Faculty of Science

Statistics Department

Bachelor of Science Honours Degree in Statistics

Research Project

Statistical Analysis on Causes of High Schools Dropouts in Windhoek (2009-2012)

Student: Mr. Torosi Sakeus Karupu

200734903

Supervisor: Mr. Werner Tjipueja

02 November 2012

ACKNOWLEDGEMENTS

First and foremost, I express my gratitude and give praise to the Almighty GOD for his continuous blessings for giving me opportunity to work on this challenging project. "Thank you Lord for the strength, hope and faith you gave me, when I thought I was not going to make it on time".

Moreover, I will like to thank my friends and my fellow classmates, for giving me courage in my work and helping me with some ideas to add to my project.

Furthermore, I would like to thank my supervisor Mr Werner Tjipuaeya for his support and guidance I got from him. I owe thanks to the Namibian Ministry of education and the regional office in Windhoek for providing me with Data. Finally, I would like to thank my lovely Pastor Chris, sisters, my brother and two uncles for their hard work, continuous caring that contributed towards the success of this project even though they are far, it did not hinder their support.

ABSTRACT

This project studies the statistical analysis on causes of High School dropouts in Windhoek for the period time 2009 to 2012. The objective of the study is to find out the current dropout number according to sex, schools, grades and to come up with a regression model that can be used in the description, in the prediction and control of the number of dropouts on the basis of the independent variables by using regression analysis. Such prediction can aid Namibians as a whole to make informed decision to plan for the future.

The increasing number of dropouts in Namibia especially in Windhoek, the unemployment and criminal records caused by dropouts served as a motivation for this study. SPSS software was used to analyse data, where parametric tests like independent t-test was used and by using ANOVA did work for this data and hence gave desirable results. The conditional backward logistic regression analysis was applied to model the data. Finally, a regression model was developed and they fit the data reasonably well.

CHAPTER 1

1.1 Introduction and background of the study

This study focus is on Statistical Analysis on causes of High School dropouts in Windhoek for the period time 2009 to 2012. The first chapter of the study provides more information on the research background and discusses issues such as the statement of the problem, research objectives, research questions, significant of the study and the hypothesis.

Education is the basic requirement for human development and survival of the society. This is the fundamental human right. It is necessary and a universal feature of society by which every generation transmits social heritage to the next generation (Ruchita Sharma, Shubhangna Sharma and Shipra Nagar, 2007). The ministry of education in Namibia and the regional offices in Windhoek is responsible for high schools administration in Windhoek and uses all high schools in Windhoek to look at the performances, absentisms, dropouts and total number of learners at each school in order to give a report to the government. Thus due to this, one of its applications is in the area of Statistical analysis.

Numerous definitions of dropouts appear in the educational literature, in this study a drop-out is defined as a learner who enrolled with school but fail to sit for the final exams. Here drop out is not presented as a distinct event, but rather a process where a range of supply-demand factors interact to influence schooling access.

1.2 Statement of the problem

Education is a key to developing the economic, social, scientific and political institutions of nation states. Dropping out of school is an issue faced by many teens today. This shows that there are many reasons why learners want to and do dropout of schools. The nation's young dropouts experience a wide array of labor market, earnings, social and income problems that exacerbate their ability to transition to careers and stable marriages from their mid-20s onward. It is for this purpose that the country and individuals have to take cognizance of this important

issue. This goes further to challenge the ability of the government in giving or providing the needs for individuals.

This is a problem because a country relies on these aspects to direct planning process in various aspects of life. According to Lockheed and Verspoor (1991), access to education is at the centre of the development agenda of most low-income countries. Namibia in the arenas of education and training is not exception in particular. Dropouts are less likely to be active labor force participants than their better educated peers, and they frequently experience considerably higher unemployment rates when they do seek work. As a consequence, they are much less likely to be employed than their better educated peers across the nation, and gaps typically widen as national labour markets deteriorate such as during the current recession.

1.3 Research objectives

The objective of the study is to find out the current dropout number according to sex, schools, and grade. That is;

1.3.1   To find out if there is a significant difference in the number of dropouts for secondary schools among boys and girls.

1.3.2   To find out if there is a significant difference in the number of dropouts between the Government and private secondary schools.

1.3.3   To find out which grade has the highest dropout's number.

1.3.4   To compare the current number of dropouts with the previous three years.

1.3.5   Finally to come up with a regression model that can be used in the description, in the prediction and control of the number of dropouts on the basis of the independent variables.

1.4 Research question

This study will be guided by the following questions:

1.4.1 How can statistical analysis be applied to predict secondary schools dropouts in Windhoek?

1.4.2 Is there a significant difference in dropout number of secondary school between boys and girls?

1.4.3 Is there a significant difference in the number of dropouts of secondary school children between government and private school?

1.4.4 Is dropout's number increasing every year?

1.4.5. Is there a difference in dropout rate of secondary school among the grades?

1.5 Significance of the study

Several research studies have been conducted on secondary school dropouts and have established invariable change in secondary schools regionally and nationally. Therefore, the study is important because it will bring answers to the question on how statistical methods can be applied to look at the current rate of secondary school dropouts in Windhoek. The study will also help to plan better for the future through the usage of statistical method. In the same vein, the ministry of education and the regional office can adopt statistical models (regression analysis) which have accurate predictive power in predicting secondary schools dropout in the country.

These include Insights into the process as well as the event of drop out, indicators of risk factors that Could be used to predict dropouts, opportunities and mechanisms that facilitate Children to drop back in after they drop out, the impact of over age enrolment and repetition on drop out, and school friendliness factors that may promote higher rather than lower levels of attendance. This overview of drop out is a good starting point for any researcher seeking to unpack the processes that lead to exclusion and identify the most important causes.

1.6 Hypothesis

i. H_0: There is no significant difference in the number of dropouts between boys and girls in secondary schools.

ii. H_o: There is no significant difference in dropout number of secondary school's children at government and private school.

iii. H_o: There is no difference in number of dropouts among the grades.

iv. H_o: There is no difference in the number of dropouts each year.

CHAPTER 2
LITERATURE REVIEW

2.1 Literature Review

High school dropout's prediction is important in our everyday lives for planning of various activities. It is important to know the current trend of dropouts in order to plan our future. However, no high school dropout's prediction if one cannot rely on the information.

Research has been carried out throughout the country to show the school dropouts trend. The startling statistics that one high school student drops out every 9 seconds illustrates the magnitude of the problem (Children's Defense Fund, 2005). Dropout rates are highest among learners with emotional and behavioral disabilities; half of the learners dropped out of school in 1998-1999 (U.S. Department of education, 2001) which the case in Namibia such the 2008-2009. Another research was carried out by the NAMCOL Strategic Development Plan 2009-2011 with different purposes on a similar topic with related findings.

Dropout statistics are particularly alarming because jobs that pay living wages and benefits have virtually disappeared for youth without a high school certificate. For society, the costs of dropout are staggering, estimated in billions of dollars in lost revenue, welfare programs, unemployment programs, underemployment, and crime prevention and prosecution (Christenson, Sinclair, Lehr, & Hurley, 2000). Given these individual and societal consequences, facilitating school completion for all students must be a critical concern for researchers, policymakers, and educators across the country.

A comprehensive review of dropout interventions (Lehr, Hanson, Sinclair, & Christenson, 2003) indicated that the dropout research has been overwhelmingly predictive or descriptive (i.e.., there have been few controlled studies), and the methodology used to evaluate the effectiveness of the majority of dropout interventions has been judged to be low quality or poor scientific merit.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This Chapter will present the methodology used in the study. It will review the study design and sources of data that will be used; also the techniques for data analysis and the expected contribution to knowledge generation will be explained in this chapter.

3.2. Study design and sources of data

No questionnaires were designed for this project. The study relies on the secondary data that was collected from the ministry of education and the regional office for the time period 2012(term1 and term2). This information was collected from available records at the high schools since the ministry of education and the regional office is responsible for high schools administration in Namibia. The study population comprise of different high schools across Windhoek.

3.3. Data Analysis and Methods

The t-test will be used to carry out the hypothesis test. At the same time causal forecasting method (regression analysis) will be used to quantitatively describe, explain and predict future values of dropouts. Data was analysed using computer application called SPSS. SPSS was used to organise plot graphs, carry out hypothesis test and regression analysis.

3.3.1 Regression analysis: A prediction tool

Regression analysis is a statistical methodology that is used to relate a variable of interest called dependent variable or response to one or more predictor or independent variables. This understanding can then be applied to the preparation of high school dropouts of future levels.

3.3.2 Role of Regression Analysis

The aim therefore is to build a linear model relating Y to X, this is called a regression model or predicting equation. As a result, this model can be used either in the description or in the prediction and control of Y on the basis of the independent variables.

3.4. Data Collection

Data was collected from the ministry of education and regional office in Windhoek as secondary data. Secondary data refers to aggregated information compiled by a department over an extended period for its own use (Steppingstones, 2004). In this case this information includes records of secondary schools dropout from 2009-2012. This is readymade information.

CHAPTER FOUR

RESULTS AND ANALYSIS

Only 43 secondary schools in Windhoek were used in the analysis of data. Moreover the variable dropout was selected and analyzed by using different statistical methods; Independent Sample T-test by comparing means of gender, grades, schools (Private/GRN) for the secondary school and regression analysis to build a model.

4.1 Data descriptive

Data validation

Table 1

Statistics

		Period	Gender	School	Reason	Number_of_ Dropouts	Grade	Year
N	Valid	308	308	308	308	308	308	308
	Missing	0	0	0	0	0	0	0

It's clear from Table 1 that there are no errors and missing value, there are 308 entries.

4.2 Data Summary

1. Period

It's clear from table 2 below that during this observed short period time of term1 and term2, the number of dropouts was found to be 150 and 158.

Table 2

Period

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	term 1	150	48.7	48.7	48.7
	term 2	158	51.3	51.3	100.0
	Total	308	100.0	100.0	

2. Gender

According to the department of education of regional office, Table 3 shows that the number of dropout was 151 for male and 157 for female.

Table 3

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	151	49.0	49.0	49.0
	female	157	51.0	51.0	100.0
	Total	308	100.0	100.0	

3. Grade

Table 4 below shows that, 27.9% of the dropouts observed are in grade8 for both government and private schools. In grade 9, 10, 11 and 12, dropouts are 24.7%, 21.4%, 17.2%, and 8.8%, respectively.

Table 4

Grade

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	grd8	86	27.9	27.9	27.9
	grd9	76	24.7	24.7	52.6
	grd10	66	21.4	21.4	74.0
	grd11	53	17.2	17.2	91.2
	grd12	27	8.8	8.8	100.0
	Total	308	100.0	100.0	

4. Type of School

The table 5 below shows that Government schools have a high number of dropouts compared to the private schools. At present, about 76.6% of the dropouts were enrolled in Government schools and 23.4% of the children were enrolled from Private schools.

Table 5

Typeof school

|		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Private	72	23.4	23.4	23.4
	Goverment	236	76.6	76.6	100.0
	Total	308	100.0	100.0	

5. Reasons

The data shows that out of recent 308 dropouts in Windhoek, there are major reasons behind that cause learners to dropout of schools. Table 6 below shows that the major reasons of high school dropout are; moved away (other region), teenage pregnancy, no interest in school and unknown reasons.

Table 6

Reason

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Moved away(other region)	103	33.4	33.4	33.4
	Teenage Pregnancy	46	14.9	14.9	48.4
	Social Problems/poverty	12	3.9	3.9	52.3
	Illness/passed away	6	1.9	1.9	54.2
	Started Work	3	1.0	1.0	55.2
	Criminal activities	3	1.0	1.0	56.2
	No interest in school	32	10.4	10.4	66.6
	Parental request	10	3.2	3.2	69.8
	Reasons unknown	84	27.3	27.3	97.1
	Suspension/disciplinary problems	5	1.6	1.6	98.7
	Other reasons	4	1.3	1.3	100.0
	Total	308	100.0	100.0	

4.3 Data analysis

4.3.1 Testing for normality

Normal Q-Q Plot of Number_of_Dropouts
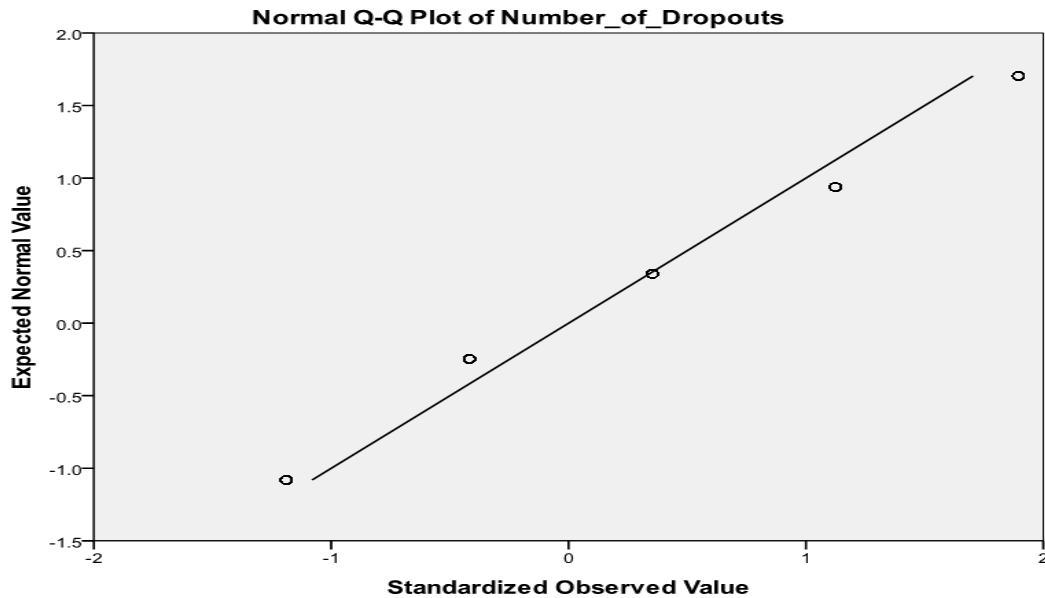
Statistics

Number_of_Dropouts

N	Valid	308
	Missing	0
Std. Deviation		1.297
Sum		783

The table and the Q-Q Plot above shows the standard deviations of the entire variables when tested for normality, this indicates that the points of the variables of interest didn't divert from a line of normality. So we carry out a parametric test.

<u>Hypothesis testing</u>

4.3.2)

H_O: There is no significant difference in the number of dropout between boys and girls in secondary schools.

H_a: There is a significant difference in the number of dropout between boys and girls in secondary schools.

$\alpha = 0.05$

RR: We will reject H_O $(null\ hypothesis)$ if the calculated p - value $< \alpha/2 = 0.025.$

Test Statistics

Table 7

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number_of_Dropouts	Equal variances assumed	1.885	.171	.648	306	.518	.139	.215	-.284	.562
	Equal variances not assumed			.646	295.716	.519	.139	.216	-.285	.564

Since from Table 7, the calculated p-value=0.518 > 0.025 we do not have enough evidence to rejectH_O. Hence there is no difference in the number of dropouts among boys and girls in secondary schools.

4.3.3)

H_O: There is no significant difference in the number of dropouts of secondary school's children between government and private school.

H_a: There is a significant difference in the number of dropouts of secondary school's children between government and private school.

$\alpha = 0.05$

RR: We will reject H_O (*null hypothesis*) if the calculated p - value $< \alpha/_2 = 0.025.$

Table 8

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number_of_Dropouts	Equal variances assumed	3.014	.084	-1.381	306	.168	-.350	.253	-.848	.149
	Equal variances not assumed			-1.604	154.770	.111	-.350	.218	-.781	.081

Table 8 provides a calculated p-value = 0.168 which is greater than 0.025, so we fail to rejectH_O (*null hypothesis*). Hence, there is no difference between the government and private secondary school's dropouts in number.

4.3.5)

H_O: There is no difference in the number of dropouts among the grades.

H_a: At least one grade has a different number of dropouts.

$\alpha = 0.05$

RR: We will reject H_O (*null hypothesis*) if the calculated p - value $< \alpha/_2 = 0.025.$

Table 9

Tests of Between-Subjects Effects

Dependent Variable:Number_of_Dropouts

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	34.027[a]	4	8.507	2.440	.047
Intercept	1177.901	1	1177.901	337.916	.000
Grade	34.027	4	8.507	2.440	.047
Error	1056.193	303	3.486		
Total	2636.000	308			
Corrected Total	1090.221	307			

a. R Squared = .031 (Adjusted R Squared = .018)

Table 9 provides a calculated p-value = 0.047 which is greater than 0.025, so we fail to reject H_0 (null hypothesis). Hence, there is no difference in number of dropouts among the grades.

4.3.4)

H_0: There is no difference in the number of dropouts each year.

H_a: At least one year has a different number of dropouts compares to other years.

$\alpha = 0.05$

RR: We will reject H_0 (null hypothesis) if the calculated p - value $< \alpha/2 = 0.025$.

Table 10

Tests of Between-Subjects Effects

Dependent Variable:Number_of_Dropouts

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	43.162ᵃ	3	14.387	5.017	.002
Intercept	6346.570	1	6346.570	2212.868	.000
Year	43.162	3	14.387	5.017	.002
Error	4259.024	1485	2.868		
Total	10819.210	1489			
Corrected Total	4302.186	1488			

a. R Squared = .010 (Adjusted R Squared = .008)

Table 9 provides a calculated p-value = 0.002 which is less than 0.025, so we do have enough evidence to reject H_0 (null hypothesis). Hence, at least one year has a different number of dropout compares to other years.

Since the null hypothesis is rejected, we carry out a post hock test to examine which year has a different number of dropouts compare to the other years.

287

Post Hoc Tests

Year

Multiple Comparisons

Number_of_Dropouts
Tukey HSD

(I) Year	(J) Year	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2009	2010	-.33*	.126	.041	-.66	-.01
	2011	-.40*	.121	.005	-.72	-.09
	2012	-.45*	.133	.004	-.79	-.11
2010	2009	.33*	.126	.041	.01	.66
	2011	-.07	.117	.928	-.37	.23
	2012	-.12	.129	.797	-.45	.21
2011	2009	.40*	.121	.005	.09	.72
	2010	.07	.117	.928	-.23	.37
	2012	-.05	.125	.983	-.37	.28
2012	2009	.45*	.133	.004	.11	.79
	2010	.12	.129	.797	-.21	.45
	2011	.05	.125	.983	-.28	.37

Based on observed means.
 The error term is Mean Square(Error) = 2.868.

  *. The mean difference is significant at the 0.05 level.

Table 11

From Table 11, it shows that the number of dropout for 2009 with 2011 and 2012 is different when we look at their significance values of 0.005 and 0.004.

Model fitting

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Typeof school, Period, Gender, Grade, Reason, School[a]	.	Enter
2	.	Gender	Backward (criterion: Probability of F-to-remove >= .100).
3	.	Period	Backward (criterion: Probability of F-to-remove >= .100).

a. All requested variables entered.

b. Dependent Variable: Number_of_Dropouts

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.429	1.200		-1.190	.235
	Period	.035	.210	.009	.169	.866
	Gender	-.014	.213	-.004	-.066	.947
	Reason	.060	.016	.357	3.710	.000
	School	.049	.030	.093	1.635	.103
	Grade	-.195	.081	-.134	-2.414	.016
	Typeof_school	1.612	.426	.363	3.781	.000
2	(Constant)	-1.456	1.123		-1.296	.196
	Period	.036	.209	.010	.171	.864
	Reason	.060	.016	.357	3.737	.000
	School	.050	.030	.094	1.675	.095
	Grade	-.196	.081	-.135	-2.422	.016
	Typeof_school	1.613	.425	.363	3.798	.000
3	(Constant)	-1.393	1.060		-1.315	.190
	Reason	.060	.016	.357	3.739	.000
	School	.049	.029	.093	1.670	.096
	Grade	-.196	.081	-.135	-2.436	.015
	Typeof_school	1.611	.424	.362	3.800	.000

a. Dependent Variable: Number_of_Dropouts

The variables Reason, School, Grade and Type of school were significant at 5% level when applied conditional backward logistic regression analysis for school dropout variable and fitted model was given by equation (1).

(Y) = -1.393 + 0.06Reason + 0.049School -.196Grade + 1.611Typeof school (1).

Model interpretation

- $\beta 1=0.06$. Any unit increase in reasons, the number of dropouts increase by 0.06 keeping other (School, Grade and type of school) constants in the model.
- B3= -.196. Any unit increase in Grade, the number of dropouts decreases by -.196 keeping other (Reason, School and type of school constants) in the model.
- $\beta 0$= -1.393. Generally on average there is no number of dropouts when all the predictors in a model are constants.

Chapter 5
Conclusions and Recommendations

This study has outlined a range of factors which can and may influence dropping out from school. Dropout is influenced by a range of interacting factors, which are specific to individual contexts (and agency) of each child.

In the present study, a total of 308 dropout children were found. 27.9% of the dropouts observed are in grade8 for both government and private schools. In grade9, 10, 11 and 12, dropouts were 24.7%, 21.4%, 17.2%, and 8.8%, respectively. It shows that the making free education is not sufficient to catch up all school grade children to continue in secondary school. It is a complex social problem. Government alone cannot reduce it. To reduce the number of dropouts of secondary school grade children, only free admission & monthly fees and book distribution will not solve the problem. The parents should be made aware & motivated regarding the importance of education, at the same time programs should be established to motivate children the importance of education at all schools. The learner dropout challenge in Namibia is solvable.

Reference

- Children's defence Fund. (2001). Twenty-five facts about American children from the state of America's Children Yearbook 2001.
- U.S. Department of education. (2001). Twenty-third annual report to Congress on the implementation of the individuals With Disabilities Education act. Washington, DC.
- Windhoek: UNICEF/SSD
- Lockheed and Verspoor (2007). Introduction to Research and Evaluation in Adult Education, Windhoek.
- NAMCOL (2009-11) NAMCOL Strategic Development Plan, Windhoek: NAMCOL
- Christenson, Sinclair, Lehr, & Hurley (2000).