Sade Ayinde

ECON-181 – Exploratory Data Analysis with R

**Write-up**

As developing economies continue to mature, labor force participation has proved necessary to sustained growth. Nigeria’s labor force is strongly defined by its’ informal sector, comprised of small businesswomen and men, skilled artisans, market vendors, tailors, and more. The main consequences of the underground economy are divided by the macro-level, market-level and household-level. On the macro level, the informal economy distorts labor market information and limits the effectiveness of public policies and services. On the market level, there is a limited access to formal support, which in turn, limits productivity, growth and innovation. Lastly, at the household level, informal employment results in a lack of social insurance. While informal employment marginally reduces unemployment, its untaxed and unmonitored nature is often seen as an impediment to actualized growth of a country’s GDP. This project aims to study the differences (quality of life, education, etc.) among those participating in informal and formal sector employment.

***Background***

The terms “informal economy” and “informal sector” were coined by Anthropologist Keith Hart, in “Informal Income Opportunities and Urban Employment in Ghana” (1973), where he conducted research on the migration patterns and urban labor markets. During his field research, Hart focused on urban areas of Southern Ghana in the 1960s. At the time, informal employment constituted as economic activities of low-income populations. Due to unattainable requirements, unskilled and illiterate workers found limitations and barriers to entry in the urban labor market. Over the years, Hart and other academics have continued to contribute to the surrounding literature on informal employment and informal sector participation. In economic and sociological literature, there are four main theories of thought that contribute to our understanding of the informal sector. As characterized by Martha Chen, HKS Lecturer and Coordinator of Women in the Informal Economy Global Organization (WIEGO), the informal economy is mainly seen through a Dualist, Structuralist, Legalist, or Voluntarist perspective.[[1]](#footnote-1) Each perspective argues different motivations for participation in the informal sector, however, this can be covered in another analysis.

***In Context: Nigeria***

As mentioned earlier, as its population increases, Nigeria’s labor force continues to depend highly on informal sector participation. As the country continues to see rapid urbanization in its major cities of Lagos and Abuja, rural to urban migration continues as well. As one of Africa’s top urbanizing countries[[2]](#footnote-2), Nigeria makes for an ideal case study. With this in mind, I wanted to use household level data that allowed a deeper analysis specific to the countries regions and provinces, as opposed to country level data. I would like to compare the returns of education in both sectors of employment, coupled with different measures of quality of life. In this case, quality of life refers to incidences of food insecurity, inconsistent income, inconsistent medication, and shortage of fuel for cooking[[3]](#footnote-3). I am expecting to see that the higher levels of education completed might result in increased participation in the formal sector. However, I have a few hesitations because of the limited amount of jobs in Nigeria. Because of this, what tends to happen is that people complete their undergraduate education, but take different jobs that are not as relevant to their degree or become self-employed (i.e. informal employment). Additionally, I do not have any presumptions about quality of life indicators measured against informal vs. formal sector participation.

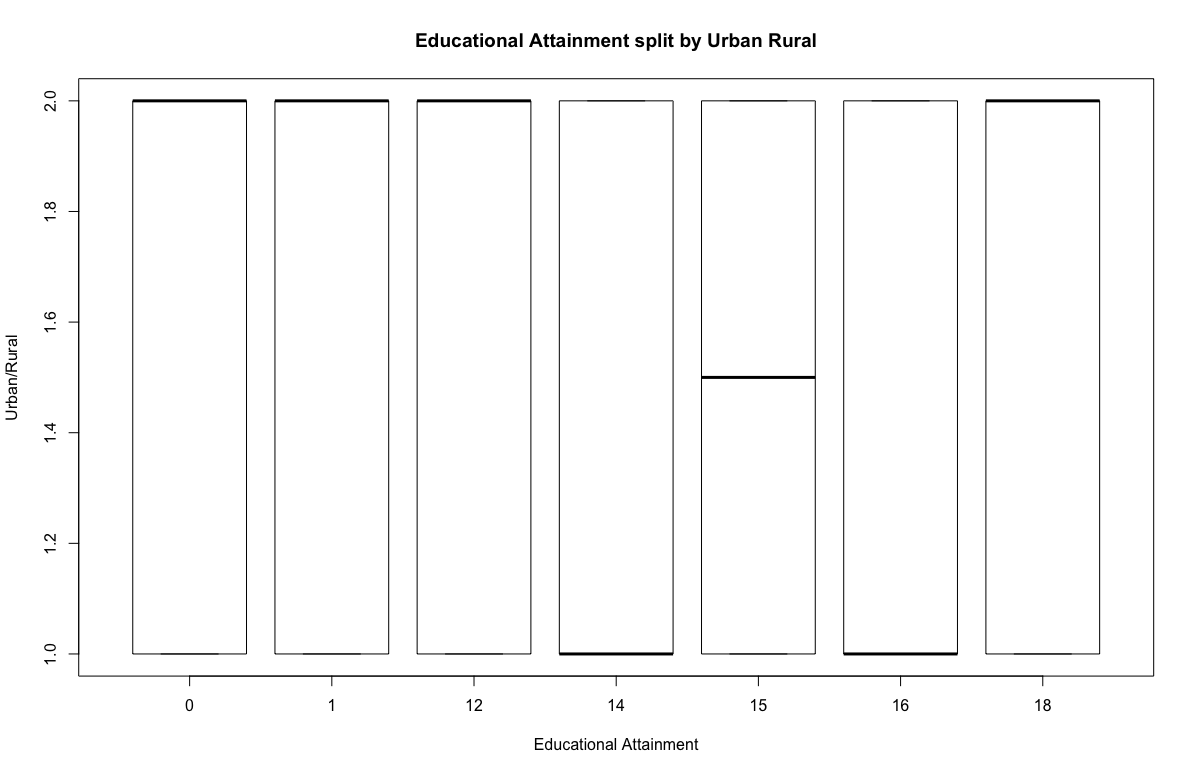
The methodology will consist of a standard OLS linear regression, with informal/formal sector participation as the independent variable. The data is from the round 6 of the Afrobarometer Survey of Nigeria in 2015. The Afrobarometer is an Africa-based, continent-wide, household survey data collection and analysis project dating back to 1999. The Afrobarometer asks households a series of questions regarding socio-demographics and characteristics (*where majority of my values come from),* civic engagement, taxation, equality and gender, institutions and leaders, crime & security, economy, democracy and so on.

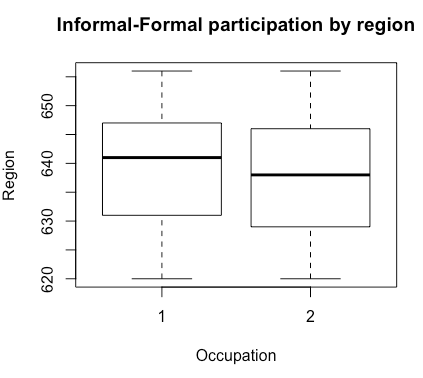
For the purposes of this project, I focused my efforts on demographic questions (region, urban/rural landscape), employment/occupation, living standards, and sample characteristics (gender, age, and education level). For employment/occupation classification of “informal sector”, I focused on legal, non-agricultural forms of informal employment. In the document titled “Codebook Values”, I was able to recode informal employment based on these characteristics. Please view the word document for additional information.

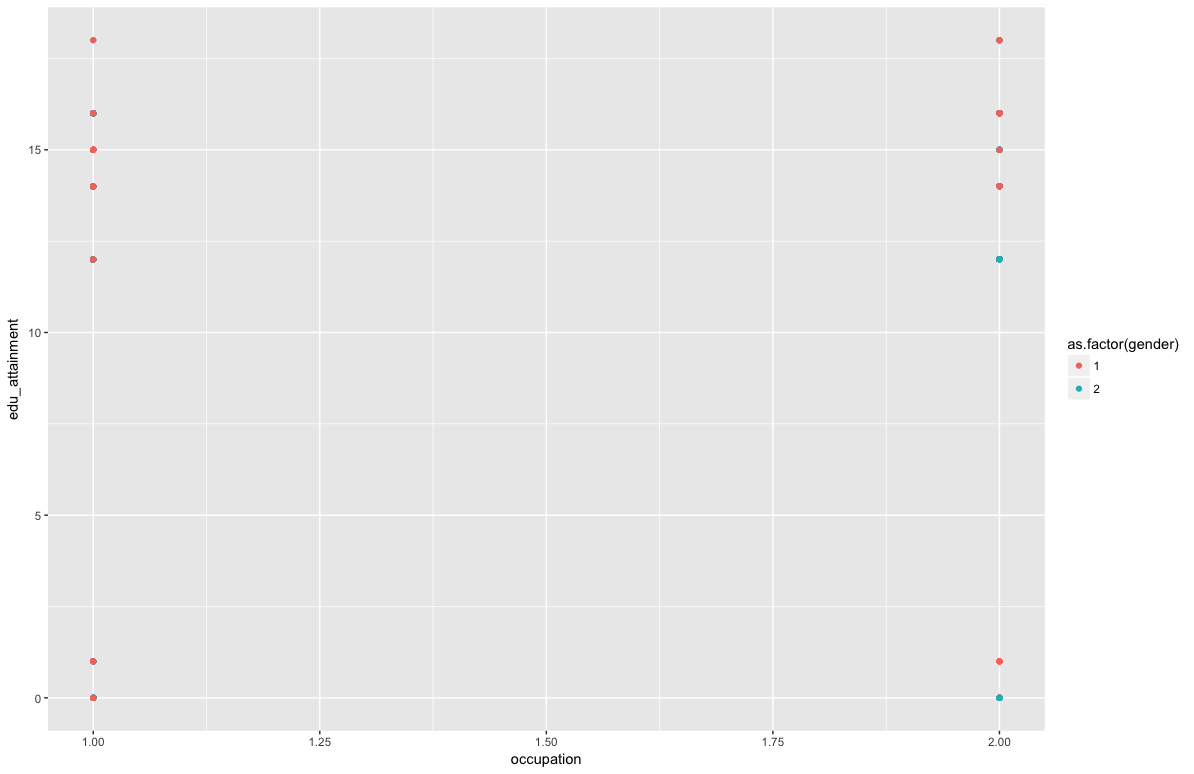
***Data (so far)***

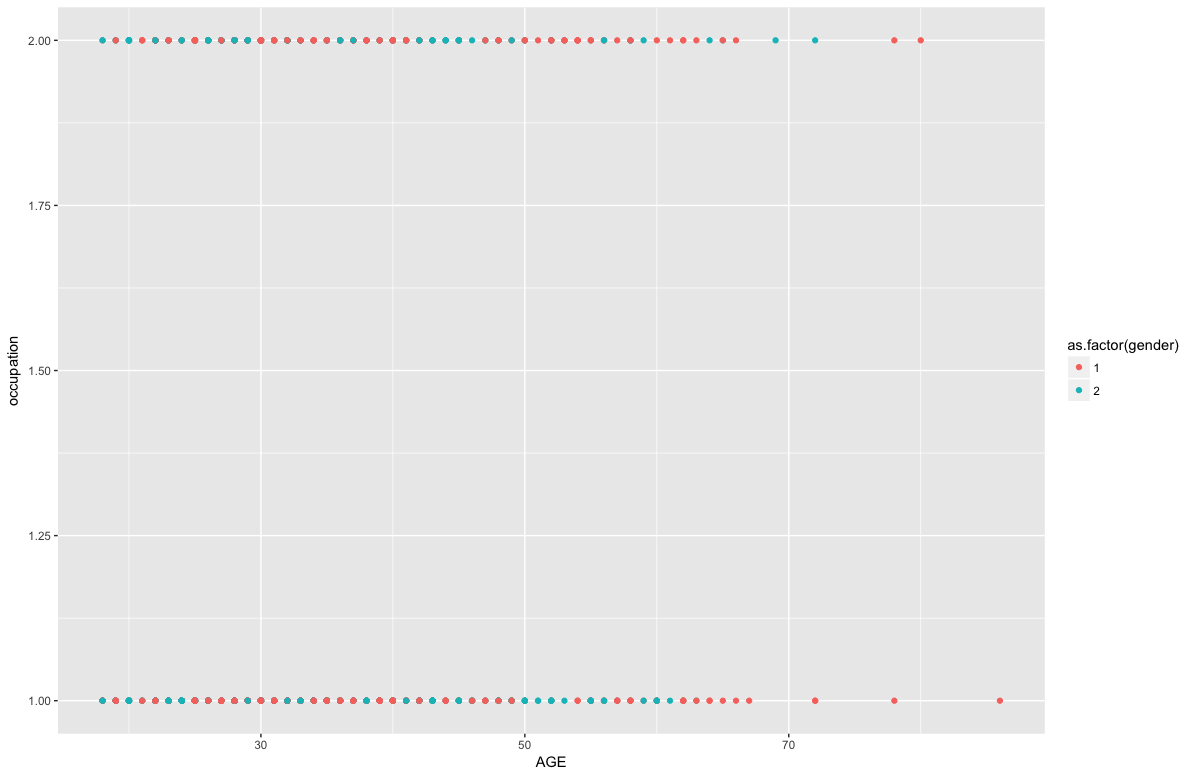
After reading in the CSV file, I renamed the variables to make the data more accessible to the user. Afterwards, I recoded the *informal employment*, based on categorizations (“vendor/hawkers”, “unskilled manual worker”, “artisan/skilled manual worker”, “housewife/homemaker”). While employment like “supervisor”, “mid-level professional”, “upper-level professional”, “clerical or secretarial” or “security services” or “retail shop” was classified as *formal employment*. These classifications are debatable (as academics spend their work defining the two types of work but I tried my best to define the two based on majority of the literature on informal employment).

After cleaning the data, recoding values and exploring it. I began to make some graph and displays. (Please see below). I hope to create more graphs and summary statistics that show the differences between the two groups. I was able to do some of this analysis by using the commands, “group by” and “summarise”, but since I have more nominal variables, I ran into some trouble with that option. I figured that the best method to use this would be through double histograms with ggplot. I hope to continue exploring the differences between the two groups before the regression analysis at the beginning of April via office hours and email.









1. Chen, Martha. WIEGO. “The Informal Economy Debate: Four Dominant Schools of Thought”. Article. N.d [↑](#footnote-ref-1)
2. Cairo, Egypt and Lagos, Nigeria are Africa’s #1 and #2 Megacities, respectively. [↑](#footnote-ref-2)
3. *Based on questions from Afrobarometer Survey* [↑](#footnote-ref-3)