 71.7% (United States Bureau of the Census, 1985, p. 79).

However, in 1954 officials in mental hospitals began using antipsychotic medication. Gronfein (1985) indicated the medication caused the patients to become submissive and docile, and it was easily administered. This addition to treatment marked the beginning of a new era in mental health. Medication facilitated treating many patients on an outpatient basis because it helped reduce the symptoms of mental illness (Markowitz, 2006). This meant the quantity of patients confined to hospitals would be reduced. Brill and Patton (1959) indicated where there had been a steady increase in the mentally ill patient population by approximately 2,000 patients per year, in 1955, it decreased by 500. Furthermore, they noted a "34% increase in patient release at New York State mental hospitals (Brill & Robert, 1959, p. 496)."

After the introduction of medication into the treatment plan of the mentally ill, a major change in the mental health system took place with the signing of the Community Mental Health Act, Public Law 88-164, 77 STAT 282, also known as "Mental Retardation Facilities and Community Health Centers Construction Act of 1963." It changed the focus of how mental illness should be handled, by taking it from the 'large institutions', and placing it into smaller 'community centers', which gave a more

humanistic perspective (http://uscode.house.gov/statutes/pl/88/164.pdf). The Act states its purpose is to "provide assistance in combating mental retardation through grants for construction of research centers and grants for facilities for the mentally retarded and assistance in improving mental health through grants for construction of community mental health centers" (http://uscode.house.gov/statutes/pl/88/164.pdf).

In addition, the act was to move funds from the large institutions (hospitals) to community centers. Although the Community Mental Health Act of 1963 detailed funding methods and dollar amounts, the funding it received was not enough (http://uscode.house.gov/statutes/pl/88/164.pdf). Fisher, Geller & Pandiani (2009), stated the population of the mental hospitals dropped by 95%. Although medication played a role in the decrease of the population hospitalized, the majority of this reduction was due to a "70% reduction in the number of psychiatric beds," which happened between 1972 and 1990" (Fisher, Geller & Pandiani, 2009). This reflected the first major step in the deinstitutionalization of the mentally ill. Deinstitutionalization for the purposes of this study is the systematic closing of mental hospitals, and the releasing of mental patients.

Conversely, the prison population began to rise significantly. In addition, for 1972, the BJS (1982) reported 196,092 incarcerated persons. However, by 1990, the number of incarcerated person was more than 771,000. This represented a percentage increase of 293% from 1972 to 1990 (Bureau of Justice Statistics, 1988, 1991). According to Blevins & Soderstrom (2015), this was a result of deinstitutionalization. Prins (2011, p 716) stated this "relationship between deinstitutionalization and the increased rates of mentally ill in prison is transinstitutionalization, (p. 1)" where

transinstitutionalization referred to the transferring of people from one institution (mental hospitals) to another (prison) (Primeau, Bowers, Harrison, & Xu, 2013).

Chapter 3

Methodology

The purpose of this study is to compare the operating characteristics of discriminant functional analysis (DFA) and logistic regression (LR) when using dichotomous predictors, with a particular application of prison data. DFA and LR classify records into categories and can answer the same research questions (Antonogeorgos et al, 2009). In addition, there have been many comparisons made between the methods (Pohar et al., 2004; Antonogeorgos et al., 2009; Liong & Foo, 2013; Taylor & Francis, 2003).

There is a concern about the incarceration of the mentally ill in the United States, since the closing of mental hospitals (Prins, 2011; Blevins & Soderstrom, 2015; Raphael & Stoll, 2013). These decisions are frequently made based on dichotomously scored criteria. Therefore, DFA will be compared with LR to determine if the two methods are comparable in confirming if there are disproportionate numbers of inmates who show signs of mental illness.

Research Design

The data used in this study are archival, gathered by the Bureau of the Census for the Bureau of Justice Statistics (BJS). The dataset is known as the Surveys of Inmates in State and Federal Correctional Facilities, 2004 (United States Department of Justice, 2016). The data were gathered using a survey, which were the results of a computer assisted personal interview. This method consisted of a computer selecting the questions, and a person conducting the interview (United States Department of Justice, 2016, pg. 8). Each interview took approximately one hour, and the participants

were randomly sampled using a two-step process. The first step involved the random sampling of the prison, and the second step involved the random sampling of the inmates/participants (United States Department of Justice, 2016, pg. 3). All of the participants were advised the survey was voluntary and their identities would be confidential.

Population

The dataset of the Surveys of Inmates in State and Federal Correctional Facilities, 2004 (United States Department of Justice, 2016) contained two samples. One sample was from the federal prisons and the second sample was from the state prisons. For this evaluation, the sample used will be the state prisons, and the sample relating to the federal prison will not be used.

There were 14,499 participants in the sample from the state prisons; it contained male and female participants. They were interviewed October 2003 through May 2004. In addition, these prisoners were from 287 of the 1,801 American State prisons, which were identified in the 2000 Census of State and Federal Correctional Facility (United States Department of Justice, 2016).

Variables

The BJS dataset contains 2,984 variables, with five potential values for each variable. However, for the purposes of this study only twenty variables will be used to formulate the dependent and independent variables. All of the dataset variables selected, with the exception of the variable relating to gender, can potentially have a value of one of five; they are 1 (yes), 2 (no), 7 (don't know), 8 (refused) and 9M (Blank, which is missing data). However, for this study the values will be limited to two potential

values; the variables will be nominal and dichotomous. Furthermore, to facilitate the understanding of the results, the dichotomous scoring will be recoded to no (0) and yes (1).

The dependent variable will be if the person showed signs of mental illness, and will be formulated by combining the values of seven dataset variables. The dataset variables and their labels are identified in Table 1.

Table 1
Survey Variables used for Dependent Variable

Survey Variable	Survey Variable Labels
V2401	S9Q9A_1: EVER DIAGNOSED - A DEPRESSIVE DISORDER
V2402	S9Q9A_2: EVER DIAGNOSED - MANIC-DEPRESSION, BIPOLAR DISORDER, OR MANIA
V2403	S9Q9A_3: EVER DIAGNOSED - SCHIZOPHRENIA OR ANOTHER PSYCHOTIC DISORDER
V2404	S9Q9A_4: EVER DIAGNOSED - POST-TRAUMATIC STRESS DISORDER
V2405	S9Q9A_5: EVER DIAGNOSED - ANOTHER ANXIETY DISORDER, SUCH AS A PANIC DISORDER
V2406	S9Q9A_6: EVER DIAGNOSED - A PERSONALITY DISORDER
V2407	S9Q9A_7: EVER DIAGNOSED - ANY OTHER MENTAL OR EMOTIONAL CONDITION

The questions related to these variables ask the participants, if they "have ever been told by a mental health professional, such as a psychiatrist or psychologist, if" they have any of the various forms of mental illnesses identified in Table 1 (United States Department of Justice, 2016, p.758-760). The responses will be combined. As previously noted, If the response was yes to any of the seven questions, the record will

be assigned a value of one. Conversely, if the response to all of the questions was no, the record will be assigned a value of zero. All other records will be removed from the population, for the remaining values represent responses of either refused to answer, don't know or left blank, which represented missing data.

Thirteen variables will be selected from the dataset to formulate the independent variables (IV). However, there will be twelve IV for this evaluation, for two of the variables will be combined. They are compiled in Table 2, which identifies the IV name, label, and potential values.

The independent variable V005_Gender will be created from the dataset variable V0005, which is labeled SEX RECODE VARIABLE. In the dataset, there is not a survey question associated with this variable, because it identifies the gender of the participant. Furthermore, the values assigned to the records were either one for male or two for female. Therefore, female will be coded zero and male will be code one.

The independent variable V0059_Military will be created from the dataset variable V0059, which is labeled S1Q6_1: U.S. MILITARY SERVICE?. The survey question was “Did you ever serve in the U.S. Armed Forces?”. Records will be assigned a value of zero for no and one for yes. If any of the records had a value other than yes or no, the record will be removed.

The independent variable V2188_addictprog will be formulated from the dataset variable V2188, which is labeled S8Q14A: EVER ATTENDED ANY KIND OF ALCOHOL OR DRUG TREATMENT PROGRAM?. The survey question for this variable was “Have you EVER attended any kind of alcohol or drug treatment program?”. The

record will be assigned a value of zero for no and one for yes. If a record contained any other value other than yes or no, it will be removed.

Table 2
Independent Variables

Variable ID	Label (Description)	Response
V0005_Gender	Sex Recode Variable	female=0; male=1
V0059_Military	S1Q6_1: U.S. Military Service	no=0; yes=1
V2188_Addictprog	S8Q14A: Ever attended any kind of alcohol or drug treatment program?	no=0; yes=1
V2379_Temper	S9Q8A_1: Lose temper more easily	no=0; yes=1
V2383_Isolation	S9Q8A_5: Not feeling close to friends or family	no=0; yes=1
V2388_Sleep	S9Q8B_10: Change in sleep	no=0; yes=1
V2396_Unreal	S9Q8D_18: Things don't seem real, like you're in a dream	no=0; yes=1
V2409_Medication	S9Q10A: For a mental or emotional problem, ever taken a medication for mental conditions?	no=0; yes=1
V2413_Hospital	S9Q11A: For a mental or emotional problem, ever admitted to a mental hospital, stayed overnight?	no=0; yes=1
V2417Counseling	S9Q12A: For a mental or emotional problem, ever receive counseling from trained professional?	no=0; yes=1
V2424_Suicide	S9Q14A: Ever attempted suicide or S9Q14B: considered suicide?	no=0; yes=1
V2517_Rules	S10Q13A: Written up or found Guilty of Breaking any rules	no=0; yes=1

The independent variable V2379_Temper corresponds to the dataset variable V2379, where the label is S9Q8A_1: LOSE TEMPER MORE EASILY. The survey question was “During the last year: Have you lost your temper easily, or had a short

fuse more often than usual?” The record will be assigned a value of zero for no and one for yes. If any records contain a value other than yes or no, the record will be removed.

The independent variable V2383_Isolation will be created from the dataset variable V2383, and its label is S9Q8A_5: NOT FEELING CLOSE TO FRIENDS OR FAMILY. The survey question was “During the last year: Have you had difficulty feeling close to friends or family members?”. The records will be assigned a value of zero for no and one for yes. If any records contain any values other than yes or no, they will be removed.

The independent variable V2388_Sleep will be developed from the dataset variable V2388; its label is S9Q8B_10: CHANGE IN SLEEP. The survey question was “During the last year: Has there been a noticeable increase or decrease in the amount of time you sleep?”. The records will be assigned a value of zero for no and one for yes. For records containing values representing values other than yes or no, the record will be removed.

The independent variable V2396_Unreal is formulated from the dataset variable V2396, and its label is S9Q8D_18: THINGS DON'T SEEM REAL, LIKE YOU'RE IN A DREAM. The survey question was “During the last year: Have you had a feeling things don't seem real, like you're in a dream?”. The records will be assigned a value of zero for no and one for yes. If any records contain any values other than yes or no, the record will be removed.

The independent variable V2409_Medication will be developed from the dataset variable V2409, where the label is S9Q10A: EVER TAKEN A MEDICATION FOR MENTAL CONDITIONS? The Survey question was “Because of an emotional or mental

problem, have you EVER taken a medication prescribed by a psychiatrist or another doctor?”. The records will be assigned a value of zero for no and one for yes. If any records contain any values other than yes or no, the record will be removed.

The independent variable V2413_Hospital relates to the dataset variable V2413, which is labeled S9Q11A: EVER ADMITTED TO A MENTAL HOSPITAL, STAYED OVERNIGHT? The survey question was “Because of an emotional or mental problem, have you EVER been admitted to a mental hospital, unit or treatment program where you stayed overnight?” The records will be assigned a value of zero for no and one for yes. For records containing values other than yes or no, it will be removed.

The independent variable V2417Counseling is formulated from the dataset variable V2417, where the label is S9Q12A: EVER RECEIVED COUNSELING FROM TRAINED PROFESSIONAL. The survey question was “Because of a mental or emotional problem have you EVER received counseling or therapy from a trained professional?” The records will be assigned a value of zero for no and one for yes. For records with values other than yes or no, they will be removed.

The independent variable V2424_Suicide is the result of combining the variable V2423 (labeled S9Q14A: Ever attempted suicide?) and variable V2425 (labeled S9Q14B: Ever considered suicide?). Their survey questions, respectively, were “Have you ever considered suicide?” and “How many times have you attempted suicide?”. Similar to the process for the DV, if the response was yes to one or both of the questions, the record will be assigned a value of one. Conversely, if the response to both of the questions was no, the record will assigned a value of zero. All other records will be removed.

The independent variable V2517_Rules corresponded to the V2517, which was labeled S10Q13A: WRITTEN UP OR FOUND GUILTY OF BREAKING ANY RULES. The survey question was “Since your admission [MOST RECENT ADMISSION DATE], have you been written up or found guilty of breaking any of the prison rules?” The records will be assigned a value of zero for no and one for yes. If any records contain any values other than yes or no, the record will be.

Analysis

The calculations will be completed using Software Package for the Social Sciences (SPSS) version 25. An alpha level of .05 will be used to establish statistical significance. According to the internet calculator provided by Roasoft, Inc. (http://www.raosoft.com/samplesize.html) using a response distribution of 50% the minimum sample size (n) required is 375. Therefore, for this analysis, a random sampling of the dataset will be performed by SPSS. Because SPSS’s “Data | Select Cases | Random sample of cases | Approximately _ # of all cases” command accepts only whole percentages, 3% will be used, which will produce slightly more than the minimum sample size required. Once the selection is made, the identified records will be saved in a separate SPSS .sav data for the calculation of the DFA and LR procedures.

Limitation

Due to the nature of archival data, it is not possible to verify the accuracy of entries in the dataset. In addition, it is not possible to assure the survey questions were delivered in the same manor to each participant and no elaboration and/or explanations was provided.

In addition, as with any archived dataset, it can only be assumed the responses were entered accurately, and not altered without indicating the particulars of the alteration in the details of the survey. Because these responses to the survey are the results of inmates providing information about themselves, it can only be assumed the responses were correct, and there was no contamination effect by participants discussing responses with other inmates.

There was no attempt to impute missing values (e.g., item was skipped, respondent refused to answer) or include the response “don’t known”. An imputation procedure may lead to different results, and the expansion of the response pattern to include “don’t know” would preclude the use of logistic regression.

The participating prisons and their inmates were not obtained via a nationally random sampling scheme, and was restricted to the time period of during October 2003 through May 2004. Therefore, the results of this study may not be generalizable.

Chapter 4

Results

The purpose of this evaluation is to compare the operating characteristics of discriminant functional analysis (DFA) and logistic regression (LR) when using dichotomous predictors, with a particular application of prison data. Both DFA and LR are used to classify subjects into a category/group based upon several explanatory variables (Liong & Foo, 2013). For this evaluation, the dependent variable is 'shows signs of mental illness', and there are twelve independent variables. The results of the DFA and LR will be present in this chapter.

The Software Package for the Social Sciences (SPSS) version 25 selected 3% of the cases in the dataset, which provided a sample size of n = 387. This exceeds the minimum sample size requirement of 375, by twelve cases. Prior to selecting the cases, any responses to the survey questions other than No or Yes were removed. This eliminated the potential for any case being identified as missing or out of range data; thereby ensuring the entire sample is included in both the DFA and LR analysis.

Discriminant Function Analysis

The output of the DFA provided by SPSS included various tables. The first two identified which variables and the quantity of cases included in the analysis; this information was seen in the notes and analysis case processing summary respectively. The details of these tables confirmed all twelve of the independent variables (IV) were selected and 100% of the 387 cases were included in the calculation.

The first analysis was displayed in a descriptive group statistics for the model. The table detail included the mean and standard deviation for each independent

variable in its assigned group and the total. The detail is displayed in Table 3. This showed the assumption of normality was violated, for the means, mode, and median were not equal (Gravetter & Wallnau, 2009). This violation was a result of positively skewed data for more of the data fell to the right of the mean than to the left of it. However, according to Sever et al (2005), DFA 'seems to be rather insensitive/fairly robust to skewness (p. 240).

Furthermore, the descriptive statistic also showed 263 of the 387 cases assigned to the group 'does not show signs of mental illness', and the remaining 124 cases were assigned to the group 'shows signs of mental illness' (see Table 3). This represents approximately 32% [(124÷387)*100] of the sample shows signs of mental illness. It was apparent the groups were unequal.

Table 3
Descriptive Statistics for Dependent Variable Showing Signs of Mental Illness

	No = 263		**Yes = 124**		**Total = 387**	
Variable ID	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
V0005_Gender	0.84	0.367	0.67	0.472	0.79	0.411
V0059_Military	0.08	0.277	0.15	0.354	0.10	0.305
V2188_Addictprog	0.61	0.489	0.66	0.475	0.63	0.485
V2379_Temper	0.28	0.449	0.48	0.501	0.34	0.475
V2383_Isolation	0.27	0.443	0.48	0.501	0.33	0.472
V2388_Sleep	0.37	0.482	0.55	0.500	0.42	0.495
V2396_Unreal	0.25	0.437	0.48	0.502	0.33	0.470
V2409_Medication	0.08	0.266	0.81	0.390	0.31	0.464
V2413_Hospital	0.03	0.182	0.42	0.495	0.16	0.365
V2417Counseling	0.09	0.283	0.69	0.463	0.28	0.450
V2424_Suicide	0.12	0.328	0.56	0.499	0.26	0.440
V2517_Rules	0.49	0.501	0.58	0.495	0.52	0.500

Box’s M test was also included in the analysis output; it was used to confirm homogeneity of covariance, which is an underlining assumption of DFA. If the observed

covariance matrices of the DV were equal across groups, the results would not be statistically significant. Because the alpha level used in this analysis was .05, a p value greater than .05 would confirm non-significance thereby validating the assumption of homogeneity has not been violated.

The results of the Box's M, was statistically significant ($p < .001$). This signified the covariance matrices were significantly different, which indicated the assumption of homogeneity was violated (Field, 2007, pg. 611). However, this test is sensitive to large sample size and unequal groups, and the descriptive statistics displayed in Table 3 showed the groups were unequal. However, Tabachnick and Fidell (2007), indicated the "violation may not invalidate the results, but the finding should be noted (p.85)."

The canonical correlation was.781, and the effect size was .61 (i.e., $.781^2$). It is a measure of the strength/power of the relationship between the groups and the discriminant scores. The common language description, using rules of thumb found throughout the statistics literature, is medium to high power, because .5 is generally taken to mean medium and .8 is high. When expressed in percentage, it was interpreted as 61% of the variation in the groups can be explained by the model, which indicated the model was a good fit.

When entering the variables into the analysis, the stepwise estimation method was not used because the results at various steps in the process were not needed for this evaluation. Thus, the simultaneous estimation method was used. Hence, the output included the results for the Wilks' Lambda. This test did not evaluate the details of the difference but evaluated if a difference between the groups existed. This was done by testing the significance of the discriminant function; it is a goodness of fit statistic.

Wilks' Lambda measured the opposite of the canonical correlation, which indicated it measured what the canonical correlation does not. Therefore, the value would be approximately one minus the squared canonical correlation. It was interpreted as the percentage of variation in the dependent variable not explained by the discriminant scores and the smaller the value the better the fit. The output showed this value was approximately .39 [1-(.781^2)], and it was statistically significant (p = 000). This indicated the model was a good fit for the data.

A standardized canonical coefficients and structure matrix was also included in the output. The results were displayed in Table 4. The standardized canonical coefficient represented the predictive ability of each IV, and this ability was based on the size (weight) of each coefficient, which was similar to the regression coefficient. This value was measured as an absolute value, and the sign indicated the direction of the relationship not the strength. This analysis showed the variable with the greatest predictive ability as V2409_Medication, V2417_Counseling and V2424_Suicide, which were .715, .323 and .164 respectively. The variable with the lowest predictive ability was V2188_Addictprog with a weight of -0.016. These coefficients were used to formulate the discriminant function equation, which is as follows:

DF = .715(V2409_Medication) + .323(V2417_Counseling) + .164(V2424_Suicide) + .112(V2413_Hospital) + .091(V2396_Unreal) + .060(V0059_Military) + .055(V2517_Rules) + .051(V2379_Temper) - .026(V0005_Gender) - .026(V2388_Sleep) + - .023(V2383_Isolation) - .016(V2188_Addictprog).

Table 4
Standardized Canonical Coefficients and Structure Matrix

Variables	Standardized Canonical Coefficients	Structure Matrix
V0005_Gender	-0.026	0.141
V0059_Military	0.060	0.170
V2188_Addictprog	-0.016	0.041
V2379_Temper	0.051	-0.158
V2383_Isolation	0.023	0.065
V2388_Sleep	-0.026	0.076
V2396_Unreal	0.091	0.187
V2409_Medication	0.715	0.888
V2413_Hospital	0.112	0.416
V2417_Counseling	0.323	0.646
V2424_Suicide	0.164	0.453
V2517_Rules	0.055	0.159

The structure matrix (see Table 4) showed the correlation between the variables. This information permitted a comparison of the correlations and indicated how closely a variable was related to the others. The variables V2409_Medication and V2417_Counseling had the highest correlation, which was .888 and .646 respectively, whereas .041 was the lowest correlation, which was for variable V2517_Rules.

The analysis also provided a 'Functions at Group Centroids' table. This table contained the at group means, which was -.857 for No (does not show signs of mental illness) and 1.817 for Yes (shows signs of mental illness). From these values a cut score was derived, and the score was 1.165 = $\frac{(263x\text{-}.857)+(124x1.817)}{387}$. The cut score

was a weighted average, which represented the mathematical point where the two groups were separated.

The output included a table containing tests of equality of group means (see Table 5). It compared the mean values to identified which variables were statistically significant predictors independent of the others, which was an ANOVA. For this analysis, all but three (V0059_Military, V2188_Addictprog, and V2517_Rules) were statistically significant; their p values were greater than .05. This information was interpreted as the predictors (IV) significantly discriminated between the groups.

Table 5
Tests of Equality of Group Means

	Wilks' Lambda	**F**	**df1**	**df2**	**Sig.**
V0005_Gender	0.962	15.113	1	385	0.000
V0059_Military	0.991	3.453	1	385	0.064
V2188_Addictprog	0.997	1.005	1	385	0.317
V2379_Temper	0.962	15.239	1	385	0.000
V2383_Isolation	0.957	17.327	1	385	0.000
V2388_Sleep	0.970	11.900	1	385	0.001
V2396_Unreal	0.948	21.052	1	385	0.000
V2409_Medication	0.447	475.444	1	385	0.000
V2413_Hospital	0.757	123.755	1	385	0.000
V2417Counseling	0.605	251.739	1	385	0.000
V2424_Suicide	0.787	104.458	1	385	0.000
V2517_Rules	0.993	2.522	1	385	0.113

The last table the analysis was the Classification Results, displayed in Table 6. This table summarized the output of the analysis, by identifies the quantity and percentages of cases correctly classified in each one of the groups/categories

The results of this analysis revealed overall 88.9% of the cases were correctly classified. The classification detail (see Table 6) showed 92% were correctly classified for did not show signs of mental illness, and 82.3% of the cases for showed signs of mental illness were correctly classified.

Table 6
DFA Classification Results (n=387)

	Shows Signs of Mental Illness	Predicted Group Membership		Total
		No	Yes	
Count	No	242	21	263
	Yes	22	102	124
%	No	92.0	8.0	100.0
	Yes	17.7	82.3	100.0

Logistic Regression

The output of logistic regression also included various tables. The first three tables displayed the quantity of records used (n=387), identified the dummy variable coding (0=No and 1=Yes) for the dependent variable (DV) and independent variables (IV). It also included a table which showed all twelve of the IV were used in the analysis, and displayed their frequencies. These tables provided confirmation of what data were used in the analysis.

The next portion of the output presented a section referred to as 'Block 0' (baseline). This section represented the analysis without including the IV; it contained only the constant. It served as point of reference for when the IVs were included. A

classification table was included in the section of the output. The results are displayed in Table 7.

Table 7
LR Baseline Classification Table (n=387)

Observed		Predicted: Shows Signs of Mental Illness: No	Predicted: Shows Signs of Mental Illness: Yes	Percentage Correct
Shows Signs of Mental Illness	No	263	0	100.0
	Yes	124	0	0.0
Overall Percentage				68.0

The analysis showed 263 of the 387 cases did not show signs of mental illness. This represented 68% [(263/387)*100%] of the records correctly classified. The 68% represented the best possible outcome for predicting the cases, which does not show signs of mental illness, without including the IV. The 'Y' intercept (-.752) was also provided, with a log odds was .471, and it was not statistically significant for the p value was greater than .05. The -2 log likelihood (-2LL) was also calculated, which was 485.438. The -2LL assesses the overall fit of the model, and the higher the number the more poorly the fit of the model.

The final table included in the first section, contained the IV not included in the equations. This table showed if any of the IV would improve the model. Although three of the twelve (V0059_Military, V2188_Addictprog and V2517_Rules) were not statistically significant, the results of the overall statistics was statistically significant. This was interpreted as the IV would improve the predictive power of the model.

The next section of the output was referred to as 'Block 1' (full model). This portion of the output represented the analysis with all of the IV included in the calculation. The section displayed several tables.

The output included an omnibus test of model coefficients, which showed the chi-square (261.093), degrees of freedom (12), and the significance level (.000). A chi-square of 261.093 represented the decrease in the -2LL (485.438), which was calculated in the baseline section. This information was confirmed in the model summary where the -2LL (224.345 = 485.438 - 261.093) for the full model containing the IV was displayed. This decrease was an indication of the model being a better fit with the IV included in the calculation. In addition these results were statistically significant because the p value was less than .05. Therefore, this was interpreted as the full model predicts the cases showing signs of mental illness more accurately than baseline model (just the constant).

Although LR does not provide R^2, it did provide a pseudo-R^2, which did indicate the strength of the relationship between the outcome (DV) and predictor (IV). The model summary showed two pseudo R^2; these values range from 0.0 to 1.0, where 1.0 was perfect. They were the Cox & Snell R square and Nagelkerke R Square, which values were displayed as .491 and .686 respectively. According to Hair et al, (2010) Cox & Snell R square does not reach 1.0.. Therefore, the Nagelkerke R square was also provided. Nevertheless, these values were interpreted as 49.1% (.491x100%) or 68.6% (.686x100%) of the variation in the outcome was explained by the full model. Therefore, it supported the results of the chi-square by showing the model is a better predictor of showing signs of mental illness.

The results of the Hosmer and Lemeshow test results were not statistically significant, for the p value was greater than .379. However, for this test, a statistically significant p value would have been an indication of a poorly fitted model. Therefore, these results also supported the model was a good fit.

A classification matrix was also included for the full model. The results are displayed in Table 8.

Table 8
LR Full Model Classification Table (n=387)

Observed		Predicted Shows Signs of Mental Illness		Percentage Correct
		No	Yes	
Shows Signs of Mental Illness	No	243	20	92.4
	Yes	26	98	79.0
Overall Percentage				88.1

The full model showed 92.4% of the cases were correctly classified for the outcome does not show signs of mental illness, and 79% correct classification for shows signs of mental illness. Overall, the percentage of cases correctly classified was 88.1% of the entire sample. This represented an increase of 20.1% in classification accuracy when compared to the baseline model of 68%. Thus, confirming the full model was a better model than the baseline model.

Although the classification table confirmed the predictors improved the accuracy of classification, the variables in the equations identified which variables made the greatest impact if any. The details of the variables in the equation are shown in Table 9.

Table 9

Variables in the Equation (p≤.05, N=387)

	B	Wald	df	Sig.	Exp(B)	95% CI for EXP(B) LL	UL
V0005_Gender	-0.291	0.399	1	0.528	0.747	0.303	1.845
V0059_Military	0.661	1.226	1	0.268	1.937	0.601	6.240
V2188_Addictprog	-0.099	0.069	1	0.793	0.905	0.431	1.900
V2379_Temper	0.269	0.486	1	0.486	1.309	0.614	2.794
V2383_Isolation	0.170	0.173	1	0.677	1.185	0.533	2.635
V2388_Sleep	-0.206	0.267	1	0.605	0.814	0.373	1.778
V2396_Unreal	0.625	2.462	1	0.117	1.868	0.856	4.075
V2409_Medication	2.943	58.244	1	0.000	18.980	8.913	40.418
V2413_Hospital	0.723	1.706	1	0.192	2.060	0.696	6.096
V2417Counseling	1.455	13.337	1	0.000	4.286	1.963	9.359
V2424_Suicide	0.816	4.005	1	0.045	2.263	1.017	5.034
V2517_Rules	0.435	1.266	1	0.261	1.544	0.724	3.293
Constant	-3.119	28.877	1	0.000	0.044		

Table 9 shows the B, the Wald Statistic, significance level, Exp(B) and the confidence interval for each variable. The B values (see Table 8), represented the coefficients for the predictors (IV); they are the log odds ratio associated with the predictors. Furthermore, they were used in the transformed equation; in this equation p represented the probability of an event occurring, $1 - p$ represented the probability of an event not occurring.

Logit p = ln $\left(\frac{p}{1-p}\right)$ = 2.943(V2409_Medication) + 1.455(V2417_Counseling) + .816(V2424_Suicide) + .723(V2413_Hospital) + .625(V2396_Unreal) + .661(V0059_Military) + .435(V2517_Rules) + .269(V2379_Temper) - .291(V0005_Gender) - .206(V2388_Sleep) + .170(V2383_Isolation) - .099(V2188_Addictprog.

When evaluating the results of LR, the Wald statistic was used. This test gave the statistical significance for each coefficient, and of the twelve variables, only three were statistically significant. Although the baseline showed all but three of the excluded variables from the equation as significant, they were not the same three which were significant when included in the full model. For the full model the statistically significant variables were V2409_Medication, V2417_Counseling, and V2424_Suicide, for the p values were less than .05. The other eight IV (predictors) did not contribute significantly to the prediction ability of the model, for their p values were greater than .05.

The Exp(B) calculated by the analysis represented the proportionate change in the odds, which was the odds ratio for each variable. The results of the Exp(B) calculation for the three significant variables were all greater than one. Hence, the interpretation of these values was as the prediction variable increase, the odds of the event occurring also increased. Therefore, for the predictors V2409_Medication, V2417_Counseling, and V2424_Suicide, the odds of showing signs of mental illness was 18.98, 4.286, and 2.263 times greater respectively than does not show signs of mental illness.

The 95% confidence interval represented the percentage of confidence the population odds ratio would fall within the lower and upper limit range for the corresponding variable. The predictors which were statistically significant had odds ratios greater than one; for the same predictors, the lower and upper limit did not fall below one. This indicated the direction of the relationship in the sample was the same as its population.

Chapter 5

Discussion and Conclusion

The primary purpose of the United States prison system is to house criminals. However, it has the highest incarceration rate in the world (United Nations, Department of Economic and Social Affairs, Population Division, 2015). An evaluation of how various factors lead to increasing the number of those incarcerated in the U. S. prison system was performed for dichotomous variables such as gender (male/female) and 'showing signs of mental illness' (no/yes).The purpose of this study was to compare the operating characteristics of discriminant function analysis (DFA) and logistic regress (LR) when using dichotomous predictors, with an application of prison data. In doing this comparison the following research questions were answered.

H1: Do the results of the DFA and/or LR show a disproportionate number of mentally ill people in the prison system?

H2: According to the DFA and LR, what are the best predictors of mental illness in the US prison system?

H3: What is the classification accuracy for DFA vs. LR?

H4: What are the similarities and/or differences between the DFA and LR?

Research Question 1: Do the results of the DFA and/or LR show a disproportionate number of mentally ill people in the prison system?

A sample of 387 cases was randomly selected from a dataset containing the results of a survey taken by prisoners confined to 287 of the American State prisons during October 2003 through May 2004. Upon running the analysis, both DFA and LR showed 124 of the 387 cases as showing signs of mental illness. This is approximately

32% (124/387) of the cases in the sample. These results were interpreted as a disproportionate number of prisoners showed signs of mental illness.

This outcome was displayed in the group statistic for DFA, and the classification table revealed overall 88.9% of the cases were correctly classified. The LR analysis also supported these result. It showed in the classification table in baseline model 263 of the 387 cases did not show signs of mental illness, which leaves 124 (387 - 263) cases showed signs of mental illness. In addition, the classification table in the full model section of the analysis for LR further supported this outcome; it calculated an overall percentage of 88.1% of the cases were correctly classified. These results were interpreted as a disproportionate number of prisoners showed signs of mental illness.

Research Question 2: According to the DFA and LR, what are the best predictors of mental illness in the US prison system?

Both DFA and LR computed corresponding coefficients for each variable; the largest coefficients represented the best predictors in the model. DFA identified these coefficients as the standardized canonical coefficients, and LR identified them as the logit coefficient, which was denoted as 'B' in the variables in the equation table.

The standardized canonical coefficients calculated by DFA as the best predictors of mental illness in the US prison system was V2409_Medication, V2417_Counseling, and V2424_Suicide, which were .715, .323 and .164 respectively. In addition, this outcome was supported by the structure matrix, which displayed the correlation between the variables. Its calculation showed the highest correlated variables V2409_Medication (.888), V2417_Counseling (.646) and V2424_Suicide (.453

The LR analysis revealed similar results; the logit coefficients also identified the variables V2409_Medication (2.943), V2417_Counseling (1.455) and V2424_Suicide (.816) as the best parameters to maximize the likelihood of predictors of mental illness in the US prison system. In addition the odds ratio [Exp(B)] of these variables was greater than those of the other variables, and the results were statistically significant ($p \leq .05$).

Research Question 3: What is the classification accuracy for DFA vs. LR?

Both DFA and LR included classification tables, and the overall classification accuracy. For this analysis, the overall percentage for DFA and LR, were within 1% point of each other. Liong and Foo (2013) provided the following formula to calculate the classification accuracy. (p. 1160.

$$\text{Percentage of correct classification} = \frac{\text{Number of observations being classified correctly in a particular group}}{\text{Total number of observations in a particular group}} \times 100$$

DFA overall classification accuracy was approximately 88.9% = [(242 + 102)/387] x 100, where 62.5% (approximate) was the overall percentage for No (does not show signs of mental illness) and 26.4% (approximate) was the overall percentage for Yes (shows signs of mental illness). In addition, the predicted group membership (see Table 6) for No was 92% (approximate) and 82.3% (approximate) for Yes; these values represented the cases correctly classified by outcome.

LR analysis consisted of two classification tables, which were baseline (constant only) and full model (included the IV). The baseline (see Table 7), showed an overall percentage correctly classified as approximately 68% = [(263 + 0)/387] x 100. In

addition the predicted group membership for No was 100% and 0% for Yes; these values represented the cases correctly classified by outcome. The full model (see Table 8) showed an overall percentage of approximately 88.1%= [(243 + 98)/387] x 100, which equated to approximately .8% less than DFA (88.9%).Similar to the other classification tables, the predicted group membership for No and Yes was also included, which were approximately 92.4% and 79% respectively. These values represented the percentage correctly classified by outcome. The difference between the classification by outcome for DFA and the full model of LR was also relatively small, which was approximately .4% for No and 3.2% for Yes.

Research Question 4: What are the similarities and/or differences between the DFA and LR?

DFA determines the unique characteristics of a group and assigns cases accordingly, whereas LR uses probability to predict group membership by determining the odds of the outcome. However, the results of the methods had more similarities than differences.

Assumptions

DFA and LR are parametric tests. These types of tests make assumptions about the distribution of data, which are normal distribution, homogeneity of variance/covariance, linearity, and multicollinearity.

The output of DFA included tables to determine if the assumptions normality and homogeneity of variance/covariance were violated, which were group statistics and Box M respectively. The results indicated these assumptions were violated. However, normality and homogeneity of variance/covariance were not assumptions of LR, and

calculations to determine if these assumptions were violated were not included in the output.

Multicollinearity is an assumption of DFA and LR; both methods included a correlation matrix. However, for this evaluation, this assumption was not a concern because its focus was on the predictive ability, which is not effected by multicollinearity (Midi, Sarkar & Rana, 2010).

Linearity is also an assumption of DFA and LR. DFA identified the discriminant function which is the summation of all of the linear combinations. It maximizes the distance of the between group means and minimum the distance within the groups. The linear equation was

DF = .715(V2409_Medication) + .323(V2417_Counseling) + .164(V2424_Suicide) + .112(V2413_Hospital) + .091(V2396_Unreal) + .060(V0059_Military) + .055(V2517_Rules) + .051(V2379_Temper) - .026(V0005_Gender) - .026(V2388_Sleep) + - .023(V2383_Isolation) - .016(V2188_Addictprog).

However, the outcome (DV) of the LR can only be one of two values, which means it is non-linear. Therefore, the equation was changed, to a logit, which transformed a non-linear relationship into a linear one. The transformed equation was

Logit p = ln $\left(\frac{p}{1-p}\right)$ = 2.943(V2409_Medication) + 1.455(V2417_Counseling) + .816(V2424_Suicide) + .723(V2413_Hospital) + .625(V2396_Unreal) + .661(V0059_Military) + .435(V2517_Rules) + .269(V2379_Temper) - .291(V0005_Gender) - .206(V2388_Sleep) + .170(V2383_Isolation) - .099(V2188_Addictprog.

Predictor Variables

DFA and LR were used to classify subjects into a group/category based upon several explanatory variables (Liong & Foo, 2013). Both methods determined the weights (coefficients) of the explanatory variables, which identified the amount of influence the IV had on the DV. For this analysis, DFA and LR identified the variables V2409_Medication, V2417_Counseling, and V2424_Suicide as the best predictors of showing signs of mental illness.

However, the interpretation of these results was not the same for both methods. DFA coefficients, V2409_Medication (.715), V2417_Counseling (.323), and V2424_Suicide (.164), represented the discriminating (predictive) power of the corresponding variable, where V2409_Medication had the greatest predictive power. However the interpretation of LR coefficients was for the predictors V2409_Medication, V2417_Counseling, and V2424_Suicide, the odds of showing signs of mental illness were 18.98, 4.286 and 22.63 (respectively) times greater than does not show signs of mental illness, where V2409_Medication showed the greatest odds.

When the variables were evaluated DFA and LR used different methods. DFA used Wilks' Lambda and F – distribution, whereas LR used Wald test and Chi

distribution. However, the results of both methods showed statistical significant. DFA calculation showed Wilks' Lambda = .39, Chi squared = 357.060, df = 12, and p <.001; the calculation of LR showed Chi squared = 261.093, df = 12 and p <.001. These results indicated the predictors (IV) did have an impact on the outcome (DV) for the model did discriminate between the groups. Furthermore, the statistical significant results showed the results were unlikely to have happen by chance.

Model Fit

Both DFA and LR included a calculation, which showed the strength/power of the relationship between the outcome and predictors. These calculations range from 0 to 1. DFA used the canonical correlation squared ($.781^2$), which equaled an effect size of .6. In statistical literature, this value represented a medium to high power, where .5 is commonly interpreted as medium and .8 is high. Although LR did not include R^2, it provided pseudo-R^2, which were the Cox & Snell R squared and Nagelkerke R squared. Their values were .491 and .686 respectively.

When the canonical correlation squared ($.781^2$), the Cox & Snell R squared (.491) and the Nagelkerke R squared (.686) were expressed in percentages (multiplied by 100%), these results were interpreted as the percentage of variation explained by the model. Both methods indicated an overall good fit.

Classification

Both DFA and LR included classification tables, which are often referred to as hit ratio. This ratio identified the overall percentage and quantity of cases correctly classified, which simplified the practical application of the results (Hair et al., 2010). However, LR included a table prior to the analysis, which was referred to as the

baseline. This table was used as a point of reference, and it permitted determining how much the model improved with the IV (full model) as opposed to just the constant. Nevertheless, DFA and LR classification tables showed similar results for the overall percentage correctly classified. DFA showed 88.9% in overall percentage correctly classified, and LR showed 88.1%. These values were within 1% of each other.

Discussion

The purpose of this analysis was to compare Discriminant Function Analysis (DFA) and Logistic Regression (LR) operating characteristics when using dichotomous predictors with a particular application of prison date. DFA is often viewed as restrictive due to its assumptions (Liong & Foo, 2013; Lei & Koehly, 2003); when the assumptions are violated, LR is often recommended because its assumptions are not as restrictive (Liong & Foo, 2013; Press & Wilson 1978). Although normality and homogeneity of variance were violated, the results of both methods were virtually identical. This suggests DFA is robust to violation of normality and homogeneity of variance/covariance.

In addition, both methods included a classification table in the output. These tables provided practical application of the results by including the percentage correctly classified by category as well as an overall percentage. The overall percentage correctly classified for DFA was 88.9% and 88.1% for LR. These results were within 1% of each other. Therefore, for this analysis, the performance and results of DFA and LR are comparable, even when the assumptions are violated.

Conclusion

By comparing the Discriminant Function Analysis and the Logistic Regression, this evaluation identified differences and similarities of the two statistical models when determining the predictors of the mentally ill in prison. Substantively, the application indicated, based on the survey sample, a disproportionate number of mentally ill are housed in the US prison system.

Recommendation

This analysis was meant to serve as a benchmark for future research regarding determining the predictors of prisoners showing signs of mental illness. However, a limitation of the analysis was the data consisted of the participants (prisoners) self evaluation. It is recommended that future research contain data which consists of documented facts such as types of medication administered, services performed, and duration of the treatment by the prison. This would provide data which is objective, thereby reducing the reliance upon subjective data.

References

Administration Office of Management and Budget (n.d.). Table 5.2 Budget Authority by Agency: 1976 - 2021. Retrieved from https://www.whitehouse.gov/omb/budget/Historicals

Allen, S. (2008). Mental health treatment and the criminal justice system. *Journal of Health and Biomedical Law*, 4(1). 153-191. Retrieved from http://www.suffolk.edu/documents/Law%20Journal%20of%20H%20and%20B/Allen-153-191.pdf

Alrasheedi, M., & Alghamdi, A. (2014). Comparison of classification methods for predicting the movement direction of Saudi stock exchange index. *Journal of Applied Science*, 14, 1883-1888. doi:10.3923/jas.2014.1883.1888

Altman, D. G. & Bland, J. M. (2009). Statistics notes parametric vs. non-parametric methods for data analysis. *BMJ*, 338. doi:10.1136/bmj.a3167

Antonogeorgos, G., Panagiotakos, D. B., Priftis, K. N., & Tzonou, A. (2009). Clinical study logistic regression and linear discriminant analysis in evaluating factors associated with asthma prevalence among 10 to 12 years old children: Divergence and similarity of the two statistical methods. *International Journal of Pediatrics*, 2009, Article ID 952042. doi:10.1155/2009/952042

Balogun, O. S., Akingbade, T. J., & Oguntunde, P. E., (2015). An assessment of the performance of discriminant analysis and the logistic regression methods in classification of mode of delivery of an expectant mother. *Mathematical Theory and Modeling*, 5(5), 147-154. Retrieved from http://eprints.covenantuniversity.edu.ng/5481/1/iiste%20paper%5B1%5D.pdf

Bettany-Saltikov, J., & Whittaker, V. J. (2013, Sept.). Selecting the most appropriate inference statistical test for your quantitative research study. *Journal of Clinical Nursing*, 23, 11-12. doi:10.1111/jocn.12343

Biau, D. J., Kernais, S., & Porcher, R. (2008). Statistics in brief: The importance of sample size in the planning and interpretation of medical research. *Clinical Orthopedics and Related Research*, 466(9). 2282-2288. dio:10.1007/s11999-008-0346-9

Blevins, K. R. & Soderstrom, I. R. (2015). The mental health crisis grows on: a descriptive analysis of doc systems in America. *Journal of Offender Rehabilitation*, 54(2), 142-160. doi:10.1080/10509674.2015.1009965

Brill, H. & Patton, R. E. (1959). Analysisof Population Reduction in New York State Mental Hospitals During the First Four years of Large-Scale Therapy with Psychotropic Drugs. *American Journal of Psychiatry,* 116(6), 495-509. dio:10.1176/ajp.116.6.495

Carson, E. A. (2015). *Prisoners in 2014*. U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. Retrieved from https://www.bjs.gov/content/pub/pdf/p14.pdf

Cahalan, M. W. (1986). Historical Corrections Statistics in the United States, 1850-1984. United States Department of Justice. Bureau of Justice Statistics, Rockville, MD, NCJ-102529. Retrieved from: https://www.bjs.gov/content/pub/pdf/hcsus5084.pdf

Cohen, R. L. (1991) Bulletin: Prisoners in 1990. United States Department of Justice. Office of Justice Programs, Bureau of Justice Statistics, Washington, D. C. 20531. Retrieved from: https://www.bjs.gov/content/pub/pdf/p90.pdf.

Colony of Williamsbury Foundation (1974) Retrieved from: http://www.history.org/almanack/places/hb/hbhos.cfm

Community Mental Health Act, (1963) Public Law 88-164, 77 STAT 282. Retrieved from; http://uscode.house.gov/statutes/pl/88/164.pdf.

Coughlin, M. (2016, Apr 18). *Maximizing the Probability of Detecting an Electromagnetic Counterpart of Gravitational-wave Events.* Cambridge, MA: Department of Physics, Harvard University. doi:10.1007/s10686-016-9503-4

Currie, D. P. (2003). The Constitution in Congress: The Public Lands, 1829-1861. The University of Chicago Law Review (70(3) 783-820. Retrieved from: http://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=5218&context=uclrev

Daniel, A. E. (2007). Care of the mentally ill in prisons: Challenges and solutions. [Editorial]. *Journal of the American Academy of Psychiatry and the Law,* 35(4), 406-410. Retrieved from http://jaapl.org/content/35/4/406

Drucker, E. (2011). *A Plague of Prisons: The Epidemiology of Mass Incarceration in America.* New York, NY: The New Press.

DePuy, L. B., (1951). The Walnut Street Prison: Pennsylvania's First Penitentiary. Quarterly Journal of the Pennsylvania Historical Association, 18(2). Retrieved from https://Journals. psu.edu/phj/article/download/22087/21856

Erceg-Hurn, D. M. & Mirosevich, V. M. (2008). Modern Robust Statistical Methods an Easy Way to Maximize the Accuracy and Power of Your Research. *American Psychological Association*, 63(7), 591-601. dio:10.1037/0003-66.63.7.591.

Federal Bureau of Prisons (n.d.). Mission. Retrieved from https://www.bop.gov/about/agency/agency_pillars.jsp

Fiske, I. J., Eruno, E. M. & Bolker, B. M. (2008). Effects of Sample Size on Estimates of Population Growth Rates Calculated with Matrix Models. *PLoS One,* 3(8):e3080. doi:10.1371/journal/pone.0003080

Fisher, W. H., Geller, J. L. & Pandiani, J. A. (2009). The changing role of the state psychiatric hospital. *Health Affairs,* 28(3), 676-684. doi:10.1377/hlthaff.28.3.676. Retrieved from: http://content.healthaffairs.org/content/28/3/676.

Fuller, D. A, Sinclair, E., Geller, J., Quanbeck, C., & Snook, J. (2016). Going, Going Gone: Trends & Consequences of Eliminating State Psychiatric Beds, 2016 (updated for Q2 data). Office of Research & Public Affairs, Treatment Advocacy Center. Retrieved from http://www.treatmentadvocacycenter.org/storage/documents/going-going-gone.pdf

Ghasemi, A. & Zahediasl, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology Metabolism,* 10(2), 486-489. doi:10.5812/ijem.3505

Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for the Behavioral Sciences* (8th ed.). Belmont, CA: Wadsworth Cengage Learning.

Gronfein, W. (1985). Psychotropic drugs and the origins of deinstitutionalization. Social Problems, 32(5) 437-454.

Haggstrom, G. W. (1983). Logistic regression and discriminant analysis by ordinary least squares. *Journal of Business & Economic Statistics*, 1(3), 229-238. Retrieved from http://www.jstor.org/stable/1391344

Hair, J. F., Black, W. C., Babin, B. J & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Harrell, F. E. & Lee, K. L. (1985). *A Comparison of the Discrimination of Discriminant Analysis and Logistic Regression under Multivariate Normality*. In P. K. Sen (Ed.), Biostatistics: Statistics in Biomedical, Public Health and Environmental Sciences. North-Holland: Elsevier Science Publishers, 333-343. Retrieved from http://biostat.mc.vanderbilt.edu/wiki/pub/Main/FrankHarrell/discrim.pdf

Haney, C. (2012). Prison effects of in the age of mass incarceration. *The Prison Journal*, 20(10), 1-24. doi:10.1177/0032885512448604

Harcourt, B. E. (2011). An institutionalization effect: The impact of mental hospitalization and imprisonment on homicide in the United States, 1934 - 2001. *Journal of Legal Studies*, 40(1). 39-83. Retrieved from http://www3.law.columbia.edu/bharcourt/documents/harcourt-jls-final-2011.pdf

Harcourt, B. E. (2011). Reducing Mass Incarceration: Lessons from the Deinstitutionalization of Mental Hospitals in the 1960's. *Ohio State Journal of Criminal Law* 9(1), 53-88. Retrieved from: http://moritzlaw.osu.edu/osjcl/Articles/Volume9_1/Harcourt.pdf

Heiman, G. H. (2011). *Basic Statistics for the Behavioral Science*. Belmont, CA: Wadworth, Cengage Learning.

Hirsch, A.J. (1982) From pillory to penitentiary: The rise of criminal incarceration in early massachusetts. Michigan Law Review, 80(6), 1179-1269.

Hosmer, D. W. & Lemeshow, S. L. (2000). *Applied Logistic Regression* (2nd ed.). Hoboken, NJ: Wiley-Interscience.

Kirchhoff, S. M. (2010). Economic Impacts of Prison Growth. Congressional Research Service, 7-5700, R41177. Retrieved from: https://fas.org/sgp/crs/misc/R41177.pdf

Lia, D., & Valliant, R. (2012). Variance inflation factors in the analysis of complex survey data. *Survey Methodology*, 38(1) 53-62. Statistics Canada, Catalogue No. 12-001-X. Retrieved from http://www/statcan/gc.ca/pub/12-001-x/2012001/article/11685-eng.pdf.

Lenth, R. V. (2001). Some practical guidelines for effective sample-size determination. *The American Statistician*, 55(3), 187-193. Retrieved from http://conium.org/~maccoun/PP279_Lenth.pdf

Lei, P. W. & Koehly, L. M. (2003). Linear discriminant analysis versus logistic regression: A comparison of classification errors in the two-group case. *The Journal of Experimental Education*, 72(1), 25-49. doi:10.1080/00220970309600878

Liong, C., & Foo, S. (2013). Comparison of Linear Discriminant Analysis and Logistic Regression for Data Classification. *AIP Conference Proceedings* 1522(1), 1159-1165. doi:10.1063/1.4801262

Markowitz, F. E. (2006). Psychiatric hospital capacity, homelessness, and crime and arrest rates. Criminology 44(1), 45-72.

Meskell, M. W. (1999). An American resolution: The history of prisons in the United States from 1777 to 1877. Stanford Law Review, 51(4) 839-865. Retrieved from: http://www.jstor.org/stable/1229442

Midi, H., Sarkar, S. K., & Rena, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics*, (13(3), 253-267.

Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: a mini-review. Frontiers in Psychology, 3(322), 1-5. doi:10.3389/fpsyg.2012.00322

Osborn, L. A. (2009). From Beauty to Despair: The Rise and Fall of the American State Mental Hospital. *Psychiatry Q*, 80, 2109-231. dio:10.1007/s11126=009-9109-3. Retrieved from: http://www.psychodyssey.net/wp-content/uploads/2011/05/From-Beauty-to-Despair-The-Rise-and-Fall.pdf

Penn Medicine, (2017). In the Beginning: The Story of the Creation of the Nation's First Hospital. Retrieved from: http://www.uphs.upenn.edu/paharc/features/creation.html

Penrose, L. S. (1939). Mental disease and crime: outline of a comparative study of European statistics. *British Journal of Medical Psychology,* 28, 1-15.

Pillsbury, S. H. (1989). Understanding penal reform: The dynamic of change. Journal of Criminal Law and Criminology, 80(3), 726-780.

Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodoloski Zvezki*, 1(1), 143-161. Retrieved from http://mrvar2.fdv.uni-lj.si/pub/mz/mz1.1/pohar.pdf

Press, S. J., & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis, *Journal of the American Statistical Association*, 73(364), 699-705. Retrieved from http://math.arizona.edu/~hzhang/math574m/LogitOrLDA.pdf

Primeau, A., Bowers, T. G., Harrison, M. A. & XuXu (2013) Deinstitutionalization of the mentally ill: Evidence for transinstitutionalization from psychiatric hospitals to penal institutions. Comprehensive Psychology, 2(2), 1-10. doi:10.2466/16.02.13.CP.2.2

Prins, S. J. (2011). Does Transinstitutionalization Explain the Overrepresentation of people with serious mental illness in the criminal justice system? *Community Mental Health Journal*, 47, 716-722 dio:10.1007/s10597-011-9420.y

Rana, R. K., Singhal, R., & Dua, P. (2016). Deciphering the dilemma of parametric and nonparametric tests. *Journal of the Practice of Cardiovascular Sciences*, 2(2), 95-98. doi:10.4103/2395-5414.191521

Raphael, S. & Stoll, M. A. (2013). Assessing the contribution of the deinstitutionalization of the mentally ill to growth in the U.S. incarceration rate. *The Journal of Legal Studies*, 42(1), 187-222. doi: 0047-2530/2013/4201-0007. Retrieved from: http://www.journals.uchicago.edu/doi/pdfplus/10.1086/667773

Rencher, A. (2002). *Methods of Multivariate Analysis* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Sawilowski, S. S. (1990). Nonparametric test of interaction in experimental design. *Review of Educational Research*, 60(1), 91-126. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.890.5238&rep=rep1&type=pdf

Sever, M., Lajovic, J., & Rajer, Borut. (2005). Robustness of the fisher's discriminant function to skew-curved normal distribution. *Metodološki zvezki,* 2(2), 231-242. Retrieved from: http://www.stat-d.si/mz/mz2.1/sever.pdf

Siegel, S. (1957). Nonparametric Statistics. *The American Statistician* (11)3, 13-19. Retrieved from http://www.jstor.org/sici?sici=0003-1305%28195706%2911%3A3%3ANB2-I

Stephan, J. J. (2004). *State Prison Expenditures, 2001.* U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from http://www.bjs.gov/index.cfm?ty=pbdetail&iid=1174

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine,* 18(10), 1099-1104. doi:10.1111/j.1553-2712.2011.01185.x

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Pearson Education, Inc.

Torrey, E. F., Kennard, A. D., Eslinger, D., Lamb, R., & Pavle, J. (2010). More Mentally Ill Persons Are in Jails and Prisons Than Hospitals: A Survey of the States. Arlington, VA: Treatment Advocacy Center and Nationals Sheriffs Association, 2010. Retrieved from: http://www.treatmentadvocacycenter.org/storage/documents/final_jails_v_hospitals_study.pdf

Tsai, T. & Scommegna, P. (2012). *U.S. Has World's Highest Incarceration Rate*. Population Reference Bureau. Retrieved from http://www.prb.org/Publications/Articles/2012/us-incarceration.aspx

Tu, Y-K., Kellett, M., Clerehugh, V., & Gilthorpe, M. S., (2005). Problems of correlations between explanatory variables in multiple regression analysis in the dental literature. *British Dental Journal*, 199(7), 457-461. doi:10.1038/sj.bdj.4812743

United Nations, Department of Economic and Social Affairs, Population Division. (2015). *World Population Prospects: The 2015 Revision, DVD Edition*. Retrieved from https://esa.un.org/unpd/wpp/Download/Standard/Population/

United States Bureau of the Census, Statistical Abstract of the United States (1985), (Seventy-ninth edition). Washington, D. C. Retrieved from: https://www.census.gov/library/publications/1958/compendia/statab/79ed.html

United States Department of Justice. Office of Justice Programs. Bureau of Justice *Statistics. Survey of Inmates in State and Federal Correctional Facilities, 2004. ICPSR04572-v2.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016-04-27. http://doi.org/10.3886/ICPSR04572.v2

Unites State Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. (1990). Bulletin: Prisoners in 1989. Washington D. C., NCJ-122716. Retrieved from: https://www.bjs.gov/content/pub/pdf/p89.pdf

United States Department of Justice. Bureau of Justice Statistics, (1982). *Bulletin: Prisoners 1925-81*. Washington, D. C., NCJ-85861. Retrieved from: https://www.bjs.gov/content/pub/pdf/p2581.pdf

Walmsley, R. (2016). *World Prison Population List (eleventh edition)*. Institute of Criminal Policy Research. Retrieved from http://www.prisonstudies.org/research-publications?shs_term_node_tid_depth=27