NETWORK SAMPLING AND ITS APPLICATION TO THE PREVALENCE OF ABORTION AMONG FEMALE STUDENTS IN NIGERIA TERTIARY INSTITUTION.

ABSTRACT

This study aimed to investigate the prevalence of abortion among female students oftertiary institutions in Nigeria and its impact on the students. The network sampling technique which makes use of an initial subject to refer other subjects with similar characteristics was used as sampling method for data collection. Data was collected from 150 students using a well-structured questionnaire, and statistical analysis using the logistic regression to determine the Wald statistics and odd ratio was conducted to conclude. The chi square was also used to determine the model adequacy which shows that the model is a good fit. The study found that factors such as rape, pressure to complete education, criticism from peers, and fear of parental reaction were influencing students' decisions to undergo abortion. Additionally, the study revealed that complications and untimely deaths were among the negative consequences of abortion. The most common age bracket for students involved in abortion was 21-25 years, accounting for 48% of cases, and the estimated prevalence of students engaged in abortion was 92.7%. Based on these findings, the study suggests promoting the use of contraceptives, such as condoms, and encouraging abstinence, especially among new students. The study also recommends that parents and guardians take an active role in monitoring their children and wards.

Keywords: *Binary Logistic Regression; Chi Square; Network Sampling; Odd Ratio; Wald Statistics*

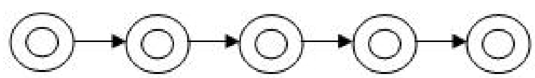
1. INTRODUCTION

Network sampling is a non-probability sampling method used when the elements of interest in surveys are rare and difficult to find. This method involves individuals selected for a study recruiting new participants from among their circle of acquaintances. It is utilized when it is challenging to find suitable participants. Unlike simple random sampling, which involves an equal chance of selection for every element of the population, network sampling does not affect probability. Instead, researchers use a first seed to choose participants, leading to the selection of several other seeds.

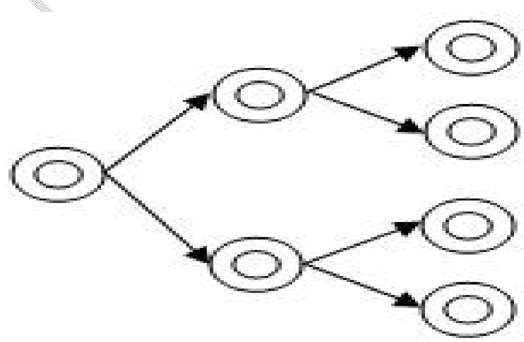
The network sampling method make use of primary data sources to recommend other potential primary data sourcesfor research. In this method, initial subjects refer additional subjects, creating a chain referral process. This sampling technique expands the sample group as existing subjects recruit future subjects from among their acquaintances. It is beneficial for studying hidden populations such as drug addicts, internet gamblers, or sex workers, who can be difficult for researchers to access. Researchers can gather adequate data for their research as the sample size grows.

Patternsof Network Sampling

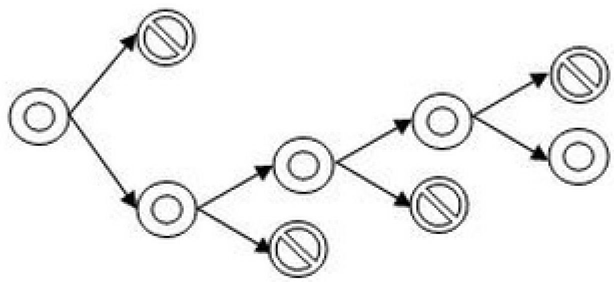
1. **Linear network sampling**: Establishing a sample group commences with a single participant introducing one referral. Upon acceptance, the referred individual is encouraged to also provide a new referral. This sequential process is repeated until the sample size reaches its desired number.



2. **Exponential non-discriminative network sampling**: The first participant introduced to thesample group provides multiple referrals, and each new individual referral provides several additional multiple referrals. This process continues until sufficient sample sizes are obtained.



3. **Exponential discriminative network sampling**: The initial person provides multiple referrals; however, only one new subject is selected among them. The choice of a new subject depends on the objective of the study and the nature of the research.



Network Sampling can be applied in various areas, such as medical practices, social research, and discord cases. This study will adopt the network sampling method to explore the rare event of "abortion" among female students. Abortion is the premature ending of a pregnancy by removing an embryo or fetus from the uterus before it reaches the stage of viability. There are controversial opinions from both supporters and opponents concerning the morality of abortion. Despite being practiced throughout human history, abortion continues to be stigmatized. It has been carried out in various ways, including using abortifacient herbs, sharpened tools, and applying abdominal pressure.

The practice of abortion, which is the termination of pregnancy, has been known since ancient times. In today's society, many people undergo the process of abortion for various reasons. Some do so to maintain their social status and career, while others seek to address their negligence and promiscuous behaviours. Additionally, some individuals undergo abortions to prevent the adverse effects of biological and chemical imbalances in the womb. In some instances, people may undergo abortion because it has been discovered that continuing the pregnancy would result in their death or cause harm to their body.

Various factors, such as social, economic, educational background, and family size, all play different roles in the context of abortion or pregnancy termination. Both male and female students should act responsibly in their sexual behavior, especially considering that many students today are sexually active. However, female students often bear the most responsibility as they are more likely to face the consequences (Alimson, 2001). These consequences include their inability to complete their education, societal stigmatization, the challenges of early motherhood, and the fear of having a child out of wedlock (Alimson, 2001).

When studying a sensitive topic such as abortion, it can be challenging to gather representative samples since the target population is very small and widely spread out. Most people involved in one way or the other are not willing to share such personal information.Hence, the population is referred to as "hard-to-reach", rare or hidden population. Researchers often use a chain-referral sampling technique to study this hidden population (Brown et al., 1999; Watters & Biernacki, 1989).

2. LITERATURE REVIEW

Network sampling gained popularity, as traditional samplingmethods like household surveys had several limitations. It is challenging to create a sampling frame for populations with specific characteristics which is against the norms or culture of the society, this makes such population hard to reach or "hidden" (Heckathorn & Cameron, 2017).

Heckathorn and Cameron (2017) in their study, made use of the respondent-driven sampling (RDS), a widely used network sampling method which employs a link-tracing design for statistical inference. Despite significant progress in RDS, several research questions remain unanswered, especially regarding the best way to calculate the variability of estimates.

According to Hassan (2016), analyzing data from different domains, such as social, communication, and information sciences, may require network sampling to draw inferences and predictions about their network. Hassan (2016) discussed various sampling methodologies (e.g., Metropolis–Hastings, random walk, importance sampling) using network sampling and applied these methods.It was noted that modelling and analyzing the large size of real-life networks could be infeasible. Additionally, many networks are inaccessible to the public due to privacy concerns.

Arkadiusz *et al.* (2021). utilized the Internet to introduce a virtual network sampling method using LinkedIn. This method has the advantage of providing quick access to respondents around the world, allowing for surveys to be conducted at a very low cost. It enables researchers to reach specific groups of respondents and conduct surveys on a global scale.

Adeleke &Ibiwoye (2011) also used network sampling to gather data on substance abuse among students in tertiary institutions. They described the sampling method as collecting seeds and generating samples from these seeds. These involve sampling from a population that shares the same social connections. According to Adeleke *et al.* (2007), the network sampling approach begins with a small number of initial respondents who then provide information leading to other members of the population with the desired characteristics. These connections then form the pool from which the second set of respondents is selected, and then the next set, and this continues.

Andrew Gelman (2016) pointed out that a small sample represents only a small portion of the population when working with atomistic data structures using traditional surveys. Conversely, a small sample from a network does not accurately represent the entire network. For instance, if you randomly sample nodes from a network and analyze the edges connecting these nodes, the resulting network will be much sparser than the original.

Abortion is a highly controversial and significant subject that has been the topic of ongoing debate for many years and will likely continue to be for years to come. There is no definitive conclusion because this issue can be viewed from numerous perspectives. Women who are faced with this decision are affected emotionally, mentally, and physically as their bodies undergo significant changes to which they are not accustomed.

The traditional sampling methods have not been effective in capturing hidden or rare events. When trying to sample hard-to-reach or elusive populations, it is often challenging to select a sampling frame using conventional techniques. These elusive populations are often termed as hidden or rare populations in the literature.

3. METHODS AND MATERIALS

3.1Research Design

This research utilized a survey method to collect data from primary sources. The data was gathered through structured questionnaires to obtain people's opinions on issues pertaining abortion. This approach was chosen to provide a systematic research framework and to test the hypotheses. Data were analyzed using binary logistic regression with the aid of Statistical Packages for Social Sciences (SPSS) version 26. The study's target population includes full-time, part-time and affiliated degree students of Yaba College of Technology in Yaba, Lagos.

3.2Sample Size Determination

The sample size for this research is derived through the recruitment of seeds. Here, we begin with an initial seed of five seeds. These five seeds would then lead to a larger sample of seeds, which forms a network; hence, it's called Network Sampling.

3.3 Logistic Regression

Logistic regression, devised by statistician David Cox in 1958, is a statistical method employed to forecast a discrete outcome from a set of predictors that may be continuous, discrete, dichotomous, or a hybrid. Unlike multiple regression, logistic regression cannot generate negative predicted probabilities (Tabachnick & Fidell, 2007). Furthermore, logistic regression yields odds ratios, serving as an effect size estimate for categorical data (Tabachnick & Fidell, 2007). In any regression analysis, the critical quantity is the mean value of the response variable given the values of the independentvariable:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + e_i, \qquad i = 1,2, \dots, n.$$

where $y \in Y$denotes the response variable, $x \in X$denotes a value of the independent variables, the $\boldsymbol{\beta}s$ values denote the model parameters, $e_i$denotes the random error term, such that, $E(e_i) = 0$, and p is the number of independent variables.

$$E(Y/x_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}$$

The quantity$E(Y/x_i)$ is called the conditional mean or the expected value of Y given the value of x. The specific form of the logistic regression model is

$$E(Y/x) = \pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{pi}}}{1 + e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{pi}}}$$

$$E(Y/x_i) = \pi(x_i) = \frac{e^{\sum_{k=0}^{p} \beta_k x_{ki}}}{1 + e^{\sum_{k=0}^{p} \beta_k x_{ki}}}$$

where$x_0 = 1$ and the transformation of the logistic function is known asthe logit transformation given by

$$e^{\sum_{k=0}^{p} \beta_k x_{ki}} = \frac{\pi(x_i)}{1 - \pi(x_i)}$$

and

$$1 - \pi(x_i) = \frac{1}{1 + e^{\sum_{k=0}^{p} \beta_k x_{ki}}}$$

then, we can have

3.3.1 Estimation of the Parameters

The maximum likelihood method is one of the most used methods for estimating a model's parameters. It is used to estimate the parameters of the logistic regression model.

Let Y_i ~ Bernoulli($\pi(x_i)$) be a random variable with$i = 1,2 \ldots, n$,and its probability mass function is given by

$$P(y_i \in Y) = [\pi(x_i)]^{y_i}[1 - \pi(x_i)]^{1-y_i}$$

The likelihood function is given by,

$$L\big(Y/\pi(x_i)\big) = \prod_{i=1}^{n} [\pi(x_i)]^{y_i}[1 - \pi(x_i)]^{1-y_i}$$

Recall that

$$1 - \pi(x_i) = \frac{1}{1 + e^{\sum_{k=0}^{p} \beta_k x_{ki}}}$$

So that,

$$L\big(Y/\pi(x_i)\big) = \prod_{i=1}^{n} \left[e^{\sum_{k=0}^{p} \beta_k x_{ki}}\right]^{y_i} \left[\frac{1}{1 + e^{\sum_{k=0}^{p} \beta_k x_{ki}}}\right]^{1-y_i}$$

Taking the log gives

$$logL\big(Y/\pi(x_i)\big) = \log\left(\prod_{i=1}^{n} \left[e^{\sum_{k=0}^{p} \beta_k x_{ki}}\right]^{y_i} \left[\frac{1}{1 + e^{\sum_{k=0}^{p} \beta_k x_{ki}}}\right]^{1-y_i}\right)$$

But the log of product is the sum of log.

$$logL[\pi(x_i)] = \sum_{i=1}^{n} \log\left(e^{y_i \sum_{k=0}^{p} \beta_k x_{ki}}\right) - \sum_{i=1}^{n} \log(1 - y_i)\left(1 + e^{\sum_{k=0}^{p} \beta_k x_{ki}}\right)$$

$$logL[\pi(x_i)] = \sum_{i=1}^{n}\left(y_i \sum_{k=0}^{p} \beta_k x_{ki}\right) - \sum_{i=1}^{n} \log(1-y_i)\left(1 + e^{\sum_{k=0}^{p}\beta_k x_{ki}}\right)$$

$$Q = logL[\pi(x_i)] = \sum_{i=1}^{n}\left(y_i \sum_{k=0}^{p} \beta_k x_{ki}\right) - \sum_{i=1}^{n} \log(1-y_i) - \sum_{i=1}^{n} \log\left(1 + e^{\sum_{k=0}^{p}\beta_k x_{ki}}\right)$$

This can be written as

$$Q = \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{pi}) - \sum_{i=1}^{n} \log(1-y_i) - \sum_{i=1}^{n} \log(1 + e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}})$$

Since

$$\pi(x_i) = \frac{e^{\sum_{k=0}^{p}\beta_k x_{ki}}}{1 + e^{\sum_{k=0}^{p}\beta_k x_{ki}}}$$

Differentiating $Q$with respect to each parameterand equating to zero gives the following systems of simultaneous equations

$$\frac{\partial Q}{\partial \beta_0} = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \frac{e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}}}{1 + e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}}} = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \pi(x_i) = 0$$

$$\frac{\partial Q}{\partial \beta_1} = \sum_{i=1}^{n} x_{1i} y_i - \sum_{i=1}^{n} x_{1i}\left(\frac{e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}}}{1 + e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}}}\right) = \sum_{i=1}^{n} x_{1i} y_i - \sum_{i=1}^{n} x_{1i}\pi(x_i) = 0$$

$$\vdots$$

$$\frac{\partial Q}{\partial \beta_p} = \sum_{i=1}^{n} x_{pi} y_i - \sum_{i=1}^{n} x_{pi}\left(\frac{e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}}}{1 + e^{\beta_0+\beta_1 x_{1i}+\cdots+\beta_k x_{pi}}}\right) = \sum_{i=1}^{n} x_{1i} y_i - \sum_{i=1}^{n} x_{1i}\pi(x_i) = 0$$

It is difficult to obtain the solution of the system of equations, so Newton Raphson numerical iteration method is used.

The procedure of the Newton Raphson iteration is given by

$$\beta_{l+1} = \beta_l - I(\beta)_l^{-1} U(\beta)_l$$

The optimal value ofβ is reached if $\beta_{l+1} \approx \beta_l$.

Where

$$U(\beta) = \begin{bmatrix} \frac{\partial Q}{\partial \beta_0} \\ \frac{\partial Q}{\partial \beta_1} \\ \vdots \\ \frac{\partial Q}{\partial \beta_p} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \pi(x_i) \\ \sum_{i=1}^{n} x_{1i} y_i - \sum_{i=1}^{n} x_{1i} \pi(x_i) \\ \vdots \\ \sum_{i=1}^{n} x_{pi} y_i - \sum_{i=1}^{n} x_{pi} \pi(x_i) \end{bmatrix}$$

This is a $p + 1$ column vector or $(p + 1) \times 1$ matrix. It is defined as a matrix of the first derivative of the log-likelihood function Q, with respect to each of the parameters, β.

And

$$I(\beta) = \begin{bmatrix} \frac{\partial^2 Q}{\partial \beta_0^2} & \frac{\partial^2 Q}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 Q}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 Q}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 Q}{\partial \beta_1^2} & \cdots & \frac{\partial^2 Q}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 Q}{\partial \beta_2 \partial \beta_0} & \frac{\partial^2 Q}{\partial \beta_2 \partial \beta_1} & \cdots & \frac{\partial^2 Q}{\partial \beta_p^2} \end{bmatrix}]$$

This is a $(p + 1) \times (p + 1)$ square matrix.

Then the inverse of $I(\beta)$is calculated, so that $I(\beta)_l^{-1}$ is also a $(p + 1) \times (p + 1)$ square matrix, so that the product $I(\beta)_l^{-1} U(\beta)_l$ is a$(p + 1) \times 1$ matrix. Since, β_l and $I(\beta)_l^{-1} U(\beta)_l$ have the same dimension, then the solutionβ_{l+1} is also a $(p + 1) \times 1$ matrix.

$$\beta_{l+1} = \beta_l - I(\beta)_l^{-1} U(\beta)_l$$

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} = \begin{bmatrix} \beta_{0,l+1} \\ \beta_{1,l+1} \\ \vdots \\ \beta_{p,l+1} \end{bmatrix} = \begin{bmatrix} \beta_{0,l} \\ \beta_{1,l} \\ \vdots \\ \beta_{p,l} \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 Q}{\partial \beta_0^2} & \frac{\partial^2 Q}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 Q}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 Q}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 Q}{\partial \beta_1^2} & \cdots & \frac{\partial^2 Q}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 Q}{\partial \beta_2 \partial \beta_0} & \frac{\partial^2 Q}{\partial \beta_2 \partial \beta_1} & \cdots & \frac{\partial^2 Q}{\partial \beta_p^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial Q}{\partial \beta_0} \\ \frac{\partial Q}{\partial \beta_1} \\ \vdots \\ \frac{\partial Q}{\partial \beta_p} \end{bmatrix}$$

Where $\hat{\beta}_0$, $\hat{\beta}_1$, …, $\hat{\beta}_p$ are the maximum likelihood estimates of β_0, β_1,…,β_p respectively.

4. RESULTS

The 150 self–administered questionnaires were distributed among the students of Yaba College of Technology, and all were recovered, giving a 100% response rate. The collected questionnaires were analyzed with the Statistical Package for Social Sciences (SPSS 26) using frequency distributions, and a Logistic regression test was used to makeinferences.

Table 1: **Demographic characteristics of the respondents**

Demographic Characteristics		**Frequency**	**Percent (%)**
Age (in years)	16-20	28	18.7
	21-25	72	48.0
	26-30	41	27.3
	31-35	4	2.7
	36+	5	3.3
	Total	***150***	***100.0***
Marital status	Single	130	86.7
	Married	19	12.7
	Divorce	1	.7
	Total	***150***	***100.0***
Level	ND1	17	11.3
	ND2	27	18.0
	ND3	13	8.7
	HND1	26	17.3
	HND2	61	40.7
	HND3	6	4.0
	Total	***150***	***100.0***
Religion	Islam	50	33.3
	Christianity	97	64.7
	Traditional	3	2.0
	Total	***150***	***100.0***

Table 1 above reveals that 19% of the respondents are in the age bracket (16 – 20) years, 48% are between 21 – 25 years, 27% are between 26 and 30 years while 3% are between 31 – 35 years and another 3% above 35 years. While 87% of the respondents were single, 12% were married and 1% are divorced.

In addition, the result reveals that 11% of the respondents were ND1, 18% were in ND II, 9% were in ND III, while 17% were HND I students, 41% were in HND II and the remaining 4% were in HND III. Lastly, 33% of the respondents are Moslems, 65% are Christians while 2% belong to traditional religion.

Table 2: Prevalence of Abortion

Statement	**Yes (%)**	**No (%)**
Ever done an abortion	11(7.3)	130 (92.7)
Do you think abortion should be legalized	50 (33.3)	100 (66.7)
Does life begin at conception	117 (78.0)	33 (22.0)
In your opinion, do you think abortion is a murder	124(82.7)	26 (17.3)
Should parents have a say in their ward(s) having an abortion	90 (60.0)	60 (40.0)

From Table.2, most female students disagree with having practised the act of abortion, which accounts for 92.7%, while 7.3% affirm it. Also, 33.3% of the population studied proposed that the law should back abortion, while 66.7% did not.

Furthermore, 60% of female students think parents should be contacted before his/her ward engages in the act of abortion, but 40% think they should not be involved.

Table 3: Effects of Abortion

	Strongly Agree	**Agree**	**Disagree**	**Strongly Disagree**
A foetus can be categorized as a person.	62 (41.3)	72 (48.0)	10 (6.7)	6 (4.0)
All forms of abortion are murder.	81 (54.0)	43 (28.7)	8 (5.3)	18 (12.0)
Vital reasoning should be taken into consideration when contemplating whether or not to have an abortion.	73 (48.7)	66 (44.0)	3 (2.0)	8 (5.3)
A woman has complete control over her own body regardless of the life that is at stake.	49 (32.7)	59 (39.3)	13 (8.7)	29 (19.3)
There are some circumstances that morally justify abortion.	46 (30.7)	66 (44.0)	14 (9.3)	24 (16.0)
Complications arise during the cause of abortion.	61 (40.7)	67 (44.6)	12 (8.0)	10 (6.7)
Abortion should be allowed, if being used as a backup contraceptive.	22 (14.7)	44 (29.3)	27 (18.0)	57 (38.0)
Most abortion leads to untimely death.	49 (32.7)	53 (35.3)	24 (16.0)	24 (16.0)

Table 3 shows that 89% of the respondents agreed that a foetus can be categorized as a person. 83% of the students agreed that all forms of abortion are murder. Also, 93% of the respondents agreed that vital reasoning should be taken into consideration when contemplating whether or not to have an abortion. In addition, 72% of the respondents agreed that a woman has complete control over her body regardless of the life at stake. While 85% agreed that some circumstances morally justify abortion.

Furthermore, 44% agreed that complications arise during the cause of abortion. Lastly, 68% agreed that abortion should be allowed if being used as a backup contraceptive.

Table 4 Reasons for Abortion

Proposition	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**
Finances should be considered as a legitimate reason to have an abortion.	18 (12.0)	37 (24.7)	27 (18.0)	68 (45.3)
Someone who got raped shouldn't be allowed to have an abortion	25 (16.7)	58 (38.7)	24 (16.0)	43 (28.6)
One shouldn't engage in the act of abortion in order to complete education.	40 (26.7)	63 (42.0)	16 (10.7)	31 (20.6)
To avoid critics from peer group, abortion is an option.	13 (8.7)	31 (20.6)	33 (22.0)	73 (48.7)
Fear of what parents will say or do when it is heard that their ward is pregnant pushes many into considering abortion.	54 (36.0)	67 (44.7)	8 (5.3)	21 (14.0)

Table 4 indicate that 37% of the respondent agreed that finances should be considered as a legitimate reason to have an abortion. Also, 55% of the respondent agreed someone who got raped shouldn't be allowed to have an abortion. In addition, 69% of the respondent agreed that one shouldn’t engage in the act of abortion in order to complete education.

Furthermore, 29% of the respondent agreed that to avoid critics from peer group, abortion is an option. Lastly, 81% of the respondent agreed that fear of what parents will say or do when it is heard that their ward is pregnant pushes many into considering abortion.

Table 5: Summary of Binary Logistic Regression Parameter Estimate

							OR 95% C.I	
		β	**S.E(β)**	**Wald**	**Df**	**Odds ratio**	**Lower**	**Upper**
Step 1[a]	Abortion should be Legalized	-0.175	0.852	0.042	1	0.840	0.158	4.458
	Life begins at conception	-0.286	0.875	0.107	1	0.751	0.135	4.170
	Abortion is a murder	2.463	0.883	7.780	1	11.743	2.080	66.299
	Parent	-0.469	0.720	0.424	1	0.626	0.152	2.567
	Constant	-3.074	0.809	14.431	1	0.046		

The Wald statistic and its associated probabilities in table 5 provides an index of the significance of each predictor in the equation. The Exp(•) column in the table presents the extent to which raising the corresponding measure by one unit influence the odds ratio and it is therefore interpreted in terms of the change in odds. If the value exceeds 1, then the odds of an outcome occurring increases, if the figure is less than 1, any increase in the predict leads to a drop in the odds of the outcome occurring.

The odds ratio associated with those people who said abortion should be legalized was 0.840. This implies that those people that said abortion should be legalized are 16% less times likely to have done abortion before than those who did not.

The odds ratio associated with those people that said life begins at conception was 0.761. This implies that those people that said life begins at conception are 24% less times likely to have done abortion before than those who did not.

The odds ratio associated with those people that said abortion is a murder was 2.080. This implies that those people that said abortion is a murder are 2.080 more times likely to have done abortion before than those who have not done it before.

Table 6: Model Adequacy

	Chi square	*Df*	*P*
Overall model significance	13.674	4	0.008
Hosmer &Lemeshow(HL)	11.394	6	0.077
R^2	0.214		

Evaluation of Binary Logistic Regression Model

The overallmodel significance, goodness of fit statistics and model summary was used to determine how effective the model expressed in Table 6 is. The overall significance is tested using Chi – square model. In this study, the model Chi – square has 5 degrees of freedom, a value of ***13.674, p < .05*** for the model**.** This indicate that the model is a good fit. The results in Table 6 shows that the ***H – L statistics*** has a significance of 0.077 which means that it is not statistically significant and therefore our model is quite a good fit. In addition, it is indicating that the logistic model explained 8.7% - 21.4% of the variation in the dependent variable.

5.Discussion of Findings

Based on the result of data analysis above, the following findings were drawn. It reveals that 19% of the students are of ages 16 – 20 years, 48% are between 21 – 25 years, 27% are between 26 and 30 years while 3% are between 31 – 35 years and 3% are above 35 years. Also, 87% of the respondents were single, 12% were married and 1% are divorced. The result reveals that 11% of the respondent were ND1, 18% were in ND II, 9% were in ND III, while 17% were HND I student, 41% were in HND II and the remaining 4% were in HND III. Lastly, 33% of the respondent are Moslems, 65% are Christians while 2% belong to traditional religion.

A logistic regression analysis was conducted to predict prevalence of abortion among 150 students using gender, family income, and family structure as predictors. A test of full model against a constant only model was statistically significant indicating that the predictors as a set of reliably model distinguished between good and poor performance (Chi – square, $p < .05$ with df = 3). In addition, the logistic model explained 11.9% - 21.6% of the variation in academic performance for single parent while it explains 18.1% - 34.2% of the variation in the academic performance for both parent family.

6. CONCLUSION

Students desire to complete their education, cases of rapes, criticism as a result of the norms from the society, and fear of what parents will do or say are some of the factors that causes the involvement of students in carrying out abortion. The study indicates that abortion has various negative consequences on the students. The findings indicated that the probability population parameters on the prevalence of students involved in abortion is very high, hence proper measures should be taken to curb it.Students should be encouraged to abstain from sex in the course of their study, while students should be advised to make use of contraceptives like condom when they find it difficult to abstain.There should also be thorough counselling for the students from school counselors, religious bodies, parents and guidance as many other diseases could bespread among from premature activity asides abortion.

Disclaimer (Artificial intelligence)
Option 1:
Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.
Option 2:
Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology
Details of the AI usage are given below:
1.
2.
3.

REFERENCES

Adeleke, I.A., Okafor, R.O. and Esan, E.O. (2007) "Horvitz-Thompson Theorem as a tool for Generalization of Probability Sampling Technique" *Journal of Scientific research and development*, Vol. 10, pp.125-134.

Adeleke, I.A, Okafor, R.O., Esan, E.O., Opara, A.I., Olayinka, O.K. and Oyeyiola, A.O. (2008) "An Adaptive Sampling Procedure for Estimating Heavy Metal Pollutants of Ground Water from Soil" *Journal of Model Assisted and Statistical Applications* Vol. 3, pp.353-

360

Adeleke, I.A& Ibiwoye, &. (2011). Network Sampling of Hard-to-Reach Population: An Application to the Substance Abuse Problem Among Tertiary Education Students in Nigeria. *International Journal of Asian Social Science, 1*, 133-141.

Alimson, A. (2001). Perceptions of sexual behaviours and knowledge about pregnancy termination among adolescents in selected states in Nigeria. *Family Planning and Reproductive Health, 25(4)*.

Andrew,G.(2016).Learning about networks using sampling.http://www.stat.columbia.edu/~gelman/research/published/hansen_paper_2.pdf.

Arkadiusz K., Adam K., Janusz D. andHanna K. (2021). Virtual network sampling method using LinkedIn.*Elsevier B.V.*: www.elsevier.com/locate/mex.

Becker H. S. (1963). *Outsiders: Studies in the Sociology of Deviance*. New York: Macmillan

Biernack P. andWaldorf D. (1981). Snowball sampling: problems and techniques of chain referral sampling. *Sociol. Methods Res.* 10:141–63.

Gelman, A. (2007). Struggles with survey weighting and regression modeling (with discussion). *Statistical Science*, 22, 153–164.

Godambe, V.P. and Thompson, M.E. (1986) "Parameters of Superpopulations and Survey Population: Their relationship and Estimation" *International Statistical review* vol. 15, pp. 3759.

Goodman, Leo. (1961). "Snowball Sampling." *Annuals of Mathematical Statistics* 32 (1): 148-70.

Granovetter, Mark. (1973). "The Strength of Weak Ties." *American Journal of Sociology* 78 (May): 1360-80, 1974. Getting a Job: A Study of Contacts and Careers. Cambridge, Mass.: Harvard University Press

Hasan, M. A. (2016). Methods and Applications of Network Sampling. *Tutorials in Operations Research* , 1-25.

Heckathorn , D.D. and Cameron, C. J. (2017). Network Sampling: From Snowball and Multiplicity to Respondent-Driven Sampling. *Annual Review of Sociology, 43*, 101-119.

Heckathorn DD. (2008). Assumptions of RDS: analytic versus functional assumptions. Presented at CDC Consult. Anal. Data Collect. Respond -Driven Sampl., Atlanta, GA

Ibrahim, I. A, Jeremiah, I, Abasi, I. J and Adda, A. O (2011). 'Patterns of complicated unsafe abortion in Niger Delta Teaching Hospital Okolobiri, a four-year review,' *the Nigerian health journal*, vol. 11, no. 4,

Sirken M. G. (1970). Household surveys with multiplicity. J. *Ann. Stat. Assoc.* 65:257–66.

Thompson S. (2020).New estimates for network sampling.

Wejnert C. (2008). An empirical test of respondent driven sampling: point estimates, variance, degree measures, and out-of-equilibrium data. *Sociol. Methodol.* 39(1):73–116.