Analysis of Tobacco Smoking Patterns in Kenya Using the Multinomial Logit Model

Samwel N. Mwenda[1], Anthony Kibira Wanjoya[2], Anthony Gichuhi Waititu[2]

[1]Department of data processing, ICT Directorate, Kenya National Bureau of Statistics, Nairobi, Kenya
[2]Department statistics and actuarial science, Jomo Kenyatta university of Agriculture and technology, Nairobi, Kenya

Email address:

Samwenda87@gmail.com (S. N Mwenda), awanjoya@gmail.com (A. K. Wanjoya), awaititu@gmail.com (A. G. Waititu)

Abstract: *Objectives:* The study aimed to determine the tobacco smoking patterns in Kenya. *Methods*: This research project used the Kenya GATS 2014 data, in which a sample of 5436 total people was interviewed. However since the research focussed on modelling tobacco smoking pattern in Kenya, data from only 4418 people was used for the analysis. Data from 1018 people in the sample was dropped because information about the individuals smoking pattern, age or work status could not be found. *Data Analysis*: The data was analysed using R-software version 3.0.2, and report presented in form of tables and graphs. *Results*: This project found out that there is likelihood of a person being a heavy smoker, light smoker or Non-smoker, if the person works in the Government and Non-government /private organization, self-employed or Unemployed. The overall effect of work status was statistically significant with a chi-square value of 129.722 (p-value<0.0001). *Conclusion*: The results show that a person's working status and their age are good predictors of a specific smoking pattern. From the results we have more people smoking as they grow old.

Keywords: GATS, Kenya, Tobacco Smoking

1. Introduction

Tobacco is the agricultural product of the leaves of plants in the genus *Nicotiana* and tobacco products are made once tobacco has been grown, harvested, cured, and processed. The tobacco products are consumed by smoking, sucking, chewing or snuffing (Gately, Iain, 2004). Tobacco use is one of the main risk factors for a number of chronic diseases, including cancer, lung diseases, and cardiovascular diseases (WHO, WHO global burden of disease report, 2008).According to World Health Organization. Non-communicable diseases (NCDs) kill more than 36 million people each year and of the NCD deaths, 29 million occur in low and middle income countries (WHO, Report on the Global Tobacco Epidemic, 2008). Report from Commonwealth health online, indicated that Non-communicable diseases (NCDs) in Kenya accounted for an estimated 29 per cent of all mortality in 2008. In the same year, the report said that most prevalent NCDs were cardiovascular diseases at twelve per cent (CHO).Cancers, non-communicable variants of respiratory diseases and

diabetes contributed six per cent, three per cent and two per cent to total mortality respectively (CHO). These figures are getting higher every day, and they may get out of control if effective interventions to contain them aren't put in place. WHO reports also indicate that NCDs are the leading causes of death in all regions except Africa, but current projections indicate that by 2020 the largest increases in NCD deaths will occur in Africa. In African nation's, deaths from NCDs are projected to exceed the combined deaths of communicable and nutritional diseases and maternal and perinatal deaths as the most common causes of death by 2030.Total tobacco – attributable deaths are projected to rise from 5.4 million in 2005 to 6.4 million in 2015 and 8.3 million in 2030 (WHO, Tobacco is the single most preventable cause of death in the world today, 2008). Lopez et al in his study found out that diseases associated with tobacco have been identified as a major cause of death in many countries and its estimated that over 350 million tobacco-related deaths resulting from similar diseases will occur in developing countries if action is not taken to prevent such an epidemic (Lopez AD; Collishaw NE; Piha T.;, 1994).The Global Youth Tobacco Survey was developed to enhance the capacity of countries to design,

implement, and evaluate tobacco control and prevention programs on students primarily between the ages of 13–15 years (GYTS, 2005). The only way that these young adults in this age group could have learnt smoking is through seeing their superior or peers smoke listened or watched advertisements in local channels through the radio or television or read in the newspapers. Research by advertising agencies has confirmed that children's personal preferences can be targeted and changed by TV advertising so one of the ways that some countries in the world have adopted to regulate smoking is to ban tobacco advertisements in public places and to restrict advertisements only during specific times, and especially when the wrong audience is watching. Solomon Nwathor give an Example, in 2004, when Nigeria signed the WHO Framework Convention on Tobacco Control (FCTC), which was ratified in 2005 and tobacco billboards in the country were taken down by 2005 and radio / television cigarette advertisement before 10pm banned. This was to ensure that public advertisement of tobacco was limited to the adults since at 10pm most children are asleep thus reducing their chance of watching advertisements on tobacco (Nwhator, 2012). The sixth session of the Conference of the parties (COP6) to the WHO Framework Convention on Tobacco Control (FCTC) was held in Moscow 18 October 2014 and In her opening speech, WHO Director-General Dr. Margaret Chan said that "as implementation of the Framework Convention reaches new heights, the tobacco industry fights back, harder and through every possible channel, no matter how devious those channels and practices are" (WHO, WHO tobacco treaty makes significant progress despite mounting pressure from tobacco industry, 2014). This shows that as much as the FCTC is trying to regulate use of tobacco, the big tobacco companies are using every resource to fight back. This has been witnessed in some countries in which the firms go to the extent of giving heavy bribes to lawmakers not to pass certain tough laws in regard to smoking tobacco. Philip Jakpor(2012) describes how British American Tobacco (BATN) undermines these efforts in Nigeria through "agricultural initiatives". He describes how through a Corporate Social Responsibility program me, BATN presents itself as a stakeholder in Nigeria's agriculture sector, targeting rural farm communities that make up 70% of Nigeria's population. Their position is a canny one – using the argument that the ban will hurt the Nigerian economy, but saying little about the harms of tobacco use (Japkor, 2012).

1.1. Statement of the Problem

According to World Health Organization Tobacco use is the leading preventable cause of death. It is on track to kill more than 8 million by 2030, by which time approximately 80% of the deaths would occur in low- and middle-income countries in which Kenya is in this category (WHO, Report on the Global Tobacco Epidemic, 2008). It is estimated that tobacco epidemic kills nearly 6million people every year that is one death every 6 seconds, including 600,000 who die from second hand smoke exposure. This means that over 10% of tobacco deaths are caused by secondary smoking. 63%of

all deaths are caused by NCDs, for which tobacco use is one of the greatest risk factors (WHO 2008). The positive fact is that all this deaths can be prevented and can be highly reduced, if proper mechanism and tough penalties are put in place. Unfortunately we have tobacco companies which are putting tens of billions of dollars each year on tobacco advertising, promotion and sponsorship. Tobacco use is a major public health problem worldwide, and it is the single most preventable cause of morbidity and mortality in humans. Tobacco use is one of the most common risk factors for non-communicable diseases (CHO). According to the Ministry of Health-Kenya, NCDs contribute to nearly 50% of all admissions in public hospitals countrywide. In Kenya, 69 per 100,000 deaths for individuals aged 30 and above are as a result of tobacco use. Five percent of all non-communicable deaths result from tobacco use, and 55% of all deaths from cancers of the trachea, bronchitis, and lung are attributable to tobacco (MoH, Kenya Health policy 2014-2030, 2008).

1.2. Objectives of the Study

1.2.1. Main Objective

To model tobacco smoking patterns in Kenya; Heavy smokers, less than daily and never smoked at all.

1.2.2. Specific Objectives

1. To find out whether work status of an individual is associated with a specific smoking pattern

2. To find out whether age is associated with a specific smoking pattern.

1.3. Research Questions

1. Does work status of an individual predicts a specific smoking pattern?

2. Does age predict a specific smoking pattern?

1.4. Justification

Given the extent to which there is an increase in the number of tobacco smokers, and the health problems associated with this, this project will be very important in determining the patterns of smoking in Kenya and the main predictors. It is after these results that we will be able to conclude the main predictors that are associated with a smoking pattern, thus we are able to prepare the intervention to counter this and reduce the tobacco health related problems caused by smoking tobacco. The health ministry has already designed very gory images that will be printed on all cigarettes packaging in a fresh attempt to reverse the growing number of puffing Kenyans. Statistics from WHO indicate that more than eight billion sticks of cigarettes were smoked in Kenya last year from just over six billion in 2013 (Nairobi News, 2015).

1.5. Limitations

The study was conducted between January and February 2014, and during these period students who were aged 15 years and above were in school. Therefore the information on

this student's smoking pattern was not captured. However, the effect of this limitation is negligible since this information is captured by the Global Youth Tobacco Survey (GYTS) which targets youths aged 13-15 years.

2. Literature Review

2.1. The Health Sector in Kenya in Respect to Tobacco Smoking

According to the Kenya health policy 2012-2030, Health is one of the components of delivering the Vision's Social Pillar given the key role it plays in maintaining a healthy and skilled workforce necessary to drive the economy (MoH, Kenya Health Policy 2012-2030, 2010).The government of Kenya spends a huge budget on health, approximately Kshs. 34.7bn for 2013/2014 for preventive and curative health services and much of this is spend on the NCDs. Total Health Expenditure per capita is US$27, government health expenditure is 5.2% of total government expenditure (MoH, Kenya Health policy 2014-2030, 2008). 31% of Total Health Expenditure is given by donors. Kenya is yet to reach the 15% of the GDP to be used on health as agreed in the Abuja convention in 2010 (WHO, The Abuja Declaration ten years on, 2010).Tobacco use is a risk factor which either increases the risk or causes some of this NCDs. Tobacco accounts for almost 6 million deaths every year (including over 600 000 deaths from exposure to second-hand smoke), and is projected to increase to 8 million by 2030 (CDC, Global Tobacco Control).In terms of attributable deaths, tobacco use is ranked as the second with 9% of global deaths are attributed to it (CDC, Global Youth Tobacco Surveillance, 2000–2007, 2008). Centers for Disease Control and Prevention (CDC) are conducting Global Adult Tobacco Survey (GATS) in all the countries, and it's doing this after every 4 years to check on how the situation is in every country .This is mainly to analyze whether the situation about smoking is improving or getting worse with time. The GATS is the global standard to systematically monitor adult tobacco use and track key tobacco control indicators. GATS is a nationally representative household survey of adults 15years of age or older, using a standard protocol. Its intention is to generate comparable data within and across countries. GATS enhance countries' capacity to design, implement and evaluate tobacco control intervention (CDC, Global Adult Tobacco Survey (GATS) — Overview).GATS has been carried out in a number of countries in Africa like Egypt Nigeria and Kenya. According to the Kenyan fact sheet it was found out that 15.1% of current tobacco users were men with only 0.8% women. 6% overall in the sample were found to be daily tobacco smokers while current men cigarette smokers and daily men cigarette smokers were found to be 15.1% and 11.6% respectively. The overall former daily tobacco smokers were found to be 28.5% meaning that over a quarter of the Kenyan population is either using tobacco or have previously used tobacco and/or

tobacco related products (KNBS, 2015).

2.2. Tobacco Smoking Views in Different Countries

Different challenges due exist in different countries, but the aim/goal cuts across all countries: To reduce/regulate smoking tobacco. In more developed countries like Malaysia, it is illegal to advertise tobacco or tobacco products between 6am-10pm. This is to ensure the advertisements on smoking tobacco reaches to as minimal people as possible and especially not the children. However even as the governments continue to put this restriction through banning advertisements, the tobacco companies look for other ways to ensure they are relevant in the market. One of the main ways they do this is by the social corporate responsibility in which they have projects in regions especially areas where tobacco is grown (Wikipedia, 2015). They do this by building schools, hospitals etcetera, and with this they are able to encourage continued growth of tobacco. They also offer good returns to tobacco growers, to the extent of giving them free tobacco seeds; fertilizer, and insurance cover for the crops in case of a loss. Bearing in mind that most of the developing countries are still struggling to maintain their economy with less borrowing, raising taxes by the government on tobacco and related products, raises a lot of cash though payment of taxes. Some governments have however been able to reduce tobacco intake in their countries and an example according to Jamie Smyth is Australia which was having very high smoking records since 1959. The government of Australia forced tobacco companies to sell their products in plain packets carrying graphic health warnings, and today Australians are consuming less tobacco than at any other time. This prompted health campaigners to claim the world's first law forcing cigarettes to be sold in plain packs is reducing smoking rates. Ireland and New Zealand followed Australia to become the second and third country respectively to commit to this (Jamie , 2015).The inherent health hazards of smoking have caused many countries to institute high taxes on tobacco products and anti-smoking campaigns are launched every year in an attempt to curb tobacco smoking.Smoking cessation is the most reliable method of significantly decreasing the prevalence of a wide variety of serious diseases. Promoting a non-smoking environment is one of the biggest steps to contribute to good health in the society by helping both smokers and non-smokers maintain well-being and by enabling the government to decrease healthcare cost significantly. As a way of regulating smoking Zulu et al, suggested that smoking should be taught earlier in schools to ensure children knows the effects of smoking from an early age. This is according to a survey done in Zambia 2010 which recommended that since most adults who smoke may have started the habit when they were adolescents, a feasible intervention would be to control or prevent tobacco smoking among in-school adolescents which was the target for the current study (Zulu R; Siziya S; Nzala SH;, 2009). Inclusion of tobacco control in primary and secondary school syllabi should be considered. This is because there is no much subject in the syllabus to

teach about smoking and its effect, and therefore children learn this habit from parents or friends who smoke tobacco or use tobacco related products. Tougher laws with tougher penalties need to be introduced to ban smoking near a child or smoking near people who don't consume tobacco.On his study on measures taken to 'To Weaken the Giant' in Malawi, Brenya, Edward(2012) concluded that regional approach to disseminate tobacco control ideas would lead to a common tobacco control campaign, which could promote compliance with tobacco control prohibition regime in the African region (Brenya, 2012).The first known advertisement in the United States of America was for the snuff and tobacco products of P.Lorillard and Company were placed in the New York daily paper in 1789. Advertising was an emerging concept, and tobacco-related advertisements were not seen as any different from those for other products: their negative impact on health was unknown at the time. Local and regional newspapers were used because of the small-scale production and transportation of these goods (Wikipedia, 2015). The first real brand name to become known on a bigger scale in the USA was "Bull Durham" which emerged in 1868, with the advertising placing the emphasis on how easy it was "to roll your own" (Durham, 1868)

Scholars started noting the parallel rise in cigarette consumption and lung cancer and by the 1930s had begun to investigate this relationship using the methods of case-control epidemiology. According to Robert N. Proctor, Lung cancer has become a formidable disease, killing about 1.5 million people per year globally (Robert, 2012).

2.3. Tobacco Smoking Models

2.3.1. Model of Cigarette Smoking by Muller et al

Franz Hermann Müller at Cologne Hospital in 1939 published the first such study, comparing 86 lung cancer 'cases' and a similar number of cancer-free controls. Müller was able to show that people with lung cancer were far more likely than non-cancer controls to have smoked, a fact confirmed by Eberhard Schairer and Eric Schöniger at the University of Jena in an even more ambitious study from 1943. These German results were subsequently verified and amplified by UK and American scholars: in 1950 alone, five separate epidemiological studies were published, including papers by Ernst Wynder and Evarts Graham in the USA and Richard Doll and A Bradford Hill in England. All confirmed this growing suspicion, that smokers' of Cigarettes were far more likely to contract lung cancer than non-smokers. Further confirmation came shortly thereafter from a series of prospective 'cohort' studies, conducted to eliminate the possibility of recall bias (Müller, Lungencarcinom. Zeitschrift, & Krebsforschung, 2011). The theory here was that by following two separate and initially healthy groups over time, one smoking and one non-smoking, matched by age, sex, occupation and other relevant traits, you could find out whether smoking was a factor in the genesis of lung disease. The results were unequivocal: Doll and Hill in 1954 concluded that smokers of 35 or more cigarettes per day increased their odds of dying from lung cancer by a factor of 40 (Doll, R.; Hill , AB.;, 1954).Hammond and Horn, working with the American Cancer Society on another large cohort study, concluded that same year that the link had been proven 'beyond a reasonable doubt'(Hammond, EC; Horn, D.;, 1954)

2.3.2. Model of Cigarette Epidemic by Kunze

In an attempt to model cigarettes epidemic, Kunze had proposed a two phase model of the relative evolution of smoking prevalence among physicians compared with the general population according to this formulation(which would seem to be well supported by historical data from developed countries), smoking rates among doctors initially exceeded those of the general public due largely to their greater disposable income and their ignorance of, or indifference to, knowledge about the health hazards of smoking. Subsequently, as the dangers of smoking become increasingly apparent, physicians give up smoking earlier and in greater numbers than the general population with the result and their prevalence falls below that of the population at large (Kunze, 1989).

2.3.3. Model of Cigarette Epidemic by Lopez

Lopez in his abstract on the descriptive model of cigarettes epidemic for developed countries found out that cigarettes are the cause of deaths of one in two of their persistent users. This is to mean that 50% of persistent cigarettes users die of tobacco related health problems (Lopez AD; Collishaw NE; Piha T.;, 1994)

Lopez continues to say that transition through the various stages of the tobacco epidemic can be effectively characterized by changes in 3 variables:
  i. Prevalence(the percentage of the adult population who smoke regularly)
 ii. Consumption(the amount smoked per adult in a given period, eg, per year)
iii. Mortality due to smoking(numbers of deaths attributed to smoking, by age, sex, and classified cause of death)

2.4. Research Gaps

Although many models have been proposed to model tobacco smoking and its effect, not a lot has been done on the tobacco smoking pattern using work status and age. This is due to the nature of the data that was previously collected because of the structure of the questionnaire. GATS 2014 questionnaire design has helped to address this issue by having questions on smokers, divided into heavy smokers (heavy smokers) and light smokers (light smokers).Therefore this proposal will create a multinomial Logit model that will compare the odds of an individual being a daily smoker (heavy smokers) or less than daily smoker (light smokers) and compare it with 'not being a smoker at all'(non-smoker) status depending on the age and work status of an individual.

3. Methodology

3.1. Study Population

The GATS 2014 was a national household-based sample

survey of adults aged 15 years and above who were usual members of a household. By definition, a household was defined as a person or group of persons who normally reside together in the same compound under one or several roofs, are answerable to the same head, and share a common cooking arrangement. The usual members of the households included those who spent at least half of the duration in a year in the household, including students in boarding institutions who were available at the time of the survey. Persons residing in non-conventional households such as the army barracks, hotels, prisons, and hospitals among others were excluded from the sample.

3.2. Sample Design

The Kenya GATS 2014 was a cross-sectional household sample survey covering all 47 counties to provide national estimates for tobacco indicators for rural and urban areas. In addition, the sample was constructed to allow separate national estimates for male and female indicators. A three-stage cluster sample design was adopted for the survey with the first stage involving selection of sample points ('clusters') while second and third stages involved selection of households and eligible individuals, respectively. A representative sample of 5,376 households was drawn with a target of interviewing one eligible adult, aged at least 15 years, within the sampled households. The survey used the fifth National Sample Surveys and Evaluation Programme (NASSEP V) master sample frame that was developed and maintained by KNBS. The frame was developed using the enumeration areas generated from the 2009 Kenya Population and Housing Census and has 5,360 clusters split into four equal sub-samples. In the first stage, 192 clusters (102 urban and 90 rural) were selected from the NASSEP V frame. A uniform sample of 28 households from the listed households in each cluster was selected in the second stage of sampling. The last stage of sampling was done using PDAs at the time of survey, where one individual was randomly selected from all eligible listed household members. From the sample, only 4418 individual data was used for the analysis since the research focused on current smoking status. Data for 1018 people was missing either the age, work status or the specific smoking pattern. Some of the people had quit smoking.

3.3. Multinomial Logistic Regression Model and Interpretation for Data Analysis

The data was analysed using the R software version 3.0.1. Data from the 2014GATS, a nationally representative survey of households (n = 4418), was used to generate descriptive statistics and characterize tobacco use in the country. A multinomial logistic regression model was used to evaluate the relationships of tobacco smoking pattern with age and work status.

Multinomial logistic regression is for modelling nominal outcome variables, in which the log odds of the outcomes are modelled as a linear combination of the predictor variables. It is used to predict categorical placement in or the probability of category membership on a dependent variable based on multiple independent variable are either dichotomous or continuous (Friedman, Jerome, Hastie, Trevor, & Tibshirani, Rob, 2010)

When creating a multinomial logit model, a linear predictor function which constructs a score from a set of weight that are linearly combined with the explanatory variables is constructed. From the secondary data GATS 2014, this project viewed smoking tobacco as a response variable, work status and age as the predictor variables. Therefore the project looked at conditional distribution of the response variable smoking status given the 2 predictor variables work status and age.

Considering a random variable (the response variable) Y_i takes several discrete variables indexed $1,2,3,\ldots\ldots\ldots j$. The 3 discrete variables are 1=Heavy smoker, 2=Light smoker and 3=Non-smoker thus j=1,2 and 3.

Letting $\pi_{ij} = Pr(Y_i = j)$ denote the probability that the i th response falls in the j th category.

In this project

π_{i1} is the probability that the i th respodent is a Heavy smoker

π_{i2} is the probability that the i threspodent is a Light smoker

π_{i3} is the probability that the i threspodent is a Non smokersmoker

The project assumed that the response categories are mutually exclusive and exhaustive then we have

$\sum_{j=1}^{j} \pi_{ij}$ =1 for each i, in that the probabilities add up to 1 for each case, and we have only $j - 1$ parameters. From the response variable smoking status, once the probability of heavy smoker and light smoker is known it's then it is possible to calculate probability of non-smoker by subtracting from 1.

Since the data is grouped, auxiliary random variable will be introduced which involve the count of response from various categories. Letting n_i denote number of cases in the ith group and lettingY_{ij} denote the number of response from the ith group that fall in the j th category with observed value Y_{ij} . Letting i represent age group, $n_i$number of smokers in the ith age and $Y_{i1} Y_{i2}$ and Y_{i3} the number of heavy smokers, light smokers and non-smokers at all in the i th age group.

Therefore $\sum_j Y_{ij} = n_i$ in that the counts in the various response categories will add upto the number of cases in each group.

For the individual data n_i=1 and Y_{ij}becomes the indicator variable which takes the value 1 if the ith response falls in the jth categoryand 0 otherwise and $\sum_j Y_{ij} = 1$since 1 and only 1 of the indicators Y_{ij}can be on for each case. Using GATS 2014 data the 4418 cases/records will let Y_{ij}be 1 if the ith person is a heavy smoker and 0 otherwise.

The probability distributions of the counts Y_{ij}given total $n_i$is given by the multinomial distribution

$$Pr\{Y_i = y_{i1}, \ldots, Y_{ij} = y_{ij}\} = \binom{n_i}{y_{i1,\ldots,}y_{ij}}\pi_{1i}^{y_{i1}}, \ldots, \pi_{ij}^{y_{ij}}$$

Considering the model where probability depends on the vector X_i associated with the ith individual. In this project the model showed the probabilities of being heavy smoker, light smoker or non-smoker at all depends on their work status and age.

The project took one of the response variables to be the baseline/reference category and then calculated the log odds for all other categories relative to the baseline/reference. It then let the log odds be a linear functions of the predictor variables. This project used the 'non-smoker' status as the reference category and then calculated the odds that a member of group i falls in the category j as opposed to the baseline $\frac{\pi_{i1}}{\pi_{iJ}}$. The project looked at the odds of being a 'heavy smoker' rather than 'non-smoker' and odds of being 'light smoker' rather than being 'non-smoker'

Assuming that the log odds of each response follows a linear model

$$n_{ij} = log\, \frac{\pi_{ij}}{\pi_{iJ}} = \alpha_j + X_i'\beta_j$$

Where α_j is a constant and β_j is a vector of regression co-efficient for $j = 1,2,3,\ldots\ldots\ldots,J-1$

The project used $j = 3$ categories and therefore contrasted categories 'heavy smoker' vs 'non-smoker', 'light smoker' vs 'non-smoker'.

Contrast between 'heavy smoker' vs 'light smoker' can be calculated as below

$$log\left(\frac{\pi_{i1}}{\pi_{i2}}\right) = log\left(\frac{\pi_{i1}}{\pi_{i3}}\right) - log\left(\frac{\pi_{i2}}{\pi_{i3}}\right)$$

The research project also calculated the Relative Risk Ratios defined as the ratio of probability of event occurring in an exposed group to the probability of event occurring in comparison, non-exposed group.

$$RR = \frac{P\ eventexposed}{P\ eventnotexposed}$$

In the project 'event exposed' was heavy smoker or light smoker, while 'event not exposed' was the non-smoker.

3.4. Test for Statistical Significance

Test for statistical significance tell the probability that the relationship we find is due only to random chance. The project used Chi-square test to find whether the model was statistical significant. R-square value was calculated to find out whether the model fits the data well. The research hypothesis stated that there was association between tobacco smoking with age and work status, while a null hypothesis state there was no association between tobacco smoking with age and work status. The project specified the probability of committing type 1 error, alpha=0.05, meaning that it accepted the probability of 5% of making type 1 error of assuming association between tobacco smoking with age and work status when really it doesn't exist.

4. Data Analysis

Data from the GATS 2014 was analyzed using R software version 3.0.2 to determine what factors contributed to a person being a heavy smoker, light smoker or a non-smoker at all depending on their work status and their age. Multinomial logistic regression was used to estimate association between relevant prediction variables work status and age with tobacco smoking patterns. Odds ratio associated with each outcome were reported and interpreted. Relative risk ratio corresponding to the various smoking pattern were reported and interpreted.

Table 1. Descriptive statistics of smoking status, gender and work status of the individuals

Gender And Work Status of the individual	Smoking Status of the individuals			
Female	Heavy smoker	Light smoker	Non-smoker	Total
Self-employed	3	3	764	770
Student	0	0	161	161
Unemployed	13	5	1,072	1090
Working For Government and NGO/private	4	1	312	317
Male				
Self-employed	164	36	661	861
Student	1	3	210	214
Unemployed	68	14	314	396
Working For Government and NGO/private	108	29	472	609
Total	361	91	3966	4418

From the Table 1above, we have 2338 females and 2080 males representing 53% and 47% of the total sample that was used for analysis. We had 108 male and 4 female who work in government and non-government/private organization who are heavy smokers, representing 96.4% and 3.6% of males and females respectively. The highest number of heavy smokers is the self-employed males which stand at 164. The unemployed male heavy smokers are 68 while for the unemployed female smokers, we have 13 who are heavy

smokers and who are also the highest number of smokers for females in the categories. Student male heavy smoker and light smoker is 1and 3 respectively, while we have no female who is a heavy smoker or a light smoker. From the tabulation, we have the students representing a very small proportion of tobacco smokers and this was due to the fact that the survey was conducted during the period Jan 2014 & Feb 2014 when the school going term was on, and therefore since the PDA was used to select the respondent in the survey, if a student

was selected and he/she wasn't available, then they weren't interviewed. The largest portion of students who were affected were the ones in boarding high schools, university and tertiary colleges. The total smokers both heavy and light smokers was 206 for the self-employed while for the government and non-government /private employees was 142 and the unemployed was 100. The total numbers of non-smokers were 3966(89.77%) while the total number of tobacco smokers was 452(10.23%)

Table 2. Multinomial logit coefficients with respective Relative Risk Ratio (RRR)

	Coefficient's	Std. Err.	Z	P>Z	[95% Confidence. Interval] Lower	Upper	Relative Risk Ratio	Std. Err.	Z	P>Z	[95% Confidence. Interval] Lower	Upper
Smoke Status Heavy Smoker Work status												
Self-Employed	0.7009727	0.140888	4.98	0.000*	0.424838	0.977107	2.01571	0.283989	4.98	0.000*	1.529343	2.65676
Working For Government And NGO/private	0.9087335	0.153023	5.94	0.000*	0.6088139	1.208653	2.48118	0.379677	5.94	0.000*	1.83825	3.348971
Age	0.0128777	0.003282	3.92	0.000*	0.0064457	0.01931	1.01296	0.003324	3.92	0.000*	1.006466	1.019497
Intercept	-3.373899	0.183308	-18.41	0.000*	-3.733176	-3.01462	0.03426	0.006279	-18.41	0.000*	0.023917	0.049064
Light smoker Work status												
Self-Employed	0.6905837	0.282311	2.45	0.014*	0.1372652	1.243902	1.99488	0.5631756	2.45	0.014*	1.147132	3.469125
Working For Government And NGO/private	1.027964	0.296628	3.47	0.001*	0.4465845	1.609344	2.795369	0.8291841	3.47	0.001*	1.562965	4.99953
Age	0.0028621	0.006856	0.42	0.676	-0.010574	0.016299	1.002866	0.0068751	0.42	0.676	0.9894813	1.016432
Intercept	-4.403256	0.359262	-12.26	0.000*	-5.107397	-3.69912	0.012237	0.0043964	-12.26	0.000*	0.0060518	0.0247454

All assumptions are based on "non smoker" as the reference category in this analysis
All results are generated at 95% confidence level

4.1. Interpreting Variable Age in Relation to the Smoking Pattern Model

From the results in table 2 above, a one unit increase in variable age, is associated with an increase in log odds of being a heavy smoker by the amount 0.0128 than being a non-smoker at all and that of being a light smoker by the amount 0.00286 than being non-smoker at all. However, the p-value (0.000) associated with heavy smoker is highly significant but the p-value (0.676) associated with light smoker isn't significant at 0.05. Therefore this shows that age is not a good predictor of an individual being a light smoker. But as age increases, there is a very high chance of someone of either gender becoming a heavy smoker rather than becoming a light smoker. The relative risk ratio for a one-unit increase in the variable age is 1.013 for being heavy smoker vs. non-smoker at all, while the relative risk ratio for a one-unit increase in variable age is 1.0029 for being a light smoker vs nonsmoker at all. The result for heavy smoker are significant at $\alpha = 0.05$ but results for light smoker are not significant at the same value of $\alpha = 0.05$. But the results shows that there is a higher risk of converting from non-smoker to a heavy smoker than there is in converting from non-smoker to light smoker as age increases.

4.2. Interpreting Variable Working for the Government and NGO/Private in Relation to the Smoking Pattern Model

The log odds of being a daily smoker vs not being a smoker at all will increase by 0.9087943 if moving from unemployment to being employed by the government or the NGO's. On the other hand the log odds of being a less than daily smoker vs not being a smoker at all will increase by 1.0279811 if moving from unemployed to being employed by the government or the NGO's. The p-value for heavy smoker and light smoker was 0.000 and 0.001 and these findings were significant at $\alpha = 0.05$. This shows that when someone is employed by the government or the NGO's they have almost equal chances of either being daily or light smokers, but the higher chance is that they may end up being light smokers. The relative risk ratio switching from unemployed to being employed by the government or NGO is 2.48118 for being a heavy smoker than being a non smoker at all. In other words, the expected risk of being a heavy smoker is high for the people who are employed by the government and NGO. Also, The relative risk ratio switching from unemployed to being employed by the government or NGO is 2.798369 for being a light smoker than being a non smoker at all. In other words, the expected risk of being a light smoker is higher for the people who are employed by the government and NGO.

This might be due to the fact that most organization both private and government have adopted the anti-smoking policy in their work places, therefore discouraging the smoking habit.

4.3. Interpreting Variable Self-Employment in Relation to the Smoking Pattern Model

The log odds of being a daily smoker vs not being a smoker at all will increase by 0.7010161if moving from unemployment to being self-employed while the log odds of being a less than daily smoker vs not being a smoker at all will increase by 0.6905652 if moving from unemployment to being self-employed. The p-value associated with heavy smoker and light smoker was 0.000 and 0.014 therefore the findings were significant at $\alpha = 0.05$. The relative risk ratio switching from unemployed to self-employed is 2.01571 for

being a heavy smoker than being a non smoker at all. In other words, the expected risk of being a heavy smoker is high for the people who are self-employed. Also, the relative risk ratio switching from unemployed to being self-employed is 1.99488 for being a light smoker than being a non smoker at all. In other words, the expected risk of being a light smoker is high for the people who are self-employed. However the higher chance is for this people becoming heavy smokers than light smokers. This habit could be caused by the reason that since most self-employees don't has rules on smoking, a person can smoke at any given time in their work place.

Comparing self-employment with formal employment (government and NGO/private) there is a higher chance of heavy smokers in self-employment than in government employment.
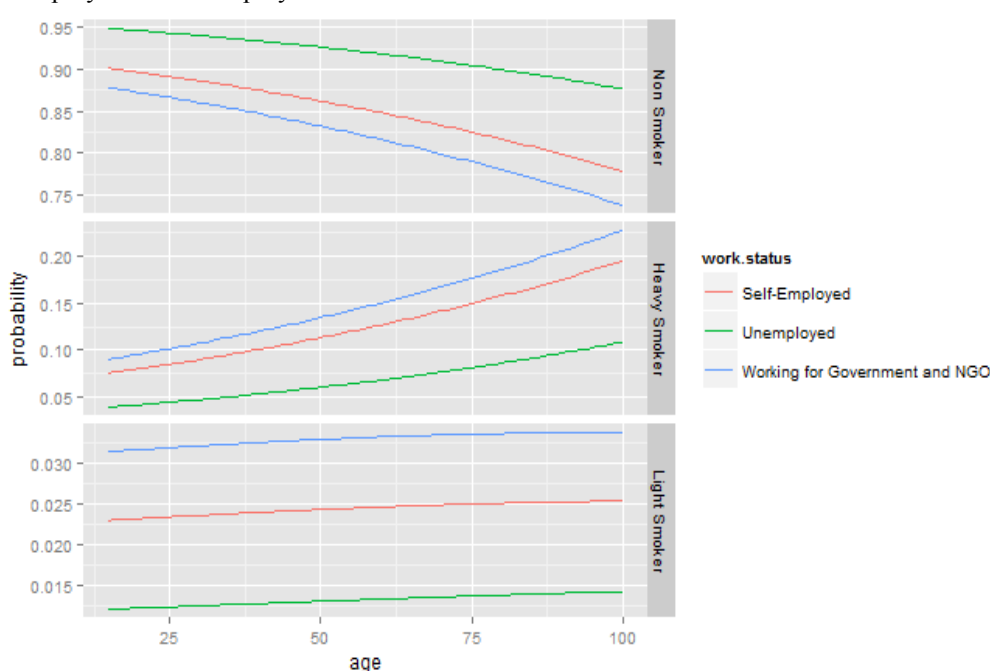
Graph 1. *predicted probabilities vs age; which shows relationship of age and smoking pattern as age increases*

Table 3. *Summary of the model fit*

Log-likelihood	
Model	-1620.65
Intercept-only	-1685.51
Chi-square	
LR (df=8)	129.722
p-value	0.0001
R2	
McFadden	0.038
Cox-Snell/ML	0.029
Cragg-Uhler/Nagelkerke	0.054

From the graph of predicted probabilities against the age by level of the 3 working status; non-government /private and government employee, self employed and unemployed it is evident that in all the 3 levels the number of smokers both daily and less than daily increases with age. However, the numbers of heavy smokers seem to increase faster than that of light smoker. This could have been caused by light

smokers converting to heavy smokers with increase in age, instead of quitting smoking tobacco.

From the summary of the model, log- likelihood value of the model with the intercept only(null model) is -1685.51 while the full model with age and work status is -1620.25. The reduction shows that the model is better at predicting a person's smoking status using work status and age. The chi-square value $x^2 = 129.722, p < 0.0001$ is statistically significant at 0.05 shows that the overall model is predicting whether someone's smoking status better than it was with intercept only. Therefore work status and age contributed significantly to fit of the model. The R-Squared values, McFadden, Cox-Snell/ML and Cragg-Uhler/Nagelkerke are treated as measure of the effect of size, however unlike in standard multiple regression, it doesn't represent the amount of variance in the outcome variable. However bigger values would indicate a better fit (Jon StarkWeather and Amanda Kay Moske)

5. Conclusion

The objective of the study were met, since from the objectives we have seen there is a positive association of smoking and age, in that the more a person grows old, the more they are likely to be smokers, whether heavy a smoker or a light smoker. Therefore the variable age is a good predictor of a person's smoking pattern. Whether a person is employed or not can also predict their smoking pattern. Depending on where one is employed can tell whether they have a higher chance of following a particular smoking status. Finally from the results, we have more people converting from not being smokers at all to being either daily or light smokers. Therefore intervention of educating people as they grow about dangers of smoking should always be emphasized so that we can have the reverse that is having more people convert from smoking to not smoking at all.

Anti-smoking campaigns need to be emphasized more on the self-employed, since most of the emphasis has been placed on government and non-government /private organization.

References

[1] Brenya, E. (2012). Struggling To Weaken the Giant: Litigation as a Measure to Compel the Adoption of Tobacco Control Instrument in Malawi. *African Journal of Political Science and International Relations, 6* (7), 155-166.

[2] CDC. (2008). *Global Youth Tobacco Surveillance, 2000–2007.* Atlanta: CDC.

[3] CDC. (n.d.). *Global Adult Tobacco Survey (GATS) — Overview.* Retrieved January 5, 2015, from CDC website: http://nccd.cdc.gov/gtssdata/Ancillary/Documentation.aspx?SUID=4&DOCT=1

[4] CDC. (n.d.). *Global Tobacco Control.* Retrieved January 6, 2015, from CDC website: http://www.cdc.gov/tobacco/global/index.htm

[5] CHO. (n.d.). *Non-communicable diseases in Kenya.* Retrieved December 14, 2014, from Commonwealth health online: http://www.commonwealthhealth.org/africa/kenya/non_communicable_diseases_in_kenya

[6] Doll, R.; Hill , AB.;. (1954). The mortality of doctors in relation to their smoking habits. *BMJ*(1), 1451-5.

[7] Durham, B. (1868). Old West Durham Neighborhood Association. North Carolina.

[8] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition.* USA: Wiley.

[9] Friedman, Jerome, Hastie, Trevor, & Tibshirani, Rob. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*(33), 1.

[10] Gately, Iain. (2004). A Cultural History of How an Exotic Plant Seduced Civilization. Diane. *Tobacco Control*, 3–7.

[11] GYTS. (2005). Global Tobacco Surveillance System (GTSS). *J Sch Health*(75), 15-24.

[12] Hammond, EC; Horn, D.;. (1954). The relationship between human smoking habits and death rates: a follow-up study of 187,766 men. *JAMA*(155), 1316-28.

[13] Jamie , S. (2015, Feb 3). Australians stub out cigarettes in wake of plain-pack law. Australian, Australian.

[14] Japkor, P. (2012). how British American Tobaco undermines the WHO FCTC through agricultural initiatives. *TobaccoControl*(21), 220.

[15] Jon StarkWeather;Amanda Kay Moske(2011). Multinomial logistic regression. Retrieved March 26,2015, from statistics website. http://www.unt.edu/rss/class/jon/benchmarks/MLR_JDS_Aug2011.pdf

[16] KNBS. (2015). *Global Adult Tobacco Survey.* Retrieved Jan 4, 2015, from KNBS website: http://www.knbs.or.ke

[17] Kunze, M. (1989). Current smoking habits in Europe. *European Conference on Tobacco priorities and strategies* (pp. 1-3). The Hague, The Netherlands: international Union Against Cancer and the Dutch Foundation on Smoking and Health

[18] Lopez AD; Collishaw NE; Piha T.;. (1994). A descriptive model of the cigarette epidemic in developed countries. *Tob Control*(3), 242-7.

[19] MoH. (2008). *Kenya Health policy 2014-2030.* Nairobi-Kenya: MOH Kenya.

[20] MoH. (2010). *Kenya Health Policy 2012-2030.* Nairobi: MOH Kenya.

[21] Müller, F., Lungencarcinom.Zeitschrift, & Krebsforschung, f. (2011). Commentary: Lung cancer and tobacco consumption. *international journal of Epidemiology*.

[22] Nairobi News. (2015, Feb 2). Health ministry to introduce new scary packs for cigarettes. Nairobi, Nairobi, Kenya: Nation Media Group. Retrieved Feb 6, 2015, from http://nairobinews.co.ke/health-ministry-to-introduce-new-scary-packs-for-cigarettes/

[23] Nwhator, S. (2012). Nigeria's costly complacency and the global tobacco epidemic. *J Public Health Policy, 33*(1), 16-33.

[24] Robert , P. N. (2012). The history of the discovery of the cigarette-lung cancer link: evidentiary traditions, corporate denial, global toll, Tobacco Control. *Tobacco Control, 87*(e91), 21.

[25] WHO. (2008). *Report on the Global Tobacco Epidemic.* Geneva: World Health Organization.

[26] WHO. (2008). *Tobacco is the single most preventable cause of death in the world today.* Geneva: WHO.

[27] WHO. (2008). *WHO global burden of disease report.* WHO.

[28] WHO. (2010). *The Abuja Declaration ten years on.* Abuja: WHO.

[29] WHO. (2014). *WHO tobacco treaty makes significant progress despite mounting pressure from tobacco industry.* Retrieved January 4, 2015, from WHO Website.

[30] Wikipedia. (2015, Jan 22). *Tobacco advertising.* (Wikipedia, Producer) Retrieved Feb 5, 2015, from Tobaco advertising website: http://en.wikipedia.org

[31] Wikipedia. (2015). *Tobacco advertising*. Retrieved Feb 6, 2015, from http://en.wikipedia.org: http://en.wikipedia.org

[32] Zulu R; Siziya S; Nzala SH;. (2009). tobacco smoking prevalence among in-school adolescents aged 13-15 years: baseline for evaluation of the implementation of the FCTC in Lusaka district,Zambia. *medical journal of Zambia, 35*(3).