Spring 2023

Understand the problem of the Haitian data management using a difference equation : A case study on the calculation method used by the Office of National Identification (ONI) to generate the list of voters for the general elections of 2015 and 2016.

Jacques A. Demezier

SHAWNEE STATE UNIVERSITY

Understand the problem of the Haitian data management using a difference equation : A case study on the calculation method used by the Office of National Identification (ONI) to generate the list of voters for the general elections of 2015 and 2016.

A Thesis

By

Jacques A Demezier

Department of Mathematical Sciences

Submitted in partial fulfillment of the requirements.

for the degree of

Master of Science, Mathematics

July 25, 2023

Accepted by the Graduate Department

7/28/2023

Graduate Director, Date

The thesis entitled **'Understand the problem of the Haitian data management using a difference equation :A case study on the calculation method used by the Office of National Identification (ONI) to generate the list of voters for the general elections of 2015 and 2016'.**

presented by **Jacques A Demezier**, a candidate for the degree of **Master of Science in Mathematics**, has been approved and is worthy of acceptance.

7/28/2023

Date

Graduate Director

07/29/2023

Date

Student

ABSTRACT

This study put emphasis on the problem of data management facing the Haitian Government. Such a problem the researcher has addressed with a case study on the calculation method used by the National Office of Identification (ONI). Since the age voting population deriving from the calculation method used by ONI was an accumulation of those expecting to attain the age voting at time (t+1) adding to those already reached the age voting at time (t) without being exposed to any type of constrains, the goal was to prove how that said population could not have been exempt for haven't being exposed to any constraints. Using a mathematical model, the researcher proved how the published electorate list by the Provisional Electorate Council (CEP) for the general elections of 2015 and 2016 failing to objectively reflects the age voting population due to lack of data management related to the omission from not subtracting the age voting population on a constant basis from death and migration.

Thus, to tackle this problem of data management , the researcher applied simultaneously an exponential equation with death and migration not being computed whose purpose is to validate the calculation method used by ONI, and a difference equation with death and migration computed to validate the hypotheses adding to the path analysis technique to check the type of correlation between each variable. By confronting the respective equations, the results confirmed that the age voting population used by ONI wasn't effectively subtracted from death and migration while the difference equation (model) as well as the path analysis results proved otherwise. Even though the results of the study exposed the invalidity of that said final list used by the CEP, further research need to be done especially a case study (at a micro level with a polling center) for a better understanding of the dynamics of the Haitian data management.

ACKNOWLEDGEMENT

Perseverance was the turning point that led to the materialization of this paper that will remain a classic in term of Haitian Data management. Since all things can be portrayed with the duality of palpable and unseeable, I, Jacques A Demezier, is pleased to show his gratitude to the "Supreme Energy" as well as my "Protectors" who gave me strength and courage throughout my master's study. Likewise, I cannot be silent for not sharing special thanks to those who stood beside me when facing hard time. From this line, I would like to thank my family especially my kids (Sephora, Jael, Jamiyah, and Benayah Demezier), BC&WC, Fritzner Auguste, Jean Senat Fleury, and Dr. Richener Noel, who played a big role by facilitating me retrieving all the related data from the National Office of Identification (ONI). Additionally, how to forget my Graduate Department Chair and thesis advisor Dr. Douglas Darbro for all his support, his guidance, and his belief in my determination to complete this thesis. All I can say is thank you; but since you have held my hands and made me see the light, I return those blessings to you and your family. Now that I am looking at a new horizon, I wish the spirit of Shawnee State University will enhance my success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF GRAPHS

Chapter I : Introduction

With the fall of the Duvalier regime in 1986 , Haiti has been transitioning from twenty-nine years with a political regime with a president for life to embrace the path of a new pollical regime called democracy where citizens have to vote to elect the state officials. The criteria on who can vote and the frequency of the election's terms are defined in the electorate decree normally taken by the government in power with regard to the Haitian constitution. For the first election following the Duvalier regime (1988), the age voting population was set to 18 and older, which will continue to be in effect ever since. In regards of the governing body that manages the elections, it is identified under the name of the " Electorate Council". Despite the first Haitian constitution post Duvalier(1987) recommend a permanent " Electorate Council", this prerogative is still hasn't been met yet. Because from 1987 to now, the Haitian government has decreed at least thirty temporary " Electorate Council" identified under the name of " Provisional Electorate Council (CEP).

But in terms of the elections process, the Haitian Government in most cases has always been assisted whether technically or financially. Here, the financial aspect and elections supervisions that are part of the technical aspect are beyond the scope of this study; for this paper, the main focus is based on the other technical aspect, which is data management. Such an assistance that usually came from specific missions from European Union (EU) and the Organizations of American States (OAS) whose purpose was to assist the Electorate Council logistically in data managing whether prior of before the elections results. Additionally, this study won't seek either the reason why, the Haitian government has been so dependent of other international countries when it comes of election management. Therefore, there is a fact in terms

of non-consistency of general census, where the Haitian government does not express too much interest in population management. For example, from 1950 to 2016, the country has held only three general censuses (1950, 1982, and 2003) used in thirty elections from 1988 to 2015.

It is also important to specify that between the elections held prior 2006, the voting age population had been separated from registered voters. That means, it was a situation where voter registration was optional , whereas it had become mandatory with the decree taken by the government Boniface/Latourtue(2005) on the creation of the National Office of Identification (ONI) whose mission was to implement the National Identification Card (NIC). With such a decree, the National Identification Card (NIC) will serve as both personal identification and voter registration card at one hand, where that said institution will become the governing body from which the Provisional Electorate Council (CEP) obtains their data that constitute the potential list of voters published prior to each election, on the other hand. Along the way, it becomes important to know the type of mechanism was put in place by the National Office of Identification (ONI) to gather or to generate their data. In other words, what calculation method or what type of procedure is used by the National Office of Identification (ONI) to generate their data? Such a question that will be clarified when testing the chosen hypotheses.

With this lack of population management, data on the age voting population will become very difficult to establish without proper statistical study. And it was according, the researcher aims to understand this problem of the Haitian data management with a case study on the calculation method used by the National Office of Identification (ONI) to generate the list of voters used in the general elections of 2015 and 2016. With this case study, the researcher is seeking to provide recommendations on the main causes characterizing the non-viability of the

age voting population data used by the Provisional Electorate Council (CEP) for the general elections of 2015 and 2016.

Thus, to address this problem of the Haitian data management, the researcher is going to use a mathematical model and path analysis technique for the purpose of proving the importance of data accuracy for better population control and election credibility. From this line, the goal will consist of comparing the general list of voters published by the Provisional Electorate Council (CEP) with the model of the study, which is going to identify by a mathematical model using a difference equation. Such a difference equation that will be used to show why is the current calculation method used by the National Office of Identification (ONI) failing to accurately predict voter registration while the path analysis technique will also be used to investigate the type of correlation existing between each variable or how each variable has been affected by each other. Along the way, by conducting this research, the goal is to describe and analyze "What are the factors making the published Haitian electorate list of 2015 and 2016 failing to accurately predict the age voting population?"

Background of the Problem

Data management is one of the greatest challenges facing the Haitian government. In this country, rare are those that their birth certificates are exempt from any type of error. This problem of an effective system capable to identify and retrace the course of the Haitians population has been addressed by the government Boniface/Latortue (2004-2006). The project had a dual purpose, providing identification number for part of population having less than 18 as well as National Identification Card (NIC) for those having 18 and over (ONI, Article2, 2005). With that decision, DGI (Direction Générale des Impôts or General Tax Office) no longer became responsible for producing personal identification cards. The National Office of

Identification (ONI) was created to assure the production of the National Identification Card (NIC) that will serve both for personal identification and voter's registration card. Along the way, Haitian citizens already possessed the traditional ID card as well as those being 18 years and plus were ability to procure or apply to the National Identification Card (NIC), which became at the same time the required identification for any legal documents.

Ever since the program was launched, the Haitian government as well as some other entities involve in the electoral data management use the available data on the National Identification Card (NIC) holders to track the total potential voters or the age voting population. For the Parliament and the Presidential elections, held respectively in August and October 2015, and November 2016, ONI reported a total of 5,871,450 and 6,189,256 potential voters. But to investigate if those lists of potential voters have been subtracted from death and migration, the researcher has discussed the matter with some officials from the National Office of Identification (ONI) throughout multiple interviews. Based on the outcomes of the conversations, it was revealed that those lists were an accumulation of all Haitians having 18 and plus starting from 2005 without being subtracted from those who might be died or migrated from 2003 to 2016. This is a situation, once registered for the National Identification Card (NIC), you became part of a counting system, where there is no mechanism that can regularly track or scrutinize the total age voting population, so it could be up to date.

Thus, considering that potential voters are the population aged more than 18 and applied for the National Identification Card (NIC), in this study, being potential voters is characterized by those who applied and received their National Identification Card (NIC). That means, the exact potential voters as well as the turnout participation rate must be related to those who received their National Identification Card (NIC). From this line, since the electorate list did not

consider the effects of deaths and migration on the age voting population, it is a particular case of the Malthusian model where only the voter inscription rate is proportional to the population size (Smith, 1977; Marion and Lawson, 2008; Luo, 2007).

Indeed, since it is common that a population cannot exponentially growth without experiencing any type of constraints, and it was accordingly that the main focus of this study is to address the issue with a balance equation. That is, since the electorate list used by the Provisional Electorate Council (CEP) hasn't been scrutinized from deaths and migration , the researcher believes that it was the omission of not reducing on constant basis the expected age voting population from death and migration that prevents the Provisional Electorate Council (CEP) to provide data that objectively reflects the age voting population.

Statement of the Problem

With the creation of National office of Identification (ONI) that became the governing body that generated data on the age voting population, we should expect an interdependence relationship between all other Haitian entities involving in data management. But when investigating those institutions , it has been noticed that entities that are supposed to manage population data are not fully operational with the exception of the Haitian Institute of Informatics and Statistics (IHSI). Despite the Haitian Institute of Informatics and Statistics (IHSI) is the only entity that is fully operational, but it doesn't show any difference in term of structural mechanism that can facilitate data sharing with the other entities (especially to the National Office of Identification that manages data of the age voting population) adding to lack of data availability on death rate and net migration. At this level, one could have asked where the National Office of Identification (ONI) got their data from and how they were computed. From this line, to know the age voting population depends on population data, the National Office of

Identification (ONI) should have been relied on the Haitian Institute of Informatics and Statistics (IHSI) that provide data on all aspects of the Haitian population, the Haitian Justice Department that should provide data on the Haitian citizens that can fully exercise their civils rights and political rights, the Civil State Office that should provide data of death certificates, and the Haitian Ministry of Interior that should provide data on migration and emigration.

But while collecting the data, it was revealed that data on death only available from the United Nations database or any international institutions involve in data management, but not from the Haitian Civil State Office. Which is similar between the National Office of Identification (ONI) and the Haitian Institute of Informatics and Statistics(IHSI) that deal with all aspects from the Haitian population data, and the Haitian Ministry of Interior that is managing migration and emigration. This lack of population tracking is a big concern in terms of data management, where one could have said that this problem could not have been the object of a study if there was no concern in term of data management or it could have had a different purpose if the Haitian government was capable to count its population.

Along the way, with the dual purposes of the National Identification Card (NIC), which is personal identification and voter registration card, it created a big concern in terms of knowing the true turnout rate. That is, since identification cards and voter registration constitute a sole statistic, it requires constant update and close relationship between all entities involved in data management. Since this study is based on the general elections of 2015 and 2016, where the last census was held in 2003, it will be very difficult with this twelve or thirteen-year gap to have accurate data on the population without a census. From this line, unlike this study , for the elections held in 1957, 1988, 1990, and 2006 the age voting population came respectively from the general census of 1950, 1982, and 2003. Hereafter, we noticed a net discrepancy between the

following censuses and the other elections, where there was a fourteen-year gap between the general census of 1982 and the general elections of 1996, which is eighteen years between the general census of 1982 and the elections of 2000, and lastly a gap of twelve or thirteen years between the general census of 2003 and the general elections of 2015 and 2016. From that perspective, the age voting population for the elections of 1957, 1988, and 1990 aligned with the common acceptation on census population that should be held every ten years, whereas it was the contrary for the general elections of 1996, 2000, 2015, and 2016.

Thus, since there is no liaison between state entities that were supposed to share population data, this study is going to investigate if whether or not the general lists of voters published by the Provisional Electorate Council (CEP) in 2015 and 2016 have been subtracted from death and migration. Since the calculation method used by the National Office of Identification (ONI) was simply an addition between those having 18 years old at time (t) and those expecting to become adult at time (t+1) without being scrutinized from the variables' death and migration, this study will use the power of mathematic and statistic to prove otherwise. That is, knowing there exists any population that can growth without being exposed to certain constraints, the researcher believes that the age voting population at time (t+1) is a function of the population at time (t) adding to the expecting age voting population at time (t+1) depreciated from death and migration (which aligns with the provided definition on population dynamics[1]).

Additionally, since the Haitian Institute of Informatics and Statistics (IHSI) is the sole Haitian institution that provide on a constant basis data on the Haitian population, this study will also investigate how the age voting population for the general elections of 2015 and 2016

[1] https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/population-dynamics

aligned with the three censuses by estimation of 2009, 2012, and 2012 from the Haitian Institute of Informatics and Statistics (IHSI) with respect to tables comparison.

Purpose of the Study

This study will be conducted to test out the reliability of the Haitian electorate list published for the general elections of 2015 and 2016 with respect to a difference equation. Considering that the last Haitian census was held in 2003, where there was a gap of twelve or thirteen years from the general elections of 2015 and 2016, the provided data on the age voting population for those said elections appears to be very problematic. Additionally, with the terrible earthquake (2010) that caused more than 300,000 deaths adding to the significant increase in the migratory flow started in 2010 with the addition of the new migratory routes towards other horizons other than North America and the Islands, it would have been statistically biased to affirm that the data on the age voting population from the National Office of Identification (ONI) wasn't exposed to any type of constraints.

Thus, since there is no liaison whatsoever between all the Haitian entities involved in data management, the accuracy of data on age voting population requires adequate statistical procedures that should consider all factors that can affect that said population whether positively or negatively. It was; accordingly, this study aims to contradict the calculation method used by the National Office of Identification (ONI) using a difference equation where the dependent variable (age voting at time t+1) equal (the population expected to become 18 at time t+1) adding to those already reached 18 depreciated from death and migration.

Significance of the Study

This study addresses one of the biggest issues facing the Haitian Government, which is data management. With the incapacity of the Haitian Government to adequately quantity the age voting population, it shows not only the poor quality of the elections but why all the elections held during the past the 25 years have been contested or linked up to certain fraud due to data management issue. In other words, if the Haitian Government is unable to adequately quantify the population how can elections results, turnout participation rate be trusted? In fact, with this study one can easily show what makes the published electorate list of 2015 and 2016 inaccurate. Likewise, other researchers may have decided to look at the impact of the internal migration on the accuracy of the electorate list during the election day. Similarly, this study can serve as an example for those involved in data management , especially the state entities, to develop services that can track the population on a daily basis. Furthermore, with this lack of population tracking that isolates all the institutions that were supposed to share population data, the findings and the recommendation from the study can serve as guidelines that can encourage those in power to bring data management among the state top priority. Additionally, with this study policy makers will have the opportunity to enforce rules on population data for preventing identity fraud or any related issues to lack of population management.

Main research question and hypothesis

This study aims to describe "What are the factors making the published Haitian electorate list of 2015 failing to accurately predict the age voting population?"

Hypothesis

Since the electorate lists used by the Provisional Electorate Council (CEP) haven't been scrutinized from deaths and migration , I believe that it was the omission of not reducing on constant basis the expected age voting population from death and migration that prevents the Provisional Electorate Council (CEP) to provide data that objectively reflects the age voting population.

Additional research questions and hypotheses.

1. **What is the real impact of the variable's death and migration on the age voting population ?**

   Hypothesis: There is no population that is not exposed to certain constraints.

2. **How can the migratory flow affect the 2015 electoral list?**

Hypothesis: The Haitian migratory flow, since 2010 has experienced a significant increase with the addition of the new migratory routes towards other horizons other than North America and the Antilles, which eventually must be considered in the study of the age voting population where there is no mechanism allowing Haitian citizens living outside the country to vote.

3. **What is the real impact of the growth rate on this electoral list?**

Hypothesis : The population already reached the age voting and those aspiring to reach the age voting are exposed to the same constraints (death and migration), but the impact of the growth rate is more significant on the electoral list than the variables death and migration .

4. **What is the gap between the variables of the model and those from the official data on the age voting population?**

Hypothesis: the difference between the variables of the model and the official data other than those from the Provisional Electorate Council (CEP) is irrelevant.

Research Design

The research main goal is to predict the " age voting population" at time (t+1) from the expected age voting population at time (t+1), age voting at time (t), death, and migration. Which makes the study explanatory or predictive. Since the study is being conducted under the " Exempt IBR review, it doesn't require any participant. In terms of instrumentation, the study will be performing using statistical software like R, Excell, and SAS. But addressing the problem of the Haitian data management consist of presenting two mathematical models and a predictive modeling using path analysis. For instance, the first mathematical models will be performed using an exponential equation with death and migration not being subtracted , which will serve to validate the calculation method used by the National Office of Identification (ONI) while the other one identified under a difference equation will be completed with death and migration computed for the purpose of testing the research hypotheses. Then concluding with a path analysis technique that will be used to test the nature of the correlation between the nine exogenous (Pop18, Pop18d, Pop18m, Pop18dm, Pn_Pop18, Pba, Pop18_d, Pop18_m, and Pop18_dm) and the three endogenous variables(RegisteredVoters, Votes, and Turnout).

Assumptions, Limitations, and Scope

Even though the explanation toward the incapability of the Haitian Government to provide accurate data on the age voting population can be interpreted one way to another. In this

paper, it was agreed that the factors involving in the State's inability to objectively quantify the number of potential voters can be addressed through these following:

- The Final lists of voters have not been subtracted from the deceases of 2005, date marking the commencement of the counting of the General list, to those who dead the last day set by the CEP (Provisional Electoral Council) to procure the NIC card, May 2015.

- The lack of connection between the Civil State Office dealing with deaths certificate, Haitian Institute of Informatics and Statistics (IHSI) dealing with all the data on the population, Haitian Ministry of Interior dealing with migration and emigration , and the National Identification Office that is managing the registered voters.

- Problem of deaths reports by many Haitians to the Civil State Office.

- The published final electorate haven't been subtracted from those leaving the national territory during the 2005-2016 periods.

With the lack of information on the second and the third points above, this study will only cover the first and the last factors. Additionally, it comes to specify with the complexity related to the internal migration, this aspect is beyond the scope of this study. Furthermore, despite the study will provide a better understanding on the dynamics of the Haitian age voting population, but it cannot substitute a case study which could have been dealing with personal information of the population under study like National Identification Card (NIC) number and so forth. But that doesn't prevent the results of this study from being used by any institutions involved in data management, where the results will depend on the chosen variables and the intent of the institution conducting the study. That is, some variables may be different from those used in this study, but in the case of the electorate population, the procedure will remain the

same where the age voting population will always depreciate from all variables that have a direct impact on the growth of that said population.

Definition of Terms.

The keys terms used in this study are mostly common, except for voter registration that vary from one country to another. Along the way, definitions that knowledgeable peers would know at a glance were not included. That is, the study will skip the definitions of the terms like : " difference equation, differential equation, path analysis, data management, population growth" and leave them at the discretion of the readers. Therefore, clarity will be brought towards the term " voter registration and electorate council". If for all countries voter registration always depend on age status, but the process of registration varies from one country to another. In the case of the study, all citizens aged 18 and more who are entitled to enjoy their civil rights can apply for a National identification Card (NIC). Once registered for the National Identification Card (NIC) and the card has been received, it can serve as both personal ID and voter registration card. Such a mechanism automatically makes all the National Identification Card (NIC) holders considered as potential voters. In regards of the notion of the " Electorate Council", it is the entity that is put in place by the Haitian Government to organize the elections. It is supposed to be permanent, but until now, the Haitian Government is still struggling to make it a reality. Because the last twenty-one general elections matched the total of temporary electorate created via decree and being identified under the name of Provisional Electorate Council (CEP).

Summary of chapter I.

Chapter 1 announced the problem statement to be explored. At this level, it was brought to the attention of the readers the main element that propelled the researcher to conduct this study, which is data management issues. With a difference equation and a path analysis technique , the study aims to address the inaccuracy of the electorate list used by the Provisional Electorate Council (CEP) in the elections of 2015 and 2016.

This chapter also proves the lack of population tracking that isolates all the institutions that were supposed to share population data. Additionally, with this study, policy makers will have the opportunity to enforce rules on population data for preventing identity fraud or any related issues to lack of population management. After the problem statement has been clarified, it was followed with the primary research question from which the main focus is to describe "What are the factors making the published Haitian electorate lists of 2015 and 2016 unreliable?"

In term of the assumptions , this chapter provides four factors that prevent the Haitian Government to objectively quantify the number of potentials . Among these four factors, two were dropped for lack of available data while two were adopted, which are the predictors death and migration. In terms of hypothetical answer , the main focus is to prove that it was the omission of not reducing on constant basis the age voting population from death and migration that prevents the Provisional Electorate Council (CEP) to provide data that objectively reflects the age voting population. But , to verify this hypothesis, the research design will be based on a mathematical model and a predictive model using respectively a difference equation and a path analysis technique. With the difference equation, the goal is to predict the dependent variable " age voting population" at time (t+1) from " age voting population" at time (t) and expecting age

voting at time (t+1) depreciated from death and migration at one hand while the path analysis will look how each variable is affected by each other, on the hand.

Additionally, it was also specified that this study will only consider the worldwide Haitian migration while the Haitian internal migration is beyond the scope of this study. Similarly, the study cannot substitute a case study which could have been dealing with personal information of the population under study like National Identification Card (NIC) number and so forth. However, in the case of the electorate population, it was also shown that the procedure will remain the same where the age voting population will always depreciate from all variables that have a direct negative impact on the growth of that said population.

Then before digging deeper in developing the theoretical framework in chapter 2, clarification was also made in chapter I to some key concepts like " voter registration " and electorate council". Once the theoretical orientation has been adopted, the next step will consist of proceeding to research methodology in chapter 3. Then follow with research results in chapter 4 by ending with conclusion and recommendations in chapter 5.

Chapter II : Literature Review

The need for population counting can be retraced far back in history, where it can be assimilated as a corollary of space management, social organization or other. For example, in 2 CE, China's Han Dynasty records the oldest surviving census data, showing a population of 57.7 million people living in 12.4 million households[2] . Thus, from ancient times to now, the need for population counting has never been an individual project, but instead an initiative deriving from those responsible to manage the living space, the state, the social organization , etc. Thus, knowing your population will continue to be a key element for all countries in terms of foreseeing all aspects of their future as nations. Despite the fact that it is statistically recommended that population counting should be held every ten years, this does not prevent it as it varies from one country to another. For example, it is mandated by Article I, Section 2 of the US Constitution to count the population every 10 years[3], whereas it resumes to the good faith of those in power from the Haitian government. For instance, from 1950 to now, the country has held only three censuses (1950,1982, and 2003). This minimum of twenty years gap between those censuses shows a big concern in terms of adequate available data on the population . It is accordingly this chapter will be divided into these following in order to confront the reality of Haitian data management to the most common accepted paradigm on data management and age voting population:

2.1 Data management.

2.2 Population voting and registration.

2.3 Difference equation

[2] retrieved from https://www.prb.org/resources/milestones-and-moments-in-global-census-history

[3] retrieved from https://www.census.gov/programs-surveys/censuses.html

2.4 Path analysis.

2.5 Population modeling.

2.1 Data Management

There is no clear statistical definition of the notion "Data management", but for this study it refers to the common mechanism used by most countries in terms of data collection via census. Even though the definition provided by " Oracle", which stated that : " Data management is the practice of collecting, keeping, and using data securely, efficiently, and cost-effectively for better decision[4] appears to be aligned to our definition , but its limitation in terms of data collection makes it secondary comparing to a census; which is normally conducted by government body where they can also provide information on any given aspect of a population. Considering that the study seeks to understand the dynamic of the age voting of the Haitian population, it is necessary as we present the mechanism of data collection by the Haitian government.

But first, let's briefly present how the United States, Jamaica, and the Dominican Republic handle population data. The reason of this choice is to compare Haiti which is approximately the same size with Jamaica and Dominican Republic and non-federate to a well-organized federate country, which is the United States. As for these three countries, all follow the recommendation set by the United Nations on population censuses to be taken at least every ten years[5], where the last census for United States dated in 2020[6], 2011 for Jamaica[7], and 2022 for the Dominican Republic[8]. Despite the U.S Census Bureau providing information on population

[4] retrieved from https://www.oracle.com/database/what-is-data-management/

[5] retrieved from https://en.wikipedia.org/wiki/Census

[6] retrieved from https://www.census.gov/programs-surveys

[7] retrieved from https://www.mof.gov.jm/update-on-the-2022-population-and-housing-census

[8] retrieved from 2022 Dominican Republic Census - Wikipedia

data, it is not the primary source for data on births and deaths. Wich belongs to the National Center for Health Statistics (NCHS) that can provide information on number of births, percent unmarried mothers, number of deaths, infant mortality and more. That means that the available data from U.S. Census Bureau derived from data sharing with NCHS and other state agency dealing with population data including the state department of vital statistics[9]. Like United States, it is required by law for the Dominican Republic census[10] to be conducted every 10 years, where the scenario it also similar as United States in terms of which institution is the primary source for data on births and deaths. Those data are managed by the Central Electoral Board (Spanish: Junta Central Electoral, JCE) of the Dominican Republic , which is a special body of the government of the Dominican Republic. Such institution is also responsible for ensuring the elections as well as the electoral process, and also administer the civil registry , marital status, granting various documents to prove the identification of nationals and citizens, registering, saving, and managing the data of all Dominicans since birth[11] . For the Jamaican country, the Statistical Institute of Jamaica (STATIN), an agency of the Ministry of Finance and Planning, is the one responsible to conduct any census in Jamica[12] . But STATIN is like the US. Census Bureau where it does not have the primary source on data on deaths and births; the Registrar General's Department (RGD) created in 1879 is the only repository in Jamaica for birth, marriage, still birth and death records.

In the case of the Haitian government, there exist four main institutions that involve in data management in occurrence: Civil Registry Office (CRO) from the Justice Department,

[9] retrieved from https://www.census.gov/topics/health/births-deaths.html

[10] retrieved from https://en.wikipedia.org/wiki/2022_Dominican_Republic_Census#cite_note-5

[11] retrieved from Central Electoral Board - Wikipedia

[12] retrieved from Statistical Institute of Jamaica (STATIN) – Jamaica Information Service (jis.gov.jm).

Migration and Emigration (ME) from the Ministry of Transport and Communication, Haitian Institute of Informatics and Statistics (IHSI) from the Ministry of Finance and Planning, National Office of Identification (ONI). From these institutions, the Haitian Institute of Informatics and Statistics (IHSI) is the one that is responsible to provide data on all aspects of the Haitian population. As mentioned above, the Haitian government does not follow the trend of the United Nations in terms of conducting a population census every ten years. Therefore, three censuses by estimation were conducted respectively in 2009, 2011, and 2015. Since those censuses are based on the Haitian census of 2003, we are wondering if the deaths related to the earthquake of 2010 causing more than 300000 has been considered. Because no available data can be retraced from the Civil Registry Office that also does not dispose of any archive on that matter. Like STATIN and the US. Census Bureau , the Haitian Institute of Informatics and Statistics (IHSI) does not have the primary source of data on deaths and births, which are managed by the Civil Registry Office. Thus, the structure of the Haitian data management looks alike at some degree with the one in practice in United States and Jamaica, except there exists no liaison between the Haitian entities involve in data management as it's not the case in United States and Jamaica. In fact, with the creation of the National Office of Identification (ONI) by the government Boniface/Latourtue in 2005, the goal was to count every Haitian citizen from birth to death while the provided identification will serve as both identification and voter registration cards. Ever since the creation of that said institution, the country is still unable to count every Haitian citizen from birth to death, except they still continue to provide the identification card. With the adoption of the ONI's project, which is similar with the one in practice in the Dominican Republic by the Central Electoral Board , the goal was to resolve the problem of data management related to population voting. Since this project is far from

becoming a reality, the data on population voting appears to be a big challenge facing the Haitian government, knowing there exists no liaison between the Haitian institution involving in data management adding to the lack of available data on deaths and migration that are available only from the international institutions involve in data management. In regards of the Migration and Emigration Office, despite the international migratory flow ravaging the country since 2010 they don't have a structure that track this data that exists only from the other international institutions involving in world population.

2.2 Population voting and registration.

Population voting stands before all for countries that require elections to renew the state officials. That means, the criteria for the age limit can vary from one country to another. In order words, population voting is meaningless for countries that do not require elections to renew their state officials. But in most countries[13], 18 is the most common voting age, with the exception of some countries and territories where those turn 16- or 17-year-olds can vote in at least some elections, and full enfranchisement in certain countries like Argentina, Austria, Brazil, Cuba, Ecuador, Nicaragua, East Timor, Greece, and Indonesia for those aged 16 or 17.

In the case of this study, the age voting population is set to 18 which is identical for the three mentioned countries in section 2.1. In terms of voter registration, all of these four countries require a total of three criterion, which are voting age population, voting-age citizen population, and registered population. In the case of United States, since 1972, every state has required that eligible voters be at least 18 years of age. Additionally, in this country, only native or naturalized citizens can legally vote in elections. As for the registration process, with the exception of North

[13] Retrieved from Voting age - Wikipedia

Dakota, every state requires eligible voters to formally register before casting a ballot. In terms of methods and deadlines, the registration procedures vary greatly from state to state[14] . The same criterion also holds for the Jamaican government, where every Jamaican citizen who possesses the following qualifications is entitled to register to vote:

- ✓ Is a Jamaican citizen of eighteen years or over and resident in Jamaica or
- ✓ Is a Commonwealth citizen who is eighteen (18) years of age or older and who is resident in Jamaica at the date of registration and who have been a resident for at least twelve months prior to the date of registration.
- ✓ Is not subjected to any legal incapacity to vote such as being of an unsound mind, convicted or under a suspended sentence[15] .

But the key difference between voter registration in Jamaica and United States is that in Jamica Fingerprints (4 rolled and 10 flat) are necessary because they offer a very good means of making sure that your name appears on the voter's list only once. Which is contrary for the remaining three countries adding to the registration[16] process that is performed continuously twice a year between April 1st and September 30th, and between October 1st and March 31st.

In the case of the Dominican Republic and Haiti, age voting population as well voting age citizenship are similar to the ones of United States and Jamaica, except they differentiate themselves in term of voter registration process. For the Dominican Republic, the Central Electoral Board (Spanish: Junta Central Electoral, JCE) is the special body from the Dominican Republic government that manages all data on Dominican citizens from birth to death. It was

[14] retrieved from Voting and Registration in the Election of November 2006 (apache.org)
[15] retrieved from Registration Procedures - Electoral Commission of Jamaica (ecj.com.jm)
[16] retrieved from FAQs - Electoral Commission of Jamaica (ecj.com.jm).

created on April 12 of 1923 and was incorporated into the Dominican Constitution in 1924. It is also responsible to provide two types of ID, the official ID obtained at age 18 that serve both as identification and voter registration ID, and the minor ID for those national who have reached 16 years for schools and work purposes[17]. As for the Haitian government, voter registration was a separate data until the creation of the National Office of Identification (ONI) in 2005. With the creation of the National Office of Identification (ONI), voter registration is an ongoing process, where data on age voting population is performed on a daily basis. But the actual list of the age voting population depends on the designated date set by the Provisional Electorate Council (CEP) for the current elections. The key difference between the structure used in Haiti and the one from the Dominican Republic it is that the Central Electoral Board from the Dominican Republic contrary to the Provisional Electorate Council (CEP) manages not only data on the elections and the elections process but also managing all Dominican citizens from birth. From this line, data on death does not appear to be an issue in determining the age voting population for the Dominican government as it is not the case for the Provisional Electorate Council (CEP) that gathered its data from the National Office of Identification (ONI).

2.3 Difference equation

In this section, the focus is on the calculation method used by the National Office of Identification (ONI) and the one used in the model by specifying what makes it important to use a difference equation to study the data. But first, what is a difference equation? Difference equation is a mathematical equality involving the differences between successive values of a function of a discrete variable. A discrete variable is one that is defined or of interest only for

[17] retrieved from Central Electoral Board - Wikipedia

values that differ by some finite amount, usually a constant and often 1; for example, the discrete variable x may have the values x0 = a, x1 = a + 1, x2 = a + 2, . . ., xn = a + n. The function y has the corresponding values y0, y1, y2, . . ., yn, from which the differences can be found[18]:

Equation.

$$\begin{aligned} \Delta y_0 &= y_1 - y_0 \\ \Delta y_1 &= y_2 - y_1 \\ &\cdots \\ \Delta y_n &= y_{n+1} - y_n. \end{aligned}$$

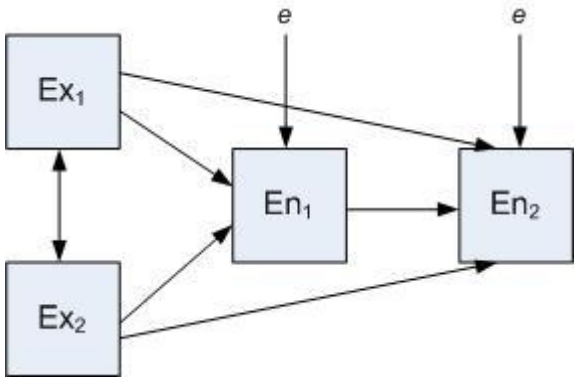
Based on the outcomes from the interviews held with some person resource from the National Office of Identification (ONI), they acknowledged that the calculation method used to generate their data is a simple addition from those who are newly becoming adult adding to those already reached the age voting. With this calculation method, variables as deaths and migration were not considered making that list a special case of the Malthus theorem where the age voting population is constantly growth without being exposed to any type of constraints. Since age voting population is discrete, where the population is exposed to death and migration, the difference equation can provide more detailed insights on the actual age voting population size after being subtracted from death and migration. Furthermore, since all populations are exposed to certain constraints, it is imperative that we consider the effect and/or the impact of those constraints on the population under study. Along the way, contrary to the calculation method used by National Office of Identification (ONI) where the age voting population is a function of those expecting to become adult at time t+1 adding to those aged 18 at time t, the research's model will expand this model by using the logistic regression model where decrease or growth of

[18] retrieved from https://www.britannica.com/science/difference-equation

population comes from the interplay of death and birth , and locally, migration[19] (Marchetti et al. - Human Evolution - 2004).

2.4 Path analysis

The lack of mechanism by the Civil Registry Office to record on a daily basis data on death and migration makes it important to clarify the correlation between the corresponding variables used in our model using a path analysis technique. But, by definition, path analysis is a statistical technique that is used to describe the direct dependencies among a set of variables. But typically, path models consist of independent and dependent variables depicted graphically by boxes or rectangles as we can see in the graph below. Here we must specify that independent variables are labeled as " exogenous" while dependent variables are called " endogenous". For example, in the graph below, we have two exogenous variables (Ex_1 and Ex_2) and two endogenous variables(En1 and En2)[20].



[19] Retrieved from Human population dynamics revisited with the logistic model: How much can be modeled and predicted? - ScienceDirect

[20] Retrieved from https://en.wikipedia.org/wiki/Path_analysis_(statistics)

With such a procedure, the researcher will be able to draw a diagram that serves as visual representation of the relationship between the variables used in the model. For instance, since the calculation method used by the National Office of Identification (ONI) did not consider the variables death and migration, it appears very crucial to use a statistical technique (path analysis) that will serve to validate the model. By using this statistical technique, the researcher we will be able to estimate both the magnitude and significance of the causal connections between the variables especially the new added variables ones, that is, death and migration[21]. Additionally, since Path analysis is closely related to multiple regression[22] ; one could have used it as a special case of path analysis for the purpose of model outcomes comparison between path analysis and multiple regression.

2.5 Population modeling

Mathematical modeling refers to the act of translating your thinking in a measurable scheme with the finality of defining a general rule aims to address a phenomenon. It can be viewed as the activity involved in finding a solution to a real-life problem by working with a mathematical structure that captures the important characteristics of the situation Hirst (2015). If for some, it is a way of describing your belief about how the world functions in real word practice (Marion and Lawson, 2008; Dym 2004), for others, it is a simplified picture of the real-world McFarland (1977). In term of population modeling, this is a type of mathematical model that is applied to the study of population dynamics[23]. As for population dynamics, it is referred

[21] retrieved from https://www.thoughtco.com/path-analysis-3026444

[22] Retrieved from http://faculty.cas.usf.edu/mbrannick/regression/Pathan.html

[23] https://en.wikipedia.org/wiki/Population_model

to the analysis of the factors that affect the increase, stability, and decrease of populations over time[24].

In the case of the Haitian electorate list, all depend on the dynamic of the cohort population 18 and plus and those expecting to become potential voters. In this line, one might be interested to study the trend of the population, whether at a demographic or socio-spatial level, and its impact on the electorate list. Such study will not new, except the researcher interest because study on population dynamics is an old discipline Cappuccino and Price (1995). For instance, Thomas Malthus can probably be credited with formulating the first population mathematical model in 1798, Pollard (1973). But critics of his model stated that a population cannot grow exponentially forever, which led to the development of the logistic model utilized by Verhulst and Peral, Freedman (1980) and Human population dynamics revisited with the logistic model: How much can be modeled and predicted from Marchetti Cesare et all (1996).

But in case of this study, either the Malthusian or the Verhults models can be considered to validate the calculation method used by the National Office of Identification (ONI). Therefore, without being exhaustive, one could have used data exploration model. That is, since registering for the National identification Card (NIC) is done daily, this implies that card registration observation is a function of time. That means, when the data consist of observations "yi at times ti, i = 1...n, we can estimate the derivative between times ti and ti+1 by di+1 = (yi+1 −yi)/(ti+1 −ti). In that case, by plotting di+1 against (yi+1 +yi)/2, we can investigate the relationship between dy/dt and y" Marion and Lawson, (p.8. 2008). Likewise, a cohort model could also be used. With that model, only persons born within a particular year are included

[24] https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/population-dynamics#

Harding (1993) cited by Li (2013). With such a model, researchers would have the opportunity to predict or calculate a new electorate population based on the previous cohort population Dangelmayr and Michael (2005). In the case of this study, the cohorts of population 5 to 17 and the population 18 and plus from the 2003 census could have been considered. By doing so, one could easily predict the maximum population 18 and plus, starting from 2003. Similarly, an agent-based model could also be used. That is, with the ability of an agent-based demographic model of portraying each individual separately Hazhir and Sterman (2008), by plugging all the agents we are likely to come up with an arithmetic value. From this line, these models appear very compelling since they are a true expression of what is happening in the calculation method used by the National Office of Identification (ONI).

Furthermore, difference equation (referred in section 2.3) is another model that could also be used to study the case of the Haitian electorate list. That is, since the newly becoming potential voters are based on age status, that implies a discrete population generation that is exposed to the constraints of death and migration. Thus, formulating a model that retraces the Haitian electorate list tends to be deterministic. Another approach could have been the population dynamic theories, where the population change is determined ultimately by only four factors: birth, death, immigration, and emigration, as referred above. Since this study is based on the age voting population, the factor birth has been omitted. Which makes it equivalent to a difference equation where the population size depends on the variation of the variables death and migration. Such an approach that is characterized from the hypothesis of the study, where the researcher believes that it was the omission of not reducing on constant basis the expected age voting population from death and migration that prevents the Provisional Electorate Council (CEP) to provide data that objectively reflects the age voting population.

Indeed, the logistic model, the difference equations, and the population dynamics theories are the true expression of the expectation of this study in terms of contradicting the calculation method used by the National Office of Identification (ONI). Even though there are some similarities between the numerical model and the difference equation, but the absence of some pertinent data on death rate and migration from the numerical model makes the difference equation or the population dynamics model more suitable to the goal of the study. Because by applying those theories , it was revealed that the results from the exponential equation proved the calculation method used by National Office of Identification (ONI) when death and migration are not computed, while those from the difference equation when death and migration are computed confirmed the inaccuracy of the provided data used in the general elections of 2015 and 2016.

Summary of chapter II.

This chapter covers the common definitions and theories on the key concepts used in the literature review. It was noticed that for the four countries listed in this chapter, they are all align on the requirements related to age-voting population, age-voting citizenship. But in terms of voter registration, all have shown some particularity. For instance, in the case of United States registration mostly vary from state-to-state operating on a daily basis while US citizens living overseas are able to vote as it is the case for the Jamaicans and the Dominican citizens with the exception of the Haitian citizens. Additionally, it was also revealed that in Dominican Republic the same institution that is managing the electoral process and voter registration is also responsible to manage data for all Dominican citizens from death which is not the case for the National office of Identification (ONI) despite is mandated to manage the data for all Haitian citizens from birth to death. In term population modeling, the most important theories were

grouped in two categories: population modeling with death and migration not computed [Population modeling theory of Malthus , Marion and Lawson theory estimating the population at time (t+1) from the variation of the population at time (t+1)/(t-1), cohort population modeling from both Dangelmayr and Michael , and Harding , and agent-based demographic model from Hazhir and Sterman) which could have been used to validate the calculation method utilized by the Office of National Identification (ONI)] , and population modeling with death and migration being computed like logistic model theory employed by Verhulst and Peral, and Human population dynamics revisited with the logistic model by Marchetti Cesare et all, difference equation, and population dynamics theories that show how all populations are exposed to certain constraints (like death and migration based on the study), which was served to contradict the published electorate lists by the Provisional Electorate Council (CEP) for the general elections of 2015 and 2016. Lastly, it was shown the importance of using the path analysis that will serve to describe the corresponding relation existing between each variable. But the next chapter will specify the methodology with regards to the literature review by presenting how it will guide the study until the completion of the last chapter.

Chapter III : Methodology

The data used in this study was taken from official publications. All sources of the cited data are reported as footnotes throughout the text. Before building the model, the researcher first conducted a series of interviews with some officials of the National Identification Office (ONI) to compare the partial list of the pre-selected variables to some pertinent factors that might be ignored during the brainstorm stage. The revelation of these interviews led the researcher delimiting the study by selecting variables age voting population, population 5+ (growth of the population from 2003 census), death, and migration from 2003 to 2016. However, for the completion of the hypotheses, the study aims to use two data set with the same variables, except that those collected from 2003 to 2016 will be used in the difference equation while variables selected from 1990 to 2026 were considered for the sole purpose of generating sufficient data observations for the application of the statistical technique of the path analysis. Thus, to build the mathematical model (the difference equation), a data set was generated based on the official list of the potential voters published by the Provisional Electorate Council (CEP) for the general elections of 2006, 2010 and 2015 with regards of the Haitian census of 2003. Furthermore, to validate the mathematical model, the study made use of two techniques in occurrence: table comparison and path analysis. With tables comparison, the study will check if the data used from the National Office of Identification aligns with other Haitian institution involve in data management while the path analysis will be used to show how each variable is affected by each other during the last 27 years.

3.1 RESEARCH DESIGN AND PROCEDURE.

The research main goal is to predict the " age voting population" at time (t+1) from the expecting age voting population at time (t+1), the age voting population at time (t) , depreciated from death and migration, which makes the study explanatory or predictive. Recall that official published electorate lists were used to build the model; using these data, the researcher generated new variables names from which derived the difference equation used by national Office of Identification (ONI) as well as the difference equation and the statistical technique used to validate the hypotheses. In regards of the variables' names, they are listed in section (3.2), but the difference equations below were used to determine the respective generated age voting population whether being affected by death and/or migration or not being exposed to both of them (see Table-A and Table-AA in Appendix-A for more details).

(1) Predictive Pndm at "t+1": $a_{n+1}=a_n+a_{n-1}$.

It is the population at time "$t+1$" **in function of the population 18+ at time "**t**"** plus the population expecting to become adult (P_{ba}) at $t-1$.

(2) Predictive Pdm a "t+1": $a_n+1= a_n+a_{n-1}-D_r-M_n$.

It is a function of the population at time "t and $t-1$" decreasing with variables death and migration.

Here, equation(1) was used to demonstrate the calculation method used by the National office of Identification ONI (3.1.1.2) while equation(2) was used to validate the model (3.1.2). Thus, since the newly National Identification Card (NIC) registered are becoming potential

voters on a daily basis , equation(1) was also used to generate the yearly potential voters as well as the true voters and the remaining potential voters from 2003 to 2016 (see Table 1 and Table 2 from Appendix-A for more details). Here, true voters are those that had a National Identification Card (NIC) and voted in the last four general elections of (2006, 2010, 2015, and 2016); whereas the remaining potential voters are the expected potential voters, that is, those that are making the remaining population 18+ and who are eligible to subscribe. Mathematically, the remaining potential voters (Pvoters) are a function of true voters (TrueVoters) minus Pop 18+ i.e:

RemainingPvoters = $TrueVoters - Pop\ 18+$ (3), where Pop 18 only indicates the remaining expected voters (see Table1 from Appendix-A for more details).

Therefore, to validate the calculation process that led to the final electorate list used by the Provisional Electorate Council (CEP), the researcher made use of an exponential equation as well as a difference equation with variables death and migration being computed. With the exponential equation (3.1.1.3), users should be able to calculate the population 18+ at any point of time. As for the difference equation, it will allow the users to validate or compare the population 18+ generated by the exponential equation (section 3.1.1.3), at one hand, and provided the different steps for the true potential and expected potential voters in the calculation method used by the National Office of Identification (ONI) on the other hand (section 3.1.1.2).

In regards of the variables' death and migration, it is important to specify that they only exist for the general population. Given the complexity of the aspect of migration, where there is no mechanism to track the mobility of the migrants by the central State and the municipalities, the interest was focused on its external aspect. Excluding the internal aspect as well as migration to the Caribbean region, Europe, and recently to South America specifically to Brazil, Ecuador, Venezuela, and Chile resumed the variable migration used in this study to the Haitian migration

trend to the United States. Such a choice was made based on the proportion of the US migration to the total Haitian migration trend to North America which accounts for more than 89% with a proportion of 61.25% to the American continent, which absorbs more than 91% of the global Haitian migration trend Bidegain (2013). In term of data processing related to the variable migration, as there is no structure allowing Haitian migrants to vote in foreign countries, in this sense, migration and death were considered equally. Thus, deaths as well as migration were applied to the entire population under study (5 and plus), where deaths affecting the population less than 5-year hold were not considered since they are special cases and have their own statistics. Consequently, to avoid the analysis being biased, instead of applying the death rate on the portion of the population expecting to become adult from time (t-1) as well as the population (n) separately, the researcher first added the population (n-1) to the current population (n) at time (t) and then applied the death rate and/or migration to the general population to determine the population at (n + 1).

In all, the overall data used in this study have been processed using mathematical modeling, table comparison between generated variables and status quo data, and path analysis technique. As for the mathematical modeling, the study made use of a difference equation (3.1.1) then scrutinizing the electorate list with regard to a difference equation when death and migration are not computed(3.1.1.2); for respectively take a glance on the Exponential equation regarding the population under study (3.1.1.3), describing the difference equation when adding variables death and migration (3.1.2) while concluding the research results with the path analysis covered in section (3.1.3). In regards of data source, data collection technique, issues of reliability and validity, and sampling technique they are described in sections (3.2), (3.3), (3.4), and (3.5), respectively.

3.1.1 DIFERENCE EQUATION

This section described the calculation method used by ONI and the one used in the model by specifying what makes it important to use a difference equation to study the data. But first, clarification was made toward the notion of " Difference Equation". Which is a mathematical equality involving the differences between successive values of a function of a discrete variable, that is defined or of interest only for values that differ by some finite amount, usually a constant and often 1; for example, the discrete variable x may have the values x0 = a, x1 = a + 1, x2 = a + 2, . . ., xn = a + n. The function y has the corresponding values y0, y1, y2, . . ., yn, from which the differences can be found[25]:

Equation.

$$\begin{aligned} \Delta y_0 &= y_1 - y_0 \\ \Delta y_1 &= y_2 - y_1 \\ &\cdots \\ \Delta y_n &= y_{n+1} - y_n. \end{aligned}$$

Based on the revelation deriving from the outcomes of the interview with some officials of the National Office of Identification (ONI) acknowledging the non-scrutinization of their data from death and migration makes their calculation method a special case of the Malthus theorem where the age voting population is constantly growth without being exposed to any type of constraints. Since age voting population is discrete, where the population is exposed to death and migration, the difference equation can help researchers establishing the actual age voting population size after being subtracted from death and migration. Along the way, contrary to the calculation method used by the National Office of Identification(ONI) where the age voting

[25] retrieved from https://www.britannica.com/science/difference-equation

population is a function of those expecting to become adult at time t+1 adding to those aged 18 at time (t), this study will expand this calculation method by using the logistic regression model where the decrease or the growth of the population comes from the interplay of death and birth , and locally, migration[26] (Marchetti et al. - Human Evolution - 2004).

3.1.1.2 SCRUTINIZING THE ELECTORATE LIST WITH REGARD TO A DIFFERENCE EQUATION WHEN DEATH AND MIGRATION ARE NOT COMPUTED.

Here the goal was to demonstrate the calculation method used by the National Office of Identification (ONI). To demonstrate the calculation method used by the National Office of Identification (ONI), we used two difference equations, where the first one provides data on the yearly potential voters while the second one describes the true and the remaining potential voter. Here, equation(1) and equation(3) from (3.1) will respectively be used to the first and the second equation. Recall that the electorate list was an accumulation of all Haitians adults started from 2003 without being subtracted from death and migration. From that regard, from 2003 to 2016 the growth of the Haitian electorate population was solely based on the cohort of the population aged 5-17 in 2003, which was a case of an exponential growth were death and migration were equal zero.

Thus, to generate the yearly potential voters, the researcher used a difference equation where potential voters at time (t+1) were used as a function of those expecting to become adult (Pba) , which were respectively substituted by a_n (*Pvoters*) and a_{n-1}(*Pba*). As for the predictive population 18+ at $t+1$, it was also labeled as $a_n{}_{+1}$.

[26] Retrieved from Human population dynamics revisited with the logistic model: How much can be modeled and predicted? - ScienceDirect

Mathematically it can be displayed is this following: $a_{n+1}=a_n+a_{n-1}$ (1), where a_{n+1} is equal to the Predictive "Pndm" mentioned above, that is, the age voting population not affected by death and migration. Thus, resolving equation (1) will follow these steps:

$a_{n+1}=a_n+a_{n-1}$ (1); Where

$a_{n+1:}$ is the predictive population 18+, i.e, the population 18+ at time "T+1".

a_n : it is the population 18+ at time "T".

a_{n-1} : is the population of those expecting to become adult at time t+1 (Pba).

Since the newly becoming adults will increase at a_{n+1}, in this sense, a_{n+1} can be expressed in function of a_n, so we can have a difference equation with one inconnu. Thus, equation (1) can be rewritten as $a_{n+1} = a_n + x\ \%a_n$ (2), where $x\ \%$ is the yearly growth rate of the population under study listed in Table-AA from Appendix-A. To compute equation (2) we used the *Pba/Pn_Pndm* ratio from 2003- 2016(see Table-AA from Appendix-A), which makes our $x\%=3.6\%$. With $a_{n+1}= a_n+ x\%a_n$ and $x\%a_n=3.6\%$ which implies that $an+1= 1.036an$ (3). The results listed in Table1(see Appendix-A) can provide the same results as Table-G in terms of population 18+ (see Table-G from Appendix-A for more details). Here, equation (2) above is similar to an exponential equation, where the population is exempt from any time of constraints. As a result, the provided data from equation(2) will be a replica of those generated by an exponential equation, which will be discussed in section 3.1.1.3 below.

In terms of the true and the reaming potential voters, they are all derived from equation (3). That is, the true voters are those that were ability to register and received their National Identification Card (NIC) while the true reaming potential voters from 2003 to 2016 while the remaining potential voters is the addition of those expecting to become adult (Pba) from 2004 to

2016 to the remaining difference between the age voting population in 2005 (see Table 2 from Appendix-A).

3.1.1.3 A GLANCE ON THE EXPONENTIAL EQUATION REGARDING THE POPULATION UNDER STUDY.

It was demonstrated in section 3.1.1.2 above how the National Office of Identification (ONI) had generated their data using a difference equation without death and migration being computed. In this section, the goal is to generate the same data using an exponential equation. By doing so, readers should be able to verify the potential voter for any point of time. Thus, in the case of the Haitian electorate list, from 2003 to 2016, the population 18+ were only affected by the cohort of population aged 5-17 in 2003. From this regard, without the net-migration dN/dt; **dN/dt = B(birth)-D(death) (1)** ; The change in population size(dN) over the interval of time (dt) can be rewritten in this fallowing: **dN/dt = bN – dN or dN/dt = (b – d)N (2)**, where b is the instantaneous (per capita) birth rate and d the instantaneous (per capita) death rate. By replacing (b-d) = r which is the instantaneous (per capita) rate of increase, equation (2) will become : **dN/dt = rN (3).**

Since deaths were not considered in the calculation method used by the National Office of Identification (ONI), the rate of increase r was equal to b, which is identified in the study by the ratio **Pba/Pn-Pndm** (see Table-AA from Appendix-A). In other words, the electorate population was exponentially increased by this following exponential equation below where the observations are listed in Table-G from Appendix-A :

Exponential Equation:

$\mathbf{N_t = N_o e^{rt}}$, where: **Nt** = the number of individuals in the population after t units of time.

No = the initial population size (t = 0).

r = the exponential growth rate.

t = time unit (usually in years).

e = the base of the natural logarithms (2.72)

3.1.2 DESCRIBING THE DIFFERENCE EQUATION WHEN ADDING VARIABLES DEATH AND MIGRATION.

This section is quite similar to 3.1.1.2, except that variable death and migration were computed. By adding the variables death and migration to the equation used by National Office of Identification (ONI) from section 3.1.1.2 above, the researcher was able to generate the age voting population affected simultaneously by death or migration or by both death and migration at the same time from 2003 to 2016. Thus, contrary to the calculation method used by the National Office of Identification (ONI) , the population aged 5-17 in 2003 was considered to be exposed to the constraints of death and migration. Thus, to generate the yearly potential voters affected by death and migration, we used equation (2) from section 3.1 above, that is:

(3) Predictive Pdm a "t+1": $a_n+1 = a_n+a_{n-1} - D_r - M_n$.

Thus, resolving equation (2) will follow these steps:

(1) we expressed a_{n+1} in function of a_n and a_{n-1}, i.e.

$a_{n+1} = a_n + a_{n-1}$, where

a_{n+1}: is the age voting population at time (t+1)

a_{n-1} : is the population at time (t-1) expected to become eligible to register at time (t).

(2) We assumed that that those expecting to become 18 and plus (a_{n-1}) at time t and those already being eligible to register (a_n) at time t can be expressed as $x\%$ percentage of a_n , where $x\%$ was the yearly growth rate of the population under study. Which makes it take this mathematical form:

$a_{n+1} = a_n + x\,\%a_n$, where *x% is* the ratio of Pba/Pn_Pndm.

(3) After added the age voting population expecting to become eligible to register at time t to those already eligible to register at time t ($a_n + a_{n-1}$) , we then depreciated the population at time (t+1) simultaneously from death and migration. Which mathematically takes this following form:

$a_{n+1} = a_n + x\,\%a_n - D_n - M_n$, where

D_n is the death rate at time t , and

M_n is the migration rate at time t.

Here, we could also generate a third equation from equation (2) by subtracting ($a_n + x\,\%a_n$) either from death or migration. Which could have taken this following mathematical form:

(3)Predictive Pd and Pm: $a_{n+1} = a_n + a_{n-1} - D_r \text{ or } M_n$. Their calculations will be similar with equation (2), except they can be considered as special case of equation (2). The results are listed in Table-a and Table-AA (see Appendix-A). However, if we want to know the age voting population affected with death and migration for any point of time, we can substitute equation(2) to this following equation:

Difference Equation

$Nt + 1 = Nt + Pba – D + I - E$ (4)

Nt+1: Potential voters at time t+1

Nt: Pop 18+ at time t

Pba: Population expecting to become adult at time t+1, cohort of the population aged 5-17 in 2003.

D: death

I: Immigration

E: Emigration

I-E: Net migration

For instance, if we want to know the age voting population for 2005 and 2015, equation (4) will become:

N(2005)= 4, 908, 600 + 205,706 – (50,509 + 51,757) - (15,280 + 15,629) (3)

N(2005) = **4, 981, 131**.

N(2015) = N(2014) + Pba(2015) - (∑D4---D15) – (∑M4----M15) = **6, 313, 428 .**

The complete results are listed in Table-M1 (see Appendix-A for more details), where the population affected by both the net migration and deaths came from the Haitian census of 2003 with the age voting population and the population five and plus estimated , respectively at 4, 675,769 and 7, 380,336.

In fact, this variation of the size of the population over time is a perfect example of a differential equation. In that case, the conversion of equation (4) has led to the following differential equation.

Differential equation

$\frac{\Delta P}{\Delta t} = KP(t) - M$ (5); K= Pba- D >0 ; M = I -E > 0

The reason why, we have to substitute "Pba" to "Birth" is that the study is based on the proportion of Haitian population that could be voted in the general elections of 2015 and 2016. From that regard, (Pba- D) can be express in terms of the population as the main elements of the increment or the decrement of the population. Thus, resolving equation (5) consist of using cross multiplication and then taking the integral of both sides, i.e:

dP = (KP – M) dt (6),

divide both terms by (KP-M)

$\frac{dP}{(KP-M)} = dt$, $dP = (KP - M)\, dt$

$\int \frac{dP}{(KP-\dot{M})} = \int dt$; $\int \frac{dP}{(KP-\dot{M})} = t + c$

With U = KP-M , dU = KdP, $dP = \frac{dU}{k}$

$\frac{1}{k}\int \frac{dU}{U} = t + c$; with $\int \frac{1}{U} = \ln |U|$

$\frac{1}{k} Ln|u| = t + c$, $L_n|u| = kt + c$

$|U| = e^{\wedge kt + c}$; $U = + - (e^{kt} . e^{c})$

$U = C\, e^{kt}$; $KP - M = C\, e^{kt}$

KP = C e^{kt} + M; $P = \frac{1}{k}(Ce^{kt}) + M$

Solve for C

$Po = \frac{1}{k}(C + M)$; KPo = C + M

C = KPo – M ; $P = \frac{1}{k}(KP - M)e^{kt} + M$

Calculation for C for year zero:

Year zero: 2003

Pop 18 + : 4, 726, 561

Net migration: 14,259

Mean of (Pba- Death): 0.03

C= 0.03 x 4,726,561 – 14259

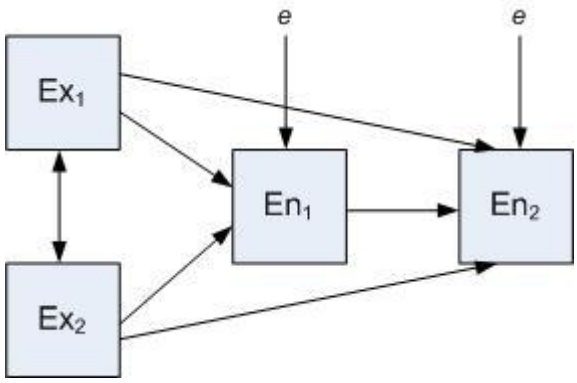
C = 127,538

The data generated from this equation is described in Table-D (see Appendix-A for details).

3.1.3 PATH ANALYSIS

The lack of mechanism by the Civil Registry Office to record on a daily basis data on death and migration makes it important as clarification being made on the nature of the correlation between the corresponding variables used in the model using the path analysis technique. By definition, path analysis is a statistical technique that is used to describe the direct dependencies among a set of variables. But typically, path models consist of independent and

dependent variables depicted graphically by boxes or rectangles as we can see in the graph below, where independent variables are labeled as " exogenous" while dependent variables are called " endogenous". For example, in the graph below, we have two exogenous variables (Ex_1 and Ex_2) and two endogenous variables(En1 and En2)[27].



With such a procedure, users will be able to draw a diagram that serves as visual representation of the relationship between the variables used in the model. For this study, since the calculation method used by the National Office of Identification (ONI) did not consider the variables death and migration, it appears very crucial as we use a statistical technique (path analysis) to validate the model. By using this statistical technique, users will be able to estimate both the magnitude and significance of the causal connections between the variables especially the new added variables ones, that is, death and migration[28]. Additionally, since Path analysis is closely related to multiple regression[29] ; once could have used it as a special case of path analysis for the purpose of model outcomes comparison between path analysis and multiple regression.

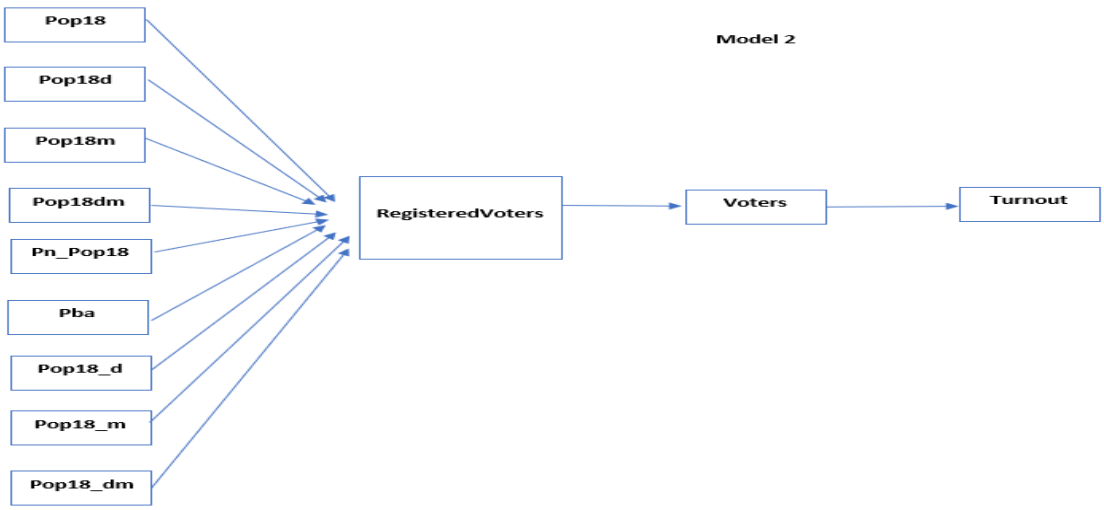
[27] Retrieved from https://en.wikipedia.org/wiki/Path_analysis_(statistics)

[28] retrieved from https://www.thoughtco.com/path-analysis-3026444

[29] Retrieved from http://faculty.cas.usf.edu/mbrannick/regression/Pathan.html

Here, by applying the path analysis, the focus was made on the last 21 elections from 1990 to 2016, which required necessary adjustment to the previous data set for the purpose of having sufficient data observations to conduct that said test. At this level, all the data prior the 2003 census came from the world population prospects from the United Nations from 1990 to 2022, except data on death rate, net migration , registered voters, total votes, population during elections, turnout that came from other websites or institutions[30] involving in world population (see footnotes 30 below for more details). Since the last census prior to the elections of 1990 was held in 1982, we used the provided registered voters of the 1990's elections as reference of the age voting population. Then with the general elections of 1995, the researcher has interpolated the age voting population of 1995 and 1990 to generate the age voting population between 1990 and 1995 likewise it was between the general elections of 1995 and 2000 and between those from 2000 and 2003. As for the other variables, they have followed the same trend, where the procedure was stopped in 2003, year of the last Haitian census that provided the current age voting population. That is, arriving in 2003, the variables were adjusted based on the available data from that said census. Once the data has been set up and manipulated, the researcher came up with the path analysis model, which can be seen in the graph below. With these this model, readers will have the choice to check how the (RegisteredVoters) affected by the age voting population take at each other by the variable death and migration at time (t) , (t+1), and (t-1) with respect to the variables death and migration.

[30] Retrieved from (Demographics of Haiti - Wikipedia , Haiti Population (2022) - Worldometer (worldometers.info); IFES Election Guide | Elections: Haiti Pres Nov 26 2000; Haiti | International IDEA).



3.2 DATA SOURCE AND VARIABLES NAMES

All the data used in this study came from public data. For the age voting population, the data was gathered from the Haitian census of 2003 from the Haitian Institute of Informatics and Statistics (IHSI). The registered voters came from the National Office of Identification (ONI). Additionally, data on death rate, net migration , registered voters, total votes, population during elections, turnout came respectively from world bank via Wikipedia, worldometers.info, IFES , and International IDEA. As for the variables' names used in this study, they are defined in these following:

TableDef: List and Definition of Variables Used.

List and definitions of Variables used	
P_{ba}	Population expecting to become adult. It is the cohort population aged 6 to 17 in 2003. It is the growth factor for the population 18+
P_{ndm}	Population 18+ with death/migration not computed
P_{dm}	Population 18+ with death/migration computed
P_{d}	Population 18+ with death computed
P_{m}	Population18+ with migration computed
P_{n_Pndm}	Population 18+ at time t-1 when computing Pndm
P_{n_Pdm}	Population 18+ at time t-1 when computing Pdm
P_{n_Pd}	Population 18+ at time t-1 when computing Pd
D_R	Death Rate
Migration (M_n)	Haitian US migration to USA. There is no available data for the periods of 2004, 2014, and 2015, so we supposed it could not equal zero. Thus, we attributed the mean value for the existing data for the missing data for 2004, 2014, and 2015
Pvs	Potential voters simulated. Its method of calculation is described in table1 below
Pssv	Potential Simulated subscribed voters. It is the difference between the Pvs at time t and t-1.
Pvsy	Official potential voters subscribed yearly. It was extract from the database of the National Identification Office (ONI)
Pvoters	Potentials voters are the total official voters computed during the last three elections (2006, 2010, and 2015). It is the general list published prior to each election. Since Pvoters is officially published on a quinquennial basis, we simulated its yearly value with their respective Pvsy. Table 1 below provides the respective simulated Pvoters values as well as those used during the last three elections (the bold ones).

List and definitions of Variables used for the Path analysis test	
P_{ba}	Population expecting to become adult. It is the cohort population aged less than 18 in 1989. It is the growth factor for the population 18+
P_{op18}	Population 18+ in 1990 with death/migration not computed
P_{n_Pop18}	Population 18+ in 1990 at time t-1 when computing Pop18
P_{op18dm}	Population 18+ in 1990 with death/migration computed
P_{op18d}	Population 18+ in 1990 with death computed
P_{op18m}	Population18+ in 1990 with migration computed
P_{op18_dm}	Population 18+ in 1990 at time t-1 when computing Pop18dm
P_{op18_d}	Population 18+ in 1990 at time t-1 when computing Pop18d
P_{op18_m}	Population 18+ in 1990 at time t-1 when computing Pop18m
D_R	Death Rate came from United Nations world Population Prospect www. macrotrends.net
(M_n)	Migration also came from United Nations world Population Prospect www. macrotrends.net
RegisteredVoters	The age voting population generated based on the published registered voters from 1990.
Votes	Those who vote in the respective elections. The data came from the Institute for Democracy and Electoral Assistance (IDEA)
Turnout	The percentage of those who vote in the respective elections. The data also came from the Institute for Democracy and Electoral Assistance (IDEA)

3.3 DATA COLLECTION TECHNIQUE.

To gather the data used in the study, those respective websites : Demographics of Haiti - Wikipedia , Haiti Population (2022) - Worldometer (worldometers.info); IFES Election Guide | Elections: Haiti Pres Nov 26, 2000; Haiti | International IDEA were used, except for the general census data as well as voter registration that were obtained directly from Haitian Institute of Informatics and Statistics (IHSI) and the National Office of Identification (ONI). Additionally, since elections are supposed to be held every five (5) years, where voters are registered daily, it was important as new data observations were generated, so it could reflect the twelve or the thirteen years span of the study.

3.4 ISSUES OF RELIABILITY AND VALIDITY

Since the study is dealing with population modeling, reliability and validity will be based on consistency and content. That means, any study on population modeling will end up taking the format used in this study. That is, the population will always be exposed to certain constraints whether death or international/ internal migration. That means, the results depending on the variables used will vary from one researcher to another. That is, some variables may be different from those used in this study, but in the case of the electorate population, the procedure will remain the same where the age voting population will always depreciate from all variables that have a direct impact of the growth of that said population.

3.5 SAMPLING TECHNIQUES.

This study does not intend to use a sampling technique. The difference equation as well as the path analysis will be used on the entire selected data.

Summary of chapter III.

This chapter put emphasis on the research design that retraced two main difference equations that are differentiated by the insertion or not of the variable death and migration. With the use of those difference equations, new variables names were generated based on the official data. Therefore, the difference equation without the variable death and migration was assimilated to an exponential equation where the age voting population was not exposed to death and migration. In the contrary, the difference equation with death and migration being computed was compared to two other equations (a difference equation and a differential equation) from which users can check the age voting population for any point of time. With the newly created variables, the researcher then described which variables will be used in the path analysis by specifying the data source and variable names as well as data collection technique collection. And when it comes to the issues related to reliability and validity, despite some variables may be different from those used in this study, but in the case of the electorate population, the procedure will remain the same where the age voting population will always depreciate from all variables that have a direct impact of the growth of that said population. Furthermore, this study does not intend to use a sampling technique, where the difference equation as well as the path analysis will be used on the entire selected data. Additionally in terms of the generated tables mentioned throughout the chapter, all will be listed in the Appendix-A, with the exception of those that revealed indispensable to insert into the body of the related chapter. In regards of the next

chapter , the goal will consist of describing how the data and the listed tables mentioned in this chapter have been gathered and generated.

CHAPTER IV : Research Results

4.1 : Data for the Haitian Population 5 and plus

According to Haitian Institute of Informatics and Informatics (IHSI), as for the Haitian census of 2003, the population 6 and plus was estimated at 7, 153,940. The cohort of population 5 to 17 representing approximately 36.65% was estimated at 2,704,567 while the population 18 and plus was estimated at 4, 675,769(63.35%). Table-A and Table-AA below show the available data about the Haitian population 5 and plus from 2004 to 2016 with and without being subtracted from death and migration.

Table-A: Simulated Pop 18+ from 2003- 2016 used to analyze the data used in the study.

	Table-A: Simulated Pop 18+ from 2003-2015 used to analyze the data used in the study								
Year	Pba	Pndm	Pdm	Pd	Pm	Drate/1000	Death	Mrate/1000	Mn
2003		4675769				**9.7**		**-3.171**	
2004	182039	4857808	4793136	4808258	4842686	10.2	49550	-3.113	15122
2005	205706	5063514	5001456	5016930	5048040	9.2	46584	-3.056	15474
2006	199871	5263385	5199709	5215488	5247605	**9.1**	47897	**-2.998**	15780
2007	201064	5464449	5399744	5415815	5448378	8.9	48634	-2.941	16071
2008	209639	5674088	5607230	5623589	5657730	8.9	50499	-2.883	16358
2009	250816	5924904	5856246	5873357	5907793	8.7	51547	-2.888	17111
2010	165314	6090218	5964808	5982421	6072605	17.7	107797	-2.892	17613
2011	248422	6338640	6264497	6282860	6320277	**8.8**	55780	**-2.897**	18363
2012	182318	6520958	6446613	6465530	6502041	8.5	55428	-2.901	18917
2013	229661	6750619	6674297	6693914	6731002	8.4	56705	-2.906	19617
2014	203540	6954159	6876564	6897135	6933589	8.2	57024	-2.958	20570
2015	199781	7153940	7073737	7095278	7132399	**8.2**	58662	**-3.011**	21541
2016	226396	7380336	7297211	7319817	7357730	**8.2**	60519	**-3.063**	22606
Total	2704567						746626		

Table-AA: Simulated Pop18+ from 2003- 2016 at time t-1 when computing the variables in Table-A

Table-AA: Simulated Pop 18+ from 2003-2015 at time t-1 when computing the variables in Table-A						
Year	Pba	Pn_Pndm	Pn_Pdm	Pn_Pd	Pn_Pm	Pba/ Pn_Pndm Exponential growth Rate
2003		4675769	4675769	4675769	4675769	
2004	182039	4675769	4675769	4675769	4675769	0.04
2005	205706	4857808	4793136	4808258	4842686	0.04
2006	199871	5063514	5001456	5016930	5048040	0.04
2007	201064	5263385	5199709	5215488	5247605	0.04
2008	209639	5464449	5399744	5415815	5448378	0.04
2009	250816	5674088	5607230	5623589	5657730	0.04
2010	165314	5924904	5856246	5873357	5907793	0.03
2011	248422	6090218	5964808	5982421	6072605	0.04
2012	182318	6338640	6264497	6282860	6320277	0.03
2013	229661	6520958	6446613	6465530	6502041	0.04
2014	203540	6750619	6674297	6693914	6731002	0.03
2015	199781	6954159	6876564	6897135	6933589	0.03
2016	226396	7153940	7073737	7095278	7132399	0.03
Total	2704567					Mean = 0.036
The values of all the variables used in Table-AA derived from Table-A						

Here, it should be noted that the values from Table-AA derived from Table-A, where all the variables, except P_{ba}, D_r, and M_n were computed based on a difference equation.

4.2 Data generated from the official data.

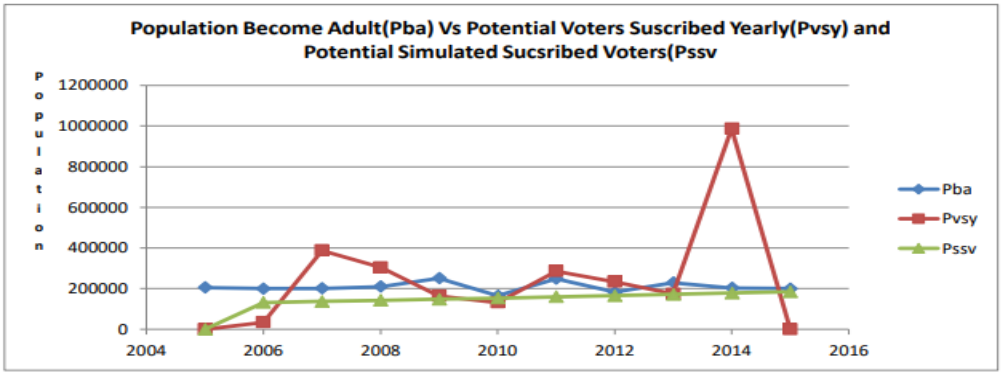
The list of the generated tables can be viewed in Appendix-A. Such a list included Table-A, Table-AA , Table-G, Table-M1 and Table-D. The calculation method used to generate those tables are described in their respective sections (Table-A, Table-AA, Table1, Table2, from section 3.1.1.2; Table-G, Table-M1, and Table-D are described respectively in section 3.1.1.3 and 3.1.2). Recall that section 3.1.1.2 which put emphasis on the difference equation when death and migration are not computed, will be used in section 4.2.3 to prove the calculation method used

by the National Office of Identification (ONI) while section 3.1.1.3 and 3.1.2, which are the contrary will be used in sections 4.2.2.1 to check how the model fit the official data while sections 4.2.1, 4.2.3.1, and 4.2.4 will be used for data visualization, test results using comparison method, and test results using path analysis for hypothesis testing , respectively.

4.2.1 Research results using data visualization.

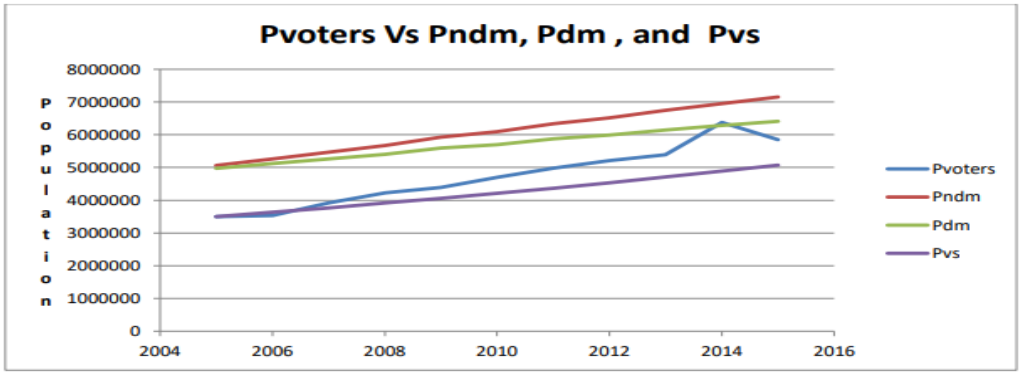
Graph 1- Figure 9: Comparison between the population expecting to becoming adult to the official yearly subscribed and the generated potential subscribed voters.

Figure 9: Comparison between the population expecting becoming adult to the official yearly subscribed and the simulated potential subscribed voters



Graph2- Figure 10: Comparison between the official potential voters (Pvoters) and the potential voters simulated (Psv) to the pop 18+ with death/migration not computed (Pndm) to the population 18+ imputed from death/migration (Pdm).

Figure 10: Comparison between the official potential voters (Pvoters) and the potential voters simulated (Pvs) to the population 18 and plus with death/migration included (Pndm) to the population 18 and plus imputed from death/migration (Pdm)

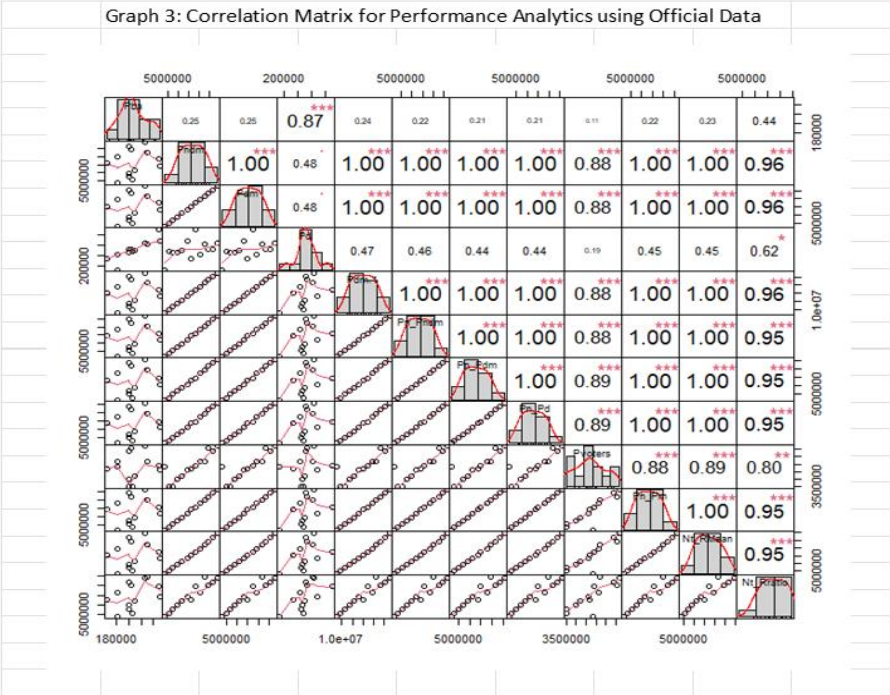


Using data visualization, it revealed that that most of the SAS outputs graphs are close to normality (see Appendix-B for more details), except the Pvsy that shows some serious data anomaly. When comparing that same variable to Pssv and Pba using excel chart (See Graph 1: Figure 9 above), the results show an equal discrepancy on both variables Pssv and Pba from 2013 to 2015. Furthermore, when comparing the official potential voters (Pvoters) to the potential voters simulated (Pvs) with Pndm and Pdm (Graph 2: Figure 10 above), the graphs show some affinity between Pvoters/Pvs and Pndm/Pdm. For instance, from 2005 to 2007, Pndm and Pdm were overlaid until they distance themselves in 2009. Such gap follows the same trend from 2009 to 2011 and being exacerbated the following years. Likewise, were Pvoters and Pvs, except for their exaggeration period that started two years later than Pndm/Pdm one.

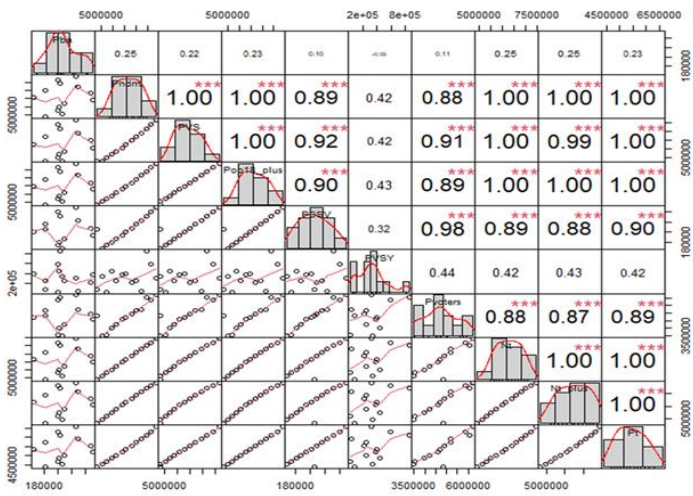
Additionally, during the time frame 2014 - 2015, Pdm and Pvoters intersected, which translated that the potential voters follow the trend of the Pdm. With respect for the observed outlier in 2014 (Graph 1: Figure 9 above), cutting it up won't do too much since Pvoters would remain closer to Pdm than Pndm. It's quite the same scenario if we switch from Pvoters to Pvs . Then when testing these values using a proc capability test, the mean of the simulated potential 18 and plus (Pdm) was 5,631,478 while it was 6,004,724 for the population 18 and plus used by the National Office of Identification (ONI). As for the simulated potential voters (Pvs) and the potential voters used by the Provisional Electorate Council (CEP), their means were 4,242,497 and 4,734,154, respectively. Finally, the means for the simulated potential voters subscribed yearly (Pssv) and the potential voters yearly subscribed used by the Provisional Electorate Council (CEP) were respectively 157,229.7 ≈ 157, 230 and 278,791.5 ≈ 278, 792.

To assure how the model fits the real data , a normality test was performed using the " Performance Analytics" from R which will also state the type of correlation between each variable. The outcomes of these tests are described in the graphs below, where the first one consists of analyzing the official data with respect to the registered voters while the second one refers to the same test , except with the variables used that came from the generated data (model).

Graph 3&4: Correlation Matrix for Performance Analytics using both Official and Generated Data.



Graph 4: Correlation Matrix for Performance Analytics using generated Data from the Model



The distribution of each variable is shown diagonally, where on the top and on the bottom are respectively the value of the correlation and their level of significance marked by a star and a bivariate scatter plot with a fitted line[31]. Here, the goal is not to analysis the bivariate scatter plot, but instead described the type of correlation for model validation. Along the way, the performance on both graphs looks similar in terms of variables normality and the tendency of the correlation for both the official and the generated data used to build the model. Thus, from both graph the age voting population is positively correlated to the population at time (t), (t-1), and (t+1) whether being subtracted from death and migration or not, where all reveal statistical significance at alpha < 0.001. Which remains to be confirmed or confronted when using the path analysis in section (4.2.4). In all, data normality and correlation between official and generated data are quite the same, except that abnormality is only shown for the variable Pvsy (see Graph4), which also confirm from the excel chart from

[31] https://bookdown.org/siju_swamy/Stat_Lab/correlation-and-regression-analysis-in-r.html

Grap 2- Figure10 above. After visualizing, presenting the descriptive statistics, and testing for performance analytics on both the official and the generated data, the next step consisted of testing how the generated data fit the official one, which is discussed in section (4.2.2.1) below.

4.2.2.1 Research results when plugin the model to the official data.

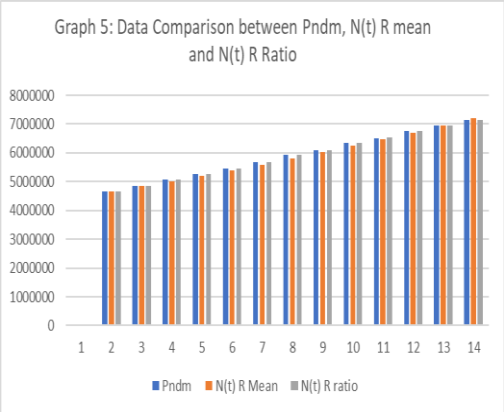
Tablecom: Comparison between official and generated data to build the model.

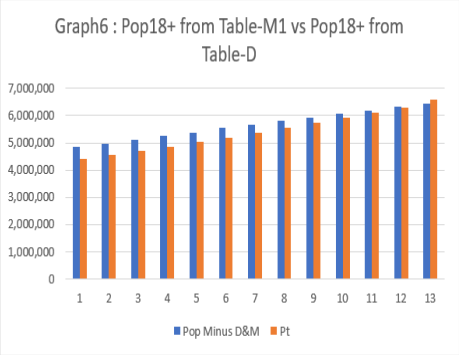
	TableComp: Comparison between official data and data generated to build the model						
Year	Pndm	Pdm	Nt	N(t) R Mean	N(t) R ratio	N(t+1)	P(t)
2003	4675769	4675769	4726561	4675769	4675769	4726561	4726561
2004	4857808	4792166	4908600	4847526	4866591	4842811	4396017
2005	5063514	4996028	5114306	5024840	5065200	4981131	4545059
2006	5263385	5187138	5314177	5209030	5271915	5112194	4698467
2007	5464449	5379097	5515241	5399971	5487066	5243209	4856350
2008	5674088	5591214	5724880	5597912	5710997	5381499	5018833
2009	5924904	5842845	5975696	5803108	5944067	5558885	5186514
2010	6090218	6006871	6141010	6015826	5768394	5649757	5359278
2011	6338640	6252766	6389432	6236341	6439131	5802588	5537513
2012	6520958	6437975	6571750	6464939	6125091	5926155	5721167
2013	6750619	6668792	6801411	6701917	6975428	6075995	5910234
2014	6954159	6876307	7004951	6947582	6503846	6197661	6106021
2015	7153940	7073223	7204732	7202251	6701917	6313428	6307804
2016	7380336	7297211	7431128	7466256	6906021	6429342	6602248
*N(t+1) from Table-M1							
** P(t) from Table-D							
*** N(t) with R ratio from Table-G							
*** Nt Population 18+ when added the data from the Haitian census of 2003 from 2003 to 2016							
**** Pndm and Pdm from Table-A							

TableComp above regroups data from Table-M1, Table-G, and Table-D. from which tables they were shown the results of the calculation method using an exponential equation [N(t)] as well as a difference equation [N(t+1)] and the integral function from Table-D [P(t)]. It is clear that the variation between the official data when the age voting is not subtracted from death and migration perfectly aligns with the generated age voting population from the exponential equation when the age voting is identified by N(t) (see Graph 5 below). Similarly, the two calculation methods used from Table-G and Table-D do not deviate from each other to the point each method will approximately provide the same results when it come to know the age

voting population subtracted from death and migration for any point of time (see Graph 6 below). In all, the graphs below show no discrepancy when plugin the model to the official data. Now, once it was shown that the model fit the data, the next step consists of proving the calculation method used by ONI, which is described in 4.2.3 below.

Graph 5&6: Data Comparison between Pndm, N(t) R Mean and N(t) R Ratio and Pop 18+ minus Death and Migration.





4.2.3 Research results based on the calculation method used by ONI.

Table 2 : Simulated Pop18+ and Potential Voters from 2003-2016

Table2: Simulated Pop 18+ and Potential voters from 2003-2016						
year	Pba	Potential voters	Registered to vote	Remaining Pvoters	Newly added Pvoters	True Voters
		(Pop18+)				Published Registered voters
2003		4675769				
2004	182039	4857808				
2005	205706	5063514	3,500,000*	1,563,514	3,500,000	3,500,000*
2006	199871	5263385	3,533,430**	1,729,955	33430	3,533,430
2007	201064	5464449		1,931,019		
2008	209639	5674088		2,140,658		
2009	250816	5924904		2,391,474		
2010	165314	6090218	4,694,961***	1,395,257	1,161,531	4,694,961
2011	248422	6338640		1,643,679		
2012	182318	6520958		1,825,997		
2013	229661	6750619		2,055,658		
2014	203540	6954159		2,259,198		
2015	199781	7153940	5,871,450****	1,282,490	1,176,489	5,871,450
2016	226396	7380336	6,189,253*****	1,191,083	317803	6,189,253
Total	2704567					
*Published by the European Mission of 2005						
** Published registered to vote by the CEP						
*** Published registered to vote by the CEP						
**** Published registered to vote by the CEP						
***** Published registered to vote by the CEP						
Potential voters sometimes referred to the Pop18+						

Referred to Table2 above, in 2003, the population 18 and plus was estimated at 4,675,769 according to the Haitian Institute of Informatics and Statistics (IHSI), but when adding the population 18 and plus from the given data from the Haitian census of 2003 it shown a total population 18 and plus of 4,726,561. This problem of data management was addressed in this following: the reported population 18 and plus of 4,675,769 was used to prove the calculation method by the National Office of Identification (ONI) while the population 18 and plus of 4,726, 561 was used to prove the difference equation or the model. That is, in this section, the population 18 and plus estimated at 4,675,769 will be used while the one (4,726,561) is mentioned or used from (TableComp) in section 4.2.2.1 above. But it is important to specify that the expected potential voters for both populations share the same calculation method, where the population at time (t+1) remains the addition of those expecting to become adult at time (t+1) to those already reached the age voting, with exception of the age voting population from 2003 that is 4,675,769 (IHSI) and 4,726,561 (generated data based on the given data from the Haitian census of 2003). Along the way, following 2003, the expected potential voters based on the reported age voting from IHSI were computed based on the Pba and the remaining non-subscribed potential voters (Pvns). Mathematically this situation had been expressed in this following:

2003: Population 18+	Epv (2003) =4, 675, 769
2004: Epv (2003) = 4,675,769 + Pba (2004) =182,039	Epv (2004) =4,857,808
2005: Epv (2004) = 4, 857, 808 +Pba (2005) =205,706	Epv (2005) =5,063,514

In 2005, ONI CIN card was launched, where the European Electorate Mission reported a total of 3,500, 000 registered voters. Accordingly, the expected potential voters for the following

year were a derivative of the difference between the population 18 and plus in 2005 minus the registered voters of the past year .

That is, in 2005: The remaining population 18 and plus was 1,563,514 i.e, potential voters of (2005) minus the total subscribed in 2005 (5,063,514- 3, 500,000); so, the expected non-subscribed potential voters for the following years took that pattern:
2006: Pvns(2005)= 1,563,514 + Pba (2005) =199,871 where the Pvns (2006) =1,763,385.

In 2006, there were general elections, where the potential voters were equal to 3,533,430 according to the Provisional Electoral Council (CEP). That means in 2006, the potential voters published by the CEP only exceeds the subscribed potential voters reported by the European Electorate Mission by 33,430. In term of the total of non-subscribed potential voters, it had once again followed the same pattern, where the potential voters non-subscribed (Pvns) were calculated in this following:

2006: Pvns (2006) =1,763,385 – 33,430	Pvns (2006) =1,729,955
2007: Pvns (2006) = 1,729,955 + Pba (2007) =201,064	Pvns (2007) =1,931,019
2008 : Pvns(2007)= 1,9301,019 +Pba(2008)=209,639	Pvns(2008)=2,140,658
2009: Pvns(2008)= 2,140,658 + Pba(2009)=250,816	Pvns(2009)=2,391,474
2010: Pvns(2009)= 2,391,474 + Pba(2010)=165,314	Pvns(2010)=2,556,788

Also, in 2010 there were general elections, where the potential voters were estimated at 4,694,961 according to the CEP. That means, by subtracting Pvoters(2006) from Pvoters(2010), we noticed that from 2006 to 2010, 1,161,531 new voters were added to the previous list (the one published for the 2006 elections). That means in 2010, the total potential voters non-subscribed (Pvns) was 1,395,257 i.e. the Pvns(2010) minus the difference between Pvoters(2010) and the

published registered voters of 2010 i.e (2,556,788-1,161,531). The calculation process for the remaining potential voters will continue by taking the same pattern:

2011: Pvns(2010)= (2,556,788-1,161,531)= 1,395,257 +Pba(2011)=248,422	Pvns (2011)=1,643,679
2012: Pvns(2011)=1,643,679 + Pba (2012)=182,318	Pvns (2012)=1,825,997
2013: Pvns(2012)=1,825,997 + Pba (2013)=229,661	Pvns=2,055,658
2014: Pvns(2013)= 2,055,658 + Pba=203,540	Pvns(2014)=2,259,198
2015: Pvns(2014)=2,259,198 + Pba=199,781	Pvns(2015)=2,458,979

In 2015, new general elections were once again held , where the potential voters were estimated at 5,871,450 according to the CEP. That means, by subtracting Pvoters(2010) from Pvoters(2015), it shown that from 2010 to 2015, a total of 1,176,489 new voters were added to the previous list (the one published for the 2010 elections). That means in 2015, the total potential voters non-subscribed (Pvns) was 1,282, 490 i.e., the Pvns(2015) minus the difference between Pvoters(2015) and the published registered voters of 2015 i.e. (2,458,979-1,176,489).

In 2016, general elections were held for the presidential and part of the senate where the potential voters were estimated at 6,189, 253 which displayed an increment of 317803 new potential voters compared to the published electorate list of 2015. Thus for 2016, the total potential voters non-subscribed (Pvns) was 1, 191,083 i.e., the difference between Pvoters(2016) and the published registered voters of 2016 i.e [(1,282,490-317803) = 964487 + Pba(2016) = 226369 = 1,191,083].

Furthermore, when added the newly becoming potential voters in 2006 to the previous published electorate list from 2005 to 2016, it revealed how the calculation method used by ONI was performed without being exposed to any type of constraints. Which shows that said calculation method was aligned with the exponential growth model when the age voting

population was a simple addition from those expecting to become adult at time (t+1) with those already reached the age voting at time (t).

Despite it was proved from Table2 above that the registered voters were a simple calculation method from Pba at time (t+1) adding to those already reached the age voting at time (t), table1 below shows some anomalies when adding the yearly subscribed potential voters (Pvsy) from 2006 to the previous published electorate list starting from 2005. But when adding the Pvsy from 2006 to 2010 from Pvoters (2006 to 2009) using Table1 below, no anomaly was observed until 2013 where the Pvsy shows an increment that was beyond the mean of the Pvsy. Consequently, contrary from Table2 where the newly added voters aligned perfectly to confirm the respective published electorate list from 2006, 2010, 2015, and 2016, the Pvoters of 2014 (Table1) which is the addition of the Pvsy of 2014 and the previous published electorate of 2013 shows a total of 6,054,233 potential voters which is bigger than the published electorate list of 2015 (5,871,450). Such abnormally that probably the reason why those elections were qualified as fraudulent and were contested to point the presidential elections was cancelled and postponed to November 2016. Since the Pvsy and Pvoters from Table1 show some anomalies from 2013 to 2016, it is required as we further investigate that said anomaly using comparison method. Knowing that the data used in this section came from the Haitian census of 2003, the next step consists of presenting the research results by comparing the official data used by Haitian institutions involving in data management, which is presented in section 4.2.3.1 below.

Table 1: Pvoters and simulated Pvs, Pssv from 2005 - 2016

	Table1: Pvoters and simulated Pvs, Pssv from 2005 to 2015 or from 2005 to 2016					
Calculated Pvs	PVS	*Pop 18+	PSSV	PVSY	**Pvoters	Year
		4,675,769				2003
an+1= an+ x%an and x%an=3.6% 4,675,769		4844097				2004
an+1= 1.036an= 1.036* 4,675,769	4675769	4675769	205999	343454	4829177	2003
	4844097	4844097	205999	85110	4829177	2004
X0 =1.036*4,675,769 = 4844097	4844097	5018484	168328	343454	3500000	2005
X1=1.036X0=1.036(4844097)	5018484	5199150	174387	33430	3533430	2006
X2=1.036X1= 1.036 * (5018484)	5199149	5386319	180665	447242	3980672	2007
X3= 1.036X2= 1.036 * (5199149)	5386319	5580227	187170	248780	4229452	2008
X4= 1.036X3= 1.036* (5386319)	5580226	5781115	193907	387610	4617062	2009
X5=1.036X4= 1.036* (5580226)	5781115	5989235	200889	394293	4694961	2010
X6=1.036X5= 1.036* (5781115)	5989235	6204848	208120	214442	4909403	2011
X7=1.036X6= 1.036*(5989235)	6204847	6428222	215612	255273	5164676	2012
X8=1.036X7= 1.036*(6204847)	6428221	6659638	223374	40852	5205528	2013
X9=1.036X8= 1.036 *(6428221)	6659637	6899385	231416	**987472**	**6193000**	2014
X10=1.036X9= 1.036*(6659637)	6899384	7147763	239747	**589563**	**5,871,450**	2015
X11=1.036X9= 1.036*(6899684)	7147762	7380336	248378	317803	6,189,253	2016
Total						
	* The calculation procedure is identical as Pvs, except the values of Pop 18+ that are different					
	** Pvoters: The bold ones are the official list published by the CEP prior the elections held in					
	the respective year while the remaining values were generated based on Pvsy					
	** *Mean for Pvsy 343454 no available data for year 2003 we used the average					
	**** Mean for PSSV 205999 , no available data for year 2003 and 2004 we used the average					
	**** Mean for Pvoters 4829177, no available data for year 2003 and 2004, we used					

4.2.3.1 Research results using comparison method between generated and official data.

Tables(3,4,5,6,7,&8) : Comparison Tables between Generated and Official Data.

Table 3: Simulated Pndm, Pdm, Pd, Pm

Year	Pndm	Pdm	Pd	Pm
2005	5063514	5001456	5016930	5048040
2006	5263385	5199744	5215488	5247605
2009	5924904	5856246	5873357	5907793
2012	6520958	6446613	6465530	6502041
2013	6750619	6674297	6693914	6731002
2015	7153940	7073737	7095278	7132399
2016	7380336	7297211	7319817	7357730

Table 4: Population 18 + by Institutions Involve in Haitian Data Management

Year	Institution	Pop18+
2003	IHSI	4675769
2005	European Electorate Mission	4500000
2006	CEP/ONI	No available data on Pop 18+, but the published potential voters for those elections were 3,533,430
2009	IHSI	5639026
2010	CEP/ONI	No available data on Pop18+, but the published potential voters for those elections were 4,694,961
2012	IHSI	6032657
2013	UNFPA	6296271
2015	IHSI	6296351
2015A	ONI	6032657
2016	No available data on Pop18+ for 2016, for IHSI, thus we used the generated Pop18+ using the growth (4.18%) from 2012 and 2015	6560045

Table 5: Pvsy Vs Pssv

Year	*Pvsy(ONI)	Pssv
2004	**85,110**	
2005		
2006		132300
2007	387,847	137301
2008	303,818	142491
2009	164,563	147877
2010	131,049	153467
2011	286,211	159268
2012	234,287	165288
2013	174,128	171536
2014	(987,472)	178020
2015		184749
Total	1,767,013	1,572,297

*** Pvsy (ONI): Total subscribed from 2004 to 2013, the value of 2014 is not included in that said total**

Table 6: Pvs Vs Pvoters

Year	Pvs	Pvoters
2005		3500000
2006	3626000	3533430
2010	4177023	4694961
2015	4985005	5871450
2016	5164465	6189253
2015 ONI card delivered		5,411,326

CIN card delivered 87% to the population 18+

***Table 7: Total Delivered CIN Card o the National Population**

FROM 2005 TO FEBRUARY 2015

DEPARTEMENTS	INSCRITS	CIN LIVREES
OUEST	2,361,997	2,243,897
ARTIBONITE	733,878	697,184
NORD OUEST	296,688	281,854
SUD EST	305,741	290,454
SUD	430,184	408,675
GRAND ANSE	229,976	218,477
NORD	551,804	524,214
CENTRE	372,347	353,730
NIPPE	185,513	176,237
NORD EST	219,609	208,629
CARAIBES	8,396	7,976
TOTAL	**5,696,133**	**5,411,326**

*** ONI: Official data on total subscribed and CIN card delivered until February 2015**

Table 8: Total Subscribed & CIN Delivered by Department

FROM 2005 TO 2013

DEPARTEMENT	*POP. 18 ANS +*	*INSCRITS*	*CIN LIVREES*
OUEST	*2,176,273*	*2,290,720*	*2,176,184*
ARTIBONITE	*1,018,067*	*710,382*	*674,863*
NORD-OUEST	*407,824*	*286,616*	*272,285*
SUD-EST	*358,090*	*293,094*	*278,439*
SUD	*450,604*	*408,954*	*388,506*
GRAND-ANSE	*268,172*	*223,492*	*212,317*
NORD	*532,488*	*540,780*	*513,741*
CENTRE	*405,356*	*359,068*	*341,115*
NIPPES	*200,116*	*178,265*	*169,352*
NORD-EST	*215,667*	*209,232*	*198,770*
TOTAL	***6,032,657***	***5,500,603***	***5,225,573***

* ONI: Official data on population 18 and plus, total subscribed and CIN card delivered until 2013

This section is quite a replica of the previous section, except that table's comparison will be used with official data from Haitian institutions involving in data management. Referred to (Table4), in 2003, the population 18 plus was estimated at 4,675,769 according to the Haitian Institute of Informatics and Statistics (IHSI). That means in 13 years, if no constraints the population 18 plus (Pndm) was expected to be no greater than 7,380,336 (Table3). Conversely, that same population was estimated at 6,032,657; 6,032,657; 6,296,351, 6,560,045 by other institutions involved in Haitian data management, respectively for 2012, 2013, 2015, and 2016 (Table4). As for the years of 2015 and 2016, the generated population 18 plus (Pndm) from

(Table3) exceeded the population 18 plus from Table4 to 857589 (7,153,940-6,296,351) and 820,291 (7,380,336- 6,560,045). But it doesn't prevent as the population 18 and plus computed by the Haitian Institute of Informatics and Statistics (IHSI) for those said years represent 88% of the Pndm for both years (2015 , 2016), which makes the validation of the generated population 18 and plus (Pndm) not being too deviate from the official data.

Meanwhile, in 2015, the potential voters were estimated at 5,871,450 (CEP, Table6) while the total delivered National identification Card by the National Office of Identification (ONI) was estimated at 5,411,326 (Table7) for a total subscribed of 5,696,133. Here the published electorate list by the Provisional Electorate Council (CEP) didn't match with the available data from the National Office of Identification (ONI) for the current year, where that said published electorate list exceed the total of delivered National Identification Card (CIN) and the subscribed reported by the National Office of Identification (ONI) by 460,124 and 175,317, respectively. This situation confirmed not only the lack of liaison existing between the Haitian institution involving in data management, but also the challenge facing the Haitian government in terms of data management. Despite the reported data from the Provisional Electorate Council (CEP) and the National Office of Identification (ONI) didn't match, the population 18 and plus used by the National Office of Identification (ONI) from (Table8) was the one published by Haitian Institute of Informatics and Statistics (IHSI) in 2012 (Table4).

Since the published population 18 plus of 2015 (IHSI, 2015) is posterior to the elections of 2015, the population 18 plus of 2012 served as the population of reference. By comparing the population 18 plus (Pndm) from (Table3) to the population 18 plus used by the National Office of Identification (ONI) in (Table8), it revealed a difference of 1,121,283 (7, 153,940-6,032,657) indicating that the population 18 and plus was exposed to certain constraints. That is, the exact

potential voters should somehow be subtracted from at least 1,121,283. By considering the total subscribed of 2015 (Table7) as the population 18 and plus and the total delivered National Identification Card (CIN) as the exact potential voters, it revealed that 94.99% of the subscribers should have received their card. Likewise, when comparing the potential voters (Pvs) to the total delivered National Identification Card (CIN) from (Table6), it shows that the simulated potential voters represent 93.73% of that said population.

Indeed, arriving in 2013 and 2015 , the National Office of Identification (ONI) reported a total of 5,500,603 and 5,696,133 registered voters for a total subscriber from 2003 to 2013 estimated at 1, 767, 013 and 5, 696,133 (Table7), respectively. That means, from 2004 to 2013 the potential voters were estimated at 5,267,013 [registered voters (2005)= 3,500,000 + subscribed voters (2003 – 2013)= 1, 767, 013). But, before the general elections of 2015, the published electorate list by the Provisional Electorate Council (CEP) reported a total of 5, 871, 450 potential voters, which exceeds the reported aged voting subscribers by the National Office of Identification (ONI) from 2015 (Table7) by 175,317 (5,871,450 – 5,696,133). But for the Pvoters being equal 5,871,450 in 2015, the new becoming potential registered voters for 2014 to 2015 should have been no greater than 604, 437 (Registered voters (2015)= 5,871,450 – Registered voters (2013)= 5,267,013). Since the potential voters for 2014 was estimated at 987,472 (Table5), which exceeded by 383,035 the expected potential voters for 2014 to 2015 (604,437), it was what characterizing the data anomaly revealed in section 4.2.2 above (Excel chart Figure 9 from section 4.2.1).

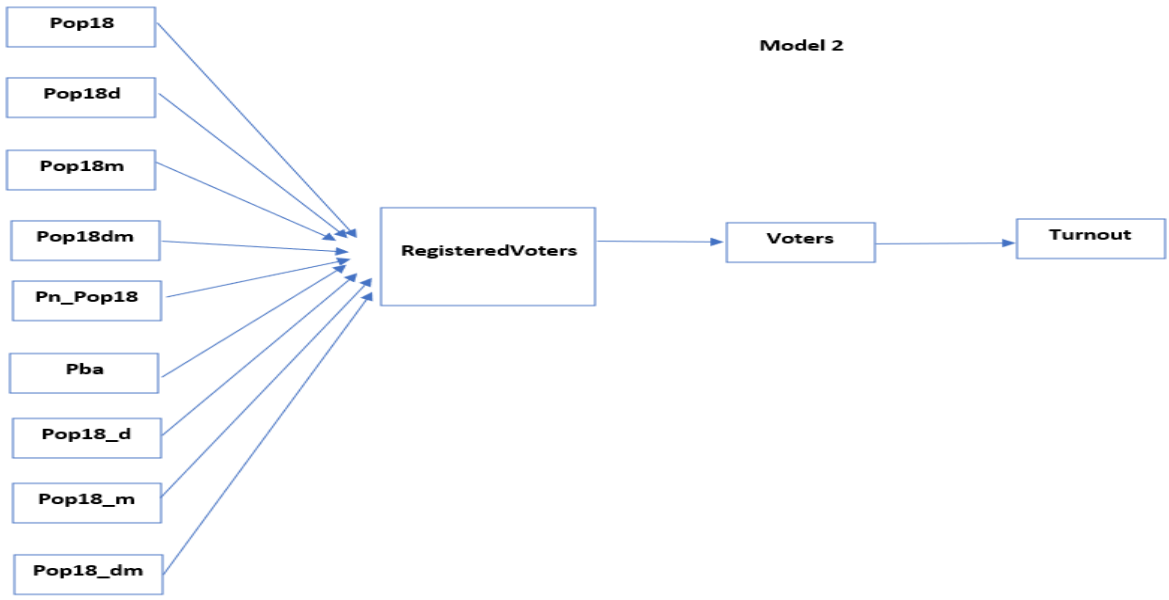
Similarly, the published electorate list by the Provisional Electorate Council (CEP) for the presidential and part of the senate elections of 2016 reported a total of 6,189, 253, that is, an increment of 317803 to the previous electorate list of 2015. That means from 2013 to 2016, for

the registered voters being equal to 6, 189,253, the total subscribers should be at least equal to 922,240, which is still less than the reported subscribed from 2014 (987, 472, Table7). Thus, by adding the registered voters of 2013 (5,267,013) to the expected subscribed voters from 2014 to 2015 (604,437) and those from 2016 (317,803), it equals to the registered voters of 2016 (6,189,253), which confirmed the calculation method from section 4.2.3.

4.2.4 Research results using path analysis.

The path analysis test was conducted using lavaan 0.06.15, where the test ended normally after 33 iterations. The test used Maximum Likelihood(ML) as estimator and Nonlinear Minimization subject to Box Constraints (NLMINB) as optimization method with 59 model parameters and 27 observations. The data used for the test has 27 observations and 12 variables. Which can be displayed in the graph below with nine exogenous variables (Pop18, Pop18d, Pop18m, Pop18dm, Pn_Pop18, Pba, Pop18_d, Pop18_m, and Pop18_dm) and three endogenous variables(RegisteredVoters, Votes, and Turnout).

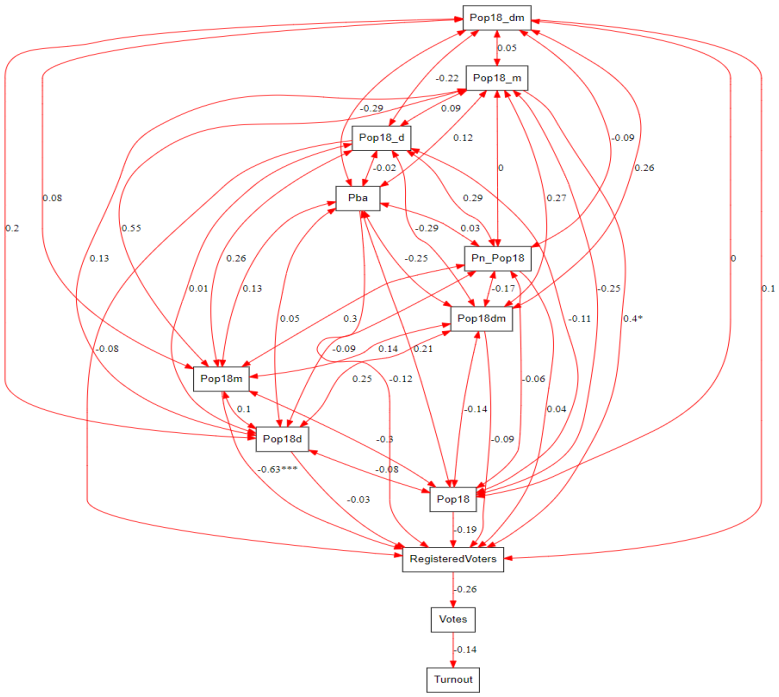
Graph 7: Graph of the Variables Used for the Path Analysis Test.



After running the test, the chi-square test (X2 (19, N = 27) = .407, p > .05) indicates a well specified model or the model fit the data. Which is similar for the comparative fit index (CFI) with (CFI = 0.948 > .90), root mean square error approximation (RMSEA) with (RMSEA = 0.04 < 0.1), Standardized Root Mean Square Residual (SRMR) with (SRMSR = 0.078 < .08) , except the Turkey-Lewis Index (TLI) with (TLI=0.87 < .95) (see the tables in Appendix-Path analysis results for more details) . In all, since only one test shows contradiction (TLI), it can be confirmed that the model is well specified.

Furthermore a " lavaanPlot" was also conducted to oversee how each data is affected or correlated by each other. The graph below summarizes the coefficients, the covariance, and the regression results. Here, the intent is not to go over a complete description of the graph, but to only describe how the variable "RegisteredVoters" is affected by the age voting population taken whether at time (t) , (t+1), and (t-1).

Graph 8 : LavaanPlot



Based on the graph above, Pop18 negatively correlated with RegisteredVoters with a coefficient of (-0.19), similarly for Pop18dm, Pop18d, and Pop18m with a respective coefficient of (-0.09, -0.03, and -0.63), where only the Pop18m is statistically significant at (alpha < 0.001). On the other hand, when those populations are taken one year back, it shows only Pop18_d that kept its negative correlation with "RegisteredVoters" with a coefficient of (-0.08) while Pop18_dm and Pop18_m show a positive correlation with a respective coefficient of (0.1 and 0.4) which is also remain statistically significant, but except this time with (alpha <.05). Therefore, when comparing the path analysis results to the linear model results (see table on linear model results in appendix-path analysis results for more details) , it shows that all the variables are statistically significant, except for Pop18 and Pba. Which indicates that age voting population not being subtracted from death and migration as well as the Pba are meaningless when comparing to the RegisteredVoters.

4.3 Discussion

The mathematical model suggests that the potential voters' lists approached the difference equation. Even though, Pvoters and Pdm were intersected from 2014 to 2015, it is unlikely that the overall potential voters have been registered during that period. As a matter of fact, contrary to what was stated by some National Office Identification (ONI) officials, the general list of the potential voters tends to approach the Pdm instead of Pndm (excel chart figure 10 above). Despite the accumulation of Pvoters of 2005 (3,500,000) and the yearly subscribed from 2006 to 2010 (1,194,961) gave the exact population 18 and plus published by the Provisional Electorate Council (CEP) for the general elections of 2010 when not being reduced from death and migration but arriving in 2014 the population 18 and plus show some deviation that raise concern in term of data management. Because the generated potential voters from 2014 (6,193,000), which is the yearly subscribed from 2014 (987, 472 see Table 5 from section 4.2.3.1) + the generated potential voters from 2013 (5,205,528), is greater than the published electorate list of 2015 (5,871,450) (see Table1 from section 4.2.3). That means by adding the reported yearly subscribed from 2015 (589,563) and 2016 (317,803) to the previous potential voters from 2014 (6,193, 000) , the potential voters for 2015 and 2016 should have been [6,779,563 = Pvoters2014(6,193,000) +Pvsy2015(589,563)] and [6, 961,599 = Pvoters2015(6,779,563) + Pvsy2016(317,803)] , respectively (see Table1 from 4.2.3).

Despite the problem of data management is a big issue for the Haitian government , but the year of 2014 differentiate itself with a high total subscribed that made the potential voters of 2014 (6,193,000) exceeded the published electorate list of 2015(5,871,450) by 321,550. Here, the goal was not to analyze the cause of this excess, but to present its significance on how it

could have affected the elections results displaying very low turnout of 17.82% and 19.20% for 2015 and 2016, respectively.

Since the 5,871,450 and 6,189, 253 represents the exact number published by the CEP for the elections held in August and October of 2015 and November 2016, there is significant evidence that their total list was effectively based on Pndm (see section 4.2.3 and 4.2.3.1). The flip side is that their data did not match those from the National Office of Identification (ONI). Because based on (Table7, section 4.2.3.1), the National Office of Identification (ONI) reported a total of 5,696,133 subscribed or potential voters, which basically contradicts those published by the Provisional Electorate Council (CEP). Even though the data from the National Office of Identification (ONI) contradict those of the Provisional Electorate Council (CEP), this did not resolve the problem of the excess of the potential voters. Moreover, the factor of the National Identification Card (CIN) delivered (Tables7&8, section 4.2.3.1) was not considered as the exact potential voters. This is a situation where the potential voters are resumed to those who subscribed but not to those who have received their National Identification Card (CIN).

This problem of data management has been addressed by the results from the " Performance Analytics and the " Path analysis" from sections (4.2.1 and 4.2.4), respectively. Despite the model perfectly aligns with the official data (section 4.2.2.1), but the results from those two tests show contradiction in terms of statistical significance and correlation type between the dependent variable (Registered voters) and the age voting population taking at time (t), (t-1), and (t+1). If for the performance analytics the registered voters are positively correlated to the age voting population taking at time (t), (t-1), and (t+1) where all are statistically significant at alpha < .001 (see graph from section 4.2.1), the path analysis show otherwise

where the age voting population is negatively correlated at all levels with only the age voting affected with migration that show statistical significance at alpha < 0.001 (see section 4.2.4).

Despite these tests showing different results, which one is more statistically significant can lead to debate. But in this study the " Linear Model" comparison was used as a referee. Thus, based on the results from that said model, all the variables are statistically significant, except for Pop18 and Pba (see table on linear model results in appendix-path analysis results for more details). Additionally, since the results from the path analysis reveal that registered voters are negatively correlated with the age voting population, where one variable increases as the other decreases or vice versa, this implies that the age voting was exposed to certain constraints which are defined in this study whether by death or migration. Also, since the path analysis only shows statistical significance for the age voting population affected by migration, this reveals some concern in term of data management where the official data shows that more people have died than migrated from 2003 to 2016, for a total death and migration of 1,208,000 and 234,969 respectively. This statistic is another concern in term of data management which reveal how the migration aspect has escaped from the control of the Haitian Government.

In all, starting from the exponential model and research comparison tables from sections 4.2.3 and 4.2.3.1, it was confirmed that the age voting deriving from the calculation method used by the Office of National Identification (ONI) was an accumulation between those who are expecting to become adult at time (t+1) and those already reached the age voting at time (t) without being exposed to the constrains of death and migration, at one hand, and the issue of data standardization from those institution involving in Haitian Data management, on the other hand, which is a big issue in term of elections credibility. Thus, this problem of data standardization as it is revealed in the tables from section 4.2.3.1 has been resolved in the three mentioned other

countries from chapter 3 in occurrence: United States, Jamaica, and the Dominican Republic. Because for those countries, the problem of data management has been resolved, whether with liaison between those involves in data management or by institution that was created to manage not only elections but all aspects of the population as it is the case for the Dominican Republic. That is, such study could have been irrelevant or being coped differently to provide significant insight for the age voting population of the Dominican Republic.

Summary of chapter IV.

In summary, the results from the exponential equation proved the calculation method used by National Office of Identification (ONI) when death and migration are not computed, while those from the difference equation when death and migration are computed confirmed the inaccuracy of the provided data used in the general elections of 2015 and 2016. In terms of the path analysis, all the tests results [chi-square test ($X2$ (19, N = 27) = .407, $p > .05$), the Comparative Fit Index (CFI) with ($CFI = 0.948 > .90$), the root mean square error approximation (RMSEA) with ($RMSEA = 0.04 < 0.1$), and the Standardized Root Mean Square Residual (SRMR) with ($SRMSR = 0.078 < .08$)] have indicated a well fitted model , except the Turkey-Lewis Index (TLI) with ($TLI=0.87 < .95$) that shows otherwise.

Additionally, when comparing the results of the performance analytics to those from the path analysis, they showed differentiation in terms of statistical significance and the nature of correlation. If for the performance analytics there exist a positive correlation between registered voters and the age voting population which was statistically significant at alpha < .001, the path

analysis has shown otherwise with only the age voting affected with migration has shown statistical significance at alpha < .001.

Therefore, when comparing the path analysis results to the linear model results , it shows that all the variables are statistically significant, except for the age voting population not being computed from death and migration (Pop18) and the expecting age voting population (Pba). Since the model (the difference equation) , the path analysis, the linear model comparison results, and other available data from other institutions involve in Haitian data management suggest that the registered voters were exposed to certain constraints namely death and migration, we failed to reject the null hypothesis.

Chapter V: Conclusion and Recommendation.

This study put emphasis on the general Haitian potential voters published by the Provisional Electorate Council (CEP) in 2015 and 2016. The study was elaborated from a mathematical perspective using a difference equation. In this research the interest consisted of analyzing the accuracy of the published electorate lists by the CEP prior the elections of 2015 and 2016 under this hypothesis:

" *Since the electorate lists used by the Provisional Electorate Council (CEP) haven't been scrutinized from deaths and migration , I believe that it was the omission of not reducing on constant basis the expected age voting population from death and migration that prevents the Provisional Electorate Council (CEP) to provide data that objectively reflects the age voting Population*". Which was mathematically described in this following: age voting population" at time (t+1) = Expected age voting population at time (t+1) + age voting population at time (t) – Death – Migration against age voting population" at time (t+1) ≠ Expected age voting population at time (t+1) + age voting population at time (t) – Death – Migration

The first finding of this study is that it revealed the lack of common definition on the notion of potential voters, where potential voters were reduced not to those that had received their National Identification Card (CIN) but on the total subscribed. It was the simplistic way of proving the inaccuracy of the turnout participation rate. Because computing the participation rate from the overall subscribed voters instead from those who received their National Identification Card (CIN) and voted made the published turnout participation rate bias.

The other revelation of that study is the excess of 321,550 potential voters for the year of 2014. This excess value can be crucial for those elections, knowing that the president was elected

with only 590, 927 votes. And this is not without reason that those elections were contested, where it is criticized by certain being tinting of fraud. From that prospective, fraud or irregularity are very complex, but a more precise study on the data anomaly observed in 2014 can be done to verify whether those elections were linked to any type of fraud or irregularity.

The other finding of that study is the discrepancy between the population 18 and plus published by the institutions involved in the Haitian Data management and the general potential voters lists published by the Provisional Electorate Council (CEP) for the years of 2015 and 2016. Although the data validation method confirmed the accuracy of that said general list, the available data of the population 18 plus proved that the potential voters were exposed to certain constraints. Because from 2003 to 2016 the population 18 plus was expected to be 7,380,336 when not reducing from death and migration, but in 2012, 2013, 2015, and 2016 the population 18 and plus were estimated at 6,032,657, 6,296,271, 6,296,351, 6,296,271, respectively. That means, the general potential voters list published by the Provisional Electorate Council (CEP) is biased. Additionally, since the considered population 18 plus by the National Office of Identification (ONI) is the one published by the Haitian Institute of Informatics and Statistics (IHSI), this was another way to prove the inaccuracy of the turnout participation rate of those elections as well as the exact potential voters for any given year.

Another interesting finding is the difference between the results of the performance analytics and the path analysis that show contradiction in terms of statistical significance and the nature of correlation. In this sense, if the for the performance analytics there exist a positive correlation between registered voters and the age voting population which was statistically significant at alpha < .001, the path analysis shows otherwise with only the age voting affected with migration has shown statistical significance at alpha < .001. At this level, a linear model

was used as referee to confirm which has provided a better result. Based on the results of the linear model, it follows that the path analysis was more statistically relevant. Because there was statistical significance between the registered voters and all the other variables, except for the population18 not being subtracted from death and migration and those who expecting to become adult at time (t+1).

Furthermore, since the exponential equation proved the accuracy of the calculation used by the National Office of Identification (ONI), where the population wasn't exposed to any type of constraint, this can lead to some temptation of rejecting the null hypothesis. But since the model (the difference equation) , the path analysis, the linear model comparison results, and other available data from other institutions involve in Haitian data management suggest that the registered voters were exposed to certain constraints namely death and migration, we cannot reject the null hypothesis.

Indeed, based on the complexity and the unavailability on some pertinent variables like internal migration, lack of control on the overall Haitian migration by age to name a few, further study needs to be done in order to have a better understanding of the dynamic of the Haitian potential voters. At this level, another approach would consist of using a case study. The aims would consist of randomly selected a polling bureau and investigating whether the listed voters for that said bureau were alive or migrate prior to the elections held in August and October 2015 and November 2016. Doing so, they could investigate the electorate list in its real-life context Yin, (2009). Also, it could serve as an evaluative purpose as clearly stated by Guba and Lincoln cited by Zucker, (2009).

Despite the Haitian government via the Haitian Institute of Informatics and Statistics (IHSI) did ton of works on internal migration, but the incapacity of the State to enforce the

application of the 2005 decree makes it difficult to evaluate its impact on the final electorate lists whether in the city of destination or the city of departure. Another alternative would be to enforce the application of the National Identification Card (NIC) decree. Doing so, the government will be able to track all Haitian citizens from birth to death as clearly stated by one of the articles of the decree leading to the creation of the National Office of Identification (ONI). Considering that the Dominican Republic shares border where the population is quite similar with Haiti, the Haitian Government can mimic the population management put in place in that said country, where there exists only one institution that manages all aspects of the population including the elections. By doing so, the Haitian government will be able to resolve this chronic data management issues characterizing by poor population control and elections irregularities.

References

4EME RECENSEMENT GÉNÉRAL DE LA POPULATION ET DE L'HABITAT. RÉSULTATS PRÉLIMINAIRES by Institut Haïtien de Statistique et d'Informatique (IHSI) et Ministre de l'Economique et des Finances: Paperback (2003) | Libros Latinos. (2003). Www.abebooks.com. https://www.abebooks.com/servlet/BookDetailsPL?bi=15616298278&searchurl=an%3Dinstitut%2Bhaitien%2Bde%2Bstatistique%2Bet%2Bd%2527informatique%2Bihsi%2Bet%2Bministre%2Bde%2Bl%2527economique%2Bet%2Bdes%2Bfinances%26sortby%3D17&cm_sp=snippet-_-srp1-_-title1

2006 Haitian general election. (2023, April 15). Wikipedia. https://en.wikipedia.org/wiki/2006_Haitian_general_election

Aberdeen, T. (2009). Yin, R. K. (2009). Case study research: Design and methods (4th Ed.). Thousand Oaks, CA: Sage. *The Canadian Journal of Action Research*, *14*(1), 69–71. https://doi.org/10.33524/cjar.v14i1.73

André M De Roos. (2014). *Modeling Population Dynamics*. https://staff.fnwi.uva.nl/a.m.deroos/downloads/pdf_readers/syllabus.pdf

Banasiak, J. (n.d.). *Difference and Differential Equations*. Retrieved July 12, 2023, from http://im0.p.lodz.pl/~jbanasiak/Sem/aims.pdf

Batalova, J. (2017, August). *Haitian Immigrants in the United States*. Migrationpolicy.org. https://www.migrationpolicy.org/article/haitian-immigrants-united-states/

Brown, T. A. (2018). Migration and Politics: The Impact of Population Mobility on American Voting Behavior. In *Google Books*. UNC Press Books. https://books.google.com/books?hl=en&lr=&id=autoDwAAQBAJ&oi=fnd&pg=PT11&d

q=population+voting+research&ots=PvDs51nGi1&sig=04sc5PsL_2bLVqcxs8AVk878VhI#v=onepage&q=population%20voting%20research&f=false

Camarota, S. A. (2010, January 25). *Fact Sheet on Haitian Immigrants in the United States*. CIS.org. https://cis.org/Fact-Sheet/Fact-Sheet-Haitian-Immigrants-United-States

Cappuccino, N., & Price, P. W. (1995). Population dynamics: new approaches and synthesis. In N. Cappuccino & P. W. Price (Eds.), *Library Catalog (Blacklight)*. Academic Press. https://searchworks.stanford.edu/view/3078813

Central Electoral Board. (2023, March 28). Wikipedia. https://en.wikipedia.org/wiki/Central_Electoral_Board

difference equation | Britannica. (n.d.). Www.britannica.com. https://www.britannica.com/science/difference-equation

Digital Scholar, M., & Luo, H. (2007). *Population Modeling by Differential Equations*. https://mds.marshall.edu/cgi/viewcontent.cgi?article=1712&context=etd

Dym, C. (2004, June 21). *Principles of Mathematical Modeling - 2nd Edition*. Shop.elsevier.com; Elsevier. https://shop.elsevier.com/books/principles-of-mathematical-modeling/dym/978-0-12-226551-8

Elections in the Dominican Republic. (2021, December 1). Wikipedia. https://en.wikipedia.org/wiki/Elections_in_the_Dominican_Republic#:~:text=Voter%20registration%20in%20the%20Dominican

Elector Registration System (ERS). (n.d.). Electoral Commission of Jamaica. Retrieved July 12, 2023, from https://ecj.com.jm/voters/voting-system/elector-registration-system-ers/

File, T. (2008). *Population Characteristics of VOTING AND REGISTRATION OF THE VOTING-AGE CITIZEN POPULATION*.

https://corpora.tika.apache.org/base/docs/govdocs1/664/664608.pdf

Freedman, H. I. (1980). Deterministic mathematical models in population ecology. In *Library Catalog (Blacklight)*. M. Dekker. https://searchworks.stanford.edu/view/826651

Haiti | International IDEA. (n.d.). Www.idea.int. https://www.idea.int/data-tools/country-view/125/40

Haiti Population 1950-2022. (n.d.). Www.macrotrends.net. https://www.macrotrends.net/countries/HTI/haiti/population

Haitian Canadians. (2023, April 11). Wikipedia. https://en.wikipedia.org/wiki/Haitian_Canadians#Haitian_Migration_to_Canada

Haitian Government: Boniface/ Latortue Administration. (2005, June). *Decree of June 2005 leading to the Creation of the National Identification Office (ONI).* Oni. http://oni.gouv.ht/580-2/

Haitian Institute of Statistics and Informatics (IHSI). (2009). *Census by Estimation of 2009*.

Haitian Institute of Statistics and Informatics (IHSI). (2012). *Census by Estimation of 2009*.

Haitian Institute of Statistics and Informatics (IHSI). (2015). *Census by Estimation of 2009*.

Ham Chande, R. (1976). J. H. Pollard. Mathematical models for the growth of human populations. London: Cambridge University Press, 1973. *Estudios Demográficos Y Urbanos*, *10*(03), 400. https://doi.org/10.24201/edu.v10i03.331

https://www.facebook.com/thoughtcodotcom. (2011). *What is Path Analysis?* ThoughtCo. https://www.thoughtco.com/path-analysis-3026444

IFES Election Guide | Elections: Haiti Pres Nov 26 2000. (2000, November 26). Www.electionguide.org. https://www.electionguide.org/elections/id/1828/

Kirby, M., & Dangelmayr, G. (n.d.). *MATHEMATICAL MODELING A Comprehensive*

Introduction. Retrieved July 12, 2023, from http://people.whitman.edu/~hundledr/courses/M250F03/ch1.pdf

Lee, R. D. (1987). Population Dynamics of Humans and Other Animals. *Demography*, *24*(4), 443. https://doi.org/10.2307/2061385

Li, J., & O'Donoghue, C. (2013, August 31). *A survey of dynamic microsimulation models: Uses, model structure and methodology | International Journal of Microsimulation*. Www.microsimulation.pub. https://www.microsimulation.pub/articles/00082

Marchetti, C., Meyer, P. S., & Ausubel, J. H. (1996). Human population dynamics revisited with the logistic model: How much can be modeled and predicted? *Technological Forecasting and Social Change*, *52*(1), 1–30. https://doi.org/10.1016/0040-1625(96)00001-7

Marion, G., Bioinformatics, & Scotland, S. (2015). *An Introduction to Mathematical Modelling*. https://people.maths.bris.ac.uk/~madjl/course_text.pdf

McClean, S., & Swamurthy, M. (1983). Growth and Structure of Human Population in the Presence of Migration. *Applied Statistics*, *32*(3), 318. https://doi.org/10.2307/2347960

Milestones and Moments in Global Census History. (n.d.). PRB. https://www.prb.org/resources/milestones-and-moments-in-global-census-history/

November 2016 Haitian presidential election. (2023, June 13). Wikipedia. https://en.wikipedia.org/wiki/November_2016_Haitian_presidential_election#cite_note-9

Oracle. (2014). *What is Data Management?* Oracle.com. https://www.oracle.com/database/what-is-data-management/

Path Analysis. (2019). Usf.edu. http://faculty.cas.usf.edu/mbrannick/regression/Pathan.html

Registration Procedures. (n.d.). Electoral Commission of Jamaica. https://ecj.com.jm/voters/voter-information/registration-procedures/

Restrepo, J. M., Rael, R. C., & Hyman, J. M. (2009). Modeling the Influence of Polls on Elections: A Population Dynamics Approach. *Public Choice*, *140*(3/4), 395–420. https://www.jstor.org/stable/40270930

Smith, D. A. (1977). Human Population Growth: Stability or Explosion? *Mathematics Magazine*, *50*(4), 186. https://doi.org/10.2307/2690216

Statistical Institute of Jamaica (STATIN) – Jamaica Information Service. (n.d.). Jis.gov.jm. https://jis.gov.jm/government/agencies/statistical-institute-of-jamaica-statin/

Tremblay, J.-M. (2005, February 2). *Gabriel Bidegain, Les Haïtiens qui sont partis. Combien sont-ils ? Où sont-ils ? Version actualisée, janvier 2013, 31 pp. L'émigration haïtienne en République Dominicaine entre le discours et la réalité*. Texte. http://classiques.uqac.ca/contemporains/Bidegain_Gabriel/Haitiens_qui_sont_partis/Haitiens_qui_sont_partis_texte.html

UNION EUROPEENNE MISSION D'OBSERVATION ELECTORALE HAÏTI Elections législatives. (2006). https://eeas.europa.eu/archives/eueom/pdf/missions/declaration_preliminaire_haiti_21_avril.pdf

United States Census Bureau. (2021, April). *Voting and Registration in the Election of November 2020*. The United States Census Bureau. https://www.census.gov/data/tables/time-series/demo/voting-and-registration/p20-585.html

US Census Bureau. (2019, May 2). *Our Censuses, U.S. Census Bureau Censuses*. The United States Census Bureau. https://www.census.gov/programs-surveys/censuses.html

Voting age. (2023, February 10). Wikipedia. https://en.wikipedia.org/wiki/Voting_age#:~:text=18%20is%20the%20most%20common

Wagner, M., Johann, D., & Kritzinger, S. (2012). Voting at 16: Turnout and the quality of voting choice. *Electoral Studies*, *31*(2), 372–383. https://doi.org/10.1016/j.electstud.2012.01.007

Wikipedia Contributors. (2019a, April 21). *Path analysis (statistics)*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Path_analysis_(statistics)

Wikipedia Contributors. (2019b, May 21). *Census*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Census

Wikipedia Contributors. (2019c, October 9). *Demographics of Haiti*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Demographics_of_Haiti

Wikipedia Contributors. (2019d, October 23). *Voter registration in the United States*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Voter_registration_in_the_United_States

Wikipedia Contributors. (2023, April 17). *2015–16 Haitian parliamentary election*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/2015%E2%80%9316_Haitian_parliamentary_election

Worldometer. (2019). *Haiti Population (2019) - Worldometers*. Worldometers.info. https://www.worldometers.info/world-population/haiti-population/

Yearbook of Immigration Statistics. (2018, September 13). Department of Homeland Security. https://www.dhs.gov/immigration-statistics/yearbook

Zucker, D. (2009). *How to Do Case Study Research*. https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1001&context=nursing_faculty_pubs

Appendix- A : Generated Tables to build the model.

	Table-A: Simulated Pop 18+ from 2003-2015 used to analyze the data used in the study								
Year	Pba	Pndm	Pdm	Pd	Pm	Drate/1000	Death	Mrate/1000	Mn
2003		4675769				**9.7**		**-3.171**	
2004	182039	4857808	4793136	4808258	4842686	10.2	49550	-3.113	15122
2005	205706	5063514	5001456	5016930	5048040	9.2	46584	-3.056	15474
2006	199871	5263385	5199709	5215488	5247605	**9.1**	47897	**-2.998**	15780
2007	201064	5464449	5399744	5415815	5448378	8.9	48634	-2.941	16071
2008	209639	5674088	5607230	5623589	5657730	8.9	50499	-2.883	16358
2009	250816	5924904	5856246	5873357	5907793	8.7	51547	-2.888	17111
2010	165314	6090218	5964808	5982421	6072605	17.7	107797	-2.892	17613
2011	248422	6338640	6264497	6282860	6320277	**8.8**	55780	**-2.897**	18363
2012	182318	6520958	6446613	6465530	6502041	8.5	55428	-2.901	18917
2013	229661	6750619	6674297	6693914	6731002	8.4	56705	-2.906	19617
2014	203540	6954159	6876564	6897135	6933589	8.2	57024	-2.958	20570
2015	199781	7153940	7073737	7095278	7132399	**8.2**	58662	**-3.011**	21541
2016	226396	7380336	7297211	7319817	7357730	**8.2**	60519	**-3.063**	22606
Total	2704567						746626		

Table-AA: Simulated Pop 18+ from 2003-2015 at time t-1 when computing the variables in Table-A						
Year	Pba	Pn_Pndm	Pn_Pdm	Pn_Pd	Pn_Pm	Pba/ Pn_Pndm Exponential growth Rate
2003		4675769	4675769	4675769	4675769	
2004	182039	4675769	4675769	4675769	4675769	0.04
2005	205706	4857808	4793136	4808258	4842686	0.04
2006	199871	5063514	5001456	5016930	5048040	0.04
2007	201064	5263385	5199709	5215488	5247605	0.04
2008	209639	5464449	5399744	5415815	5448378	0.04
2009	250816	5674088	5607230	5623589	5657730	0.04
2010	165314	5924904	5856246	5873357	5907793	0.03
2011	248422	6090218	5964808	5982421	6072605	0.04
2012	182318	6338640	6264497	6282860	6320277	0.03
2013	229661	6520958	6446613	6465530	6502041	0.04
2014	203540	6750619	6674297	6693914	6731002	0.03
2015	199781	6954159	6876564	6897135	6933589	0.03
2016	226396	7153940	7073737	7095278	7132399	0.03
Total	2704567					Mean = 0.036
The values of all the variables used in Table-AA derived from Table-A						

Table-G: Simulated Pop 18+ from 2003-2015 from the exponential equation					
year	Pba	Pndm	time	N(t)=Noe^rt R = 0.036 is the mean of Pba/Pn_Pndm from TableAA	N(t)=Noe^rt R is the ratio of Pba/Pn_Pndm Retrieved From Table AA
2003		4675769	to= 0.036	4675769	
2004	182039	4857808	t1=R*1=0.036	4847526	4866591
2005	205706	5063514	t2=R*2= 0.072	5024840	5065200
2006	199871	5263385	t3= R*3= 0.108	5209030	5271915
2007	201064	5464449	t4=R*4= 0.144	5399971	5487066
2008	209639	5674088	t5= R*5=0.18	5597912	5710997
2009	250816	5924904	t6=R*6=0.216	5803108	5944067
2010	165314	6090218	t7= R*7=0.252	6015826	5768394
2011	248422	6338640	t8= R*8=0.288	6236341	6439131
2012	182318	6520958	t9= R*9= 0.324	6464939	6125091
2013	229661	6750619	t10= R*10= 0.36	6701917	6975428
2014	203540	6954159	t11=R*11=0.396	6947582	6503846
2015	199781	7153940	t12=R*12=0.432	7202251	6701917
2016	226396	7380336	t13=R*13=0.468	7466256	6906021
Total					

Table1: Pvoters and simulated Pvs, Pssv from 2005 to 2015 or from 2005 to 2016							
Calculated Pvs	PVS	Pvs adjusted	*Pop 18+	PSSV	PVSY	**Pvoters	Year
			4,675,769				2003
X%=3.6% with Pop 18+ = 4, 675, 769			4844097		85110		2004
X% =3.6% Pvoters = 3500000	3500000	4844097	5018484			3500000	2005
X1=1.036X0=1.036(3 500 000)	3626000	5018484	5199150	122500	33430	**3,533,430**	2006
X2=1.036X1= 1.036 * (3626000)	3756536	5199149	5386319	126788	447242	3,980,672	2007
X3= 1.036X2= 1.036 * (3756536)	3891771	5386319	5580227	131225	248780	4,229,452	2008
X4= 1.036X3= 1.036* (3891771)	4031875	5580226	5781115	135818	387610	4,617,062	2009
X5=1.036X4= 1.036* (4031875)	4177023	5781115	5989235	140571	394293	**4,694,961**	2010
X6=1.036X5= 1.036* (4177023)	4327395	5989235	6204848	145492	214442	4,909,403	2011
X7=1.036X6= 1.036*(4327395)	4483182	6204847	6428222	150583	255273	5,164,676	2012
X8=1.036X7= 1.036*(4483182)	4644576	6428221	6659638	155855	40852	5,205,528	2013
X9=1.036X8= 1.036 *(4644576)	4811781	6659637	6899385	161309	848705	6,054,233	2014
X10=1.036X9= 1.036*(4811781)	4985005	6899384	7147763	166955	589563	**5,871,450**	2015
X10=1.036X9= 1.036*(4985005)	5164465	7147762	7380336	248378	317803	**6,189,253**	2016
Total				1, 437,096			
	* The calculation procedure is identical as Pvs, except the values of Pop 18+ that are different						
	** Pvoters: The bold ones are the official list published by the CEP prior the elections held in the respective year while the remaining values were simulated based on Pvsy						
	The data shows some anomalies between year 2013, 2014 and 2014						

Table2: Simulated Pop 18+ and Potential voters from 2003-2016						
year	Pba	Potential voters	Registered to vote	Remaining Pvoters	Newly added Pvoters	True Voters
		(Pop18+)				Published Registered voters
2003		4675769				
2004	182039	4857808				
2005	205706	5063514	3,500,000*	1,563,514	3,500,000	3,500,000*
2006	199871	5263385	3,533,430**	1,729,955	33430	3,533,430
2007	201064	5464449		1,931,019		
2008	209639	5674088		2,140,658		
2009	250816	5924904		2,391,474		
2010	165314	6090218	4,694,961***	1,395,257	1,161,531	4,694,961
2011	248422	6338640		1,643,679		
2012	182318	6520958		1,825,997		
2013	229661	6750619		2,055,658		
2014	203540	6954159		2,259,198		
2015	199781	7153940	5,871,450****	1,282,490	1,176,489	5,871,450
2016	226396	7380336	6,189,253*****	1,191,083	317803	6,189,253
Total	2704567					
*Published by the European Mission of 2005						
** Published registered to vote by the CEP						
*** Published registered to vote by the CEP						
**** Published registered to vote by the CEP						
***** Published registered to vote by the CEP						
Potential voters sometimes referred to the Pop18+						

Table-M1: Population 18+ minus death and migration from equation (1)					
Year	Pba	Nt	(∑Dn-1---Dn)	(∑Mn-1-----Mn)	Pop 18+ minus Death & Migration
					Nt + 1 = Nt + Pba - (∑Dn-1---Dn) - (∑Mn-1---Mn)
2003		4726561			
2004	182039	4908600	50509	15280	4,842,811
2005	205706	5114306	102266	30909	4,981,131
2006	199871	5314177	155142	46841	5,112,194
2007	201064	5515241	208971	63061	5,243,209
2008	209639	5724880	263815	79566	5,381,499
2009	250816	5975696	319987	96824	5,558,885
2010	165314	6141010	376669	114584	5,649,757
2011	248422	6389432	434685	133094	5,802,588
2012	182318	6571750	493436	152159	5,926,155
2013	229661	6801411	553492	171924	6,075,995
2014	203540	7004951	614645	192645	6,197,661
2015	199781	7204732	676966	214338	6,313,428
2016	226396	7431128	764966	236820	6,429,342
					6,429,342
Total	2,704,567	7,431,128	764,966	241,338	

Table-D: Pop18 minus death and migration when using integration				
Year	Kgrowth = (Pba-Drate)/Pn_Pndm	Nt		Pop 18+ minus
			Total	Death & Migration
			Net Migration	
				P(t)= ((KPo-M)e^kt + KM)
				K
	Mean = 0.03			P(t) = (127538)e^kt + Km
	C: 127538			K
2003		4726561	Mo= 14259	
2004	0.03	4908600	15280	4396017
2005	0.03	5114306	30909	4545059
2006	0.03	5314177	46841	4698467
2007	0.03	5515241	63061	4856350
2008	0.03	5724880	79566	5018833
2009	0.03	5975696	96824	5186514
2010	0.02	6141010	114584	5359278
2011	0.03	6389432	133094	5537513
2012	0.02	6571750	152159	5721167
2013	0.03	6801411	171924	5910234
2014	0.02	7004951	192645	6106021
2015	0.02	7204732	214338	6307804
2016	0.02	7,431,128	236820	6602248

Appendix- B: SAS outputs for data visualization using generated data.

Figure 7 Distribution of Pvs, SAS output

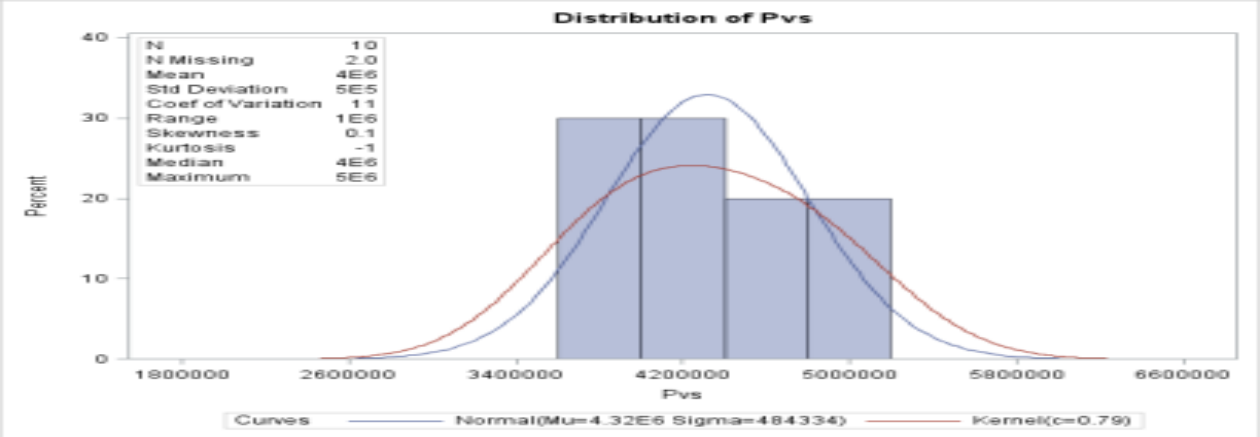


Figure 8 Distribution of Pssv, SAS output

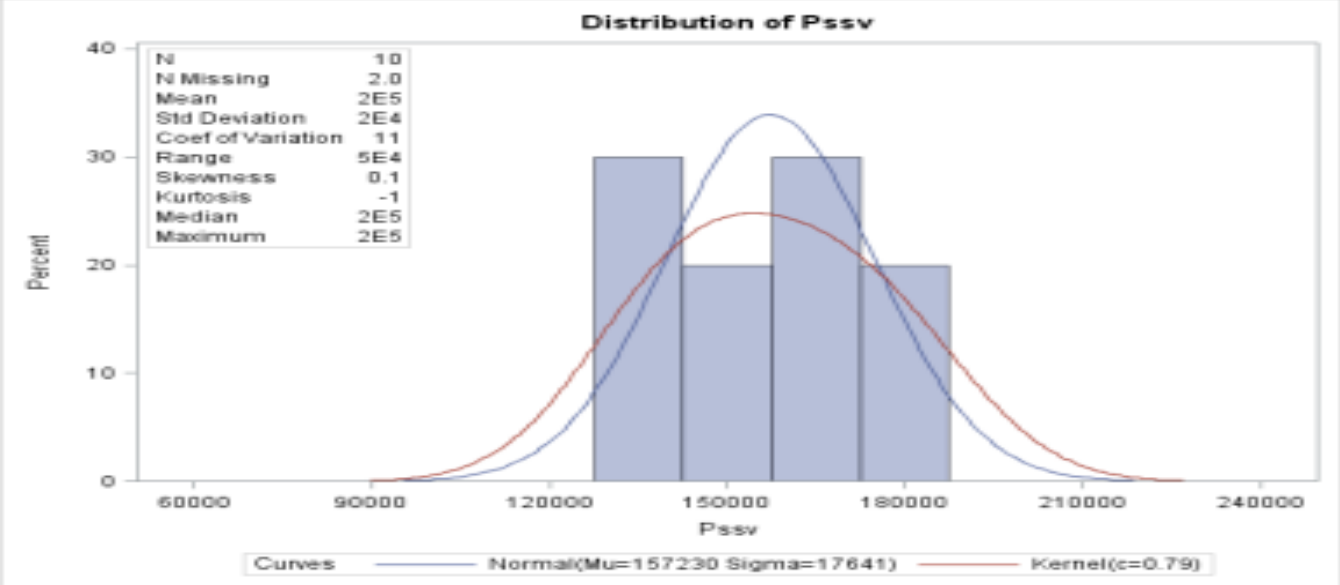


Figure 9 Distribution of Pdm, SAS output

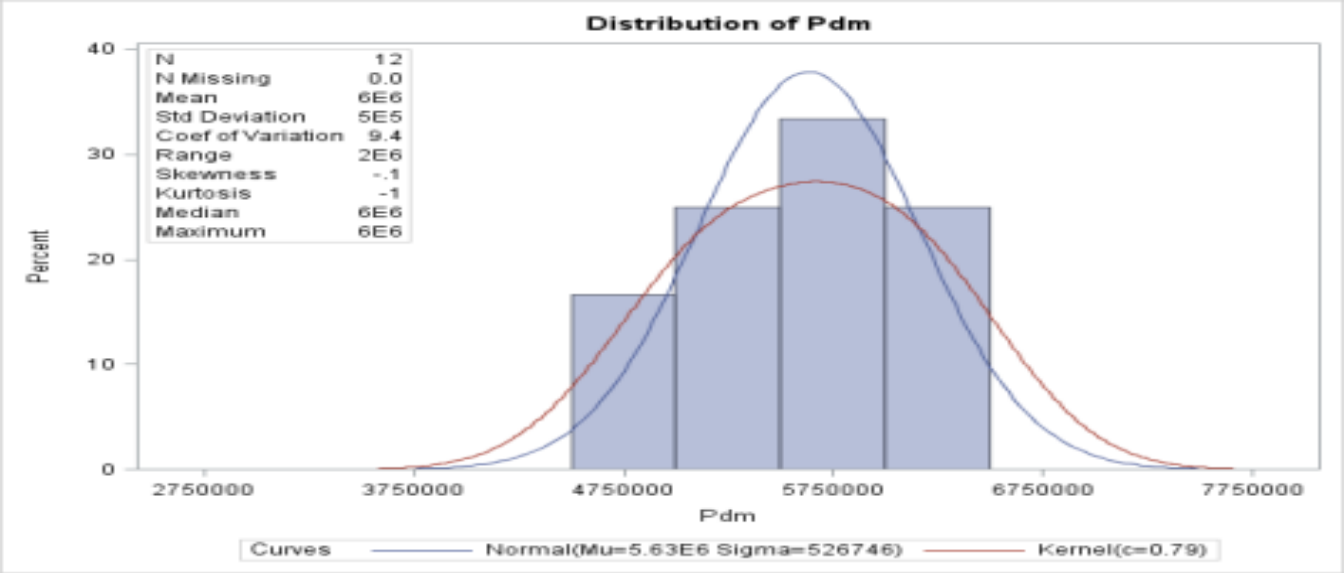


Figure 10 Distribution of Pvsy, SAS output

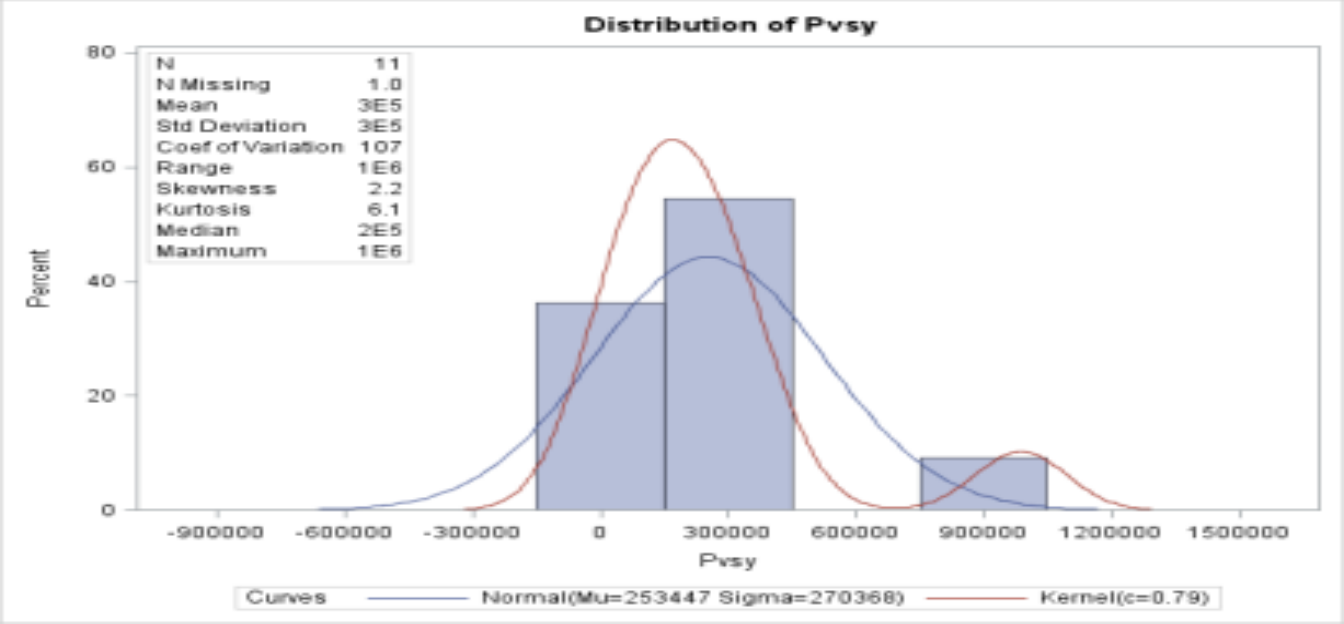


Figure 11 Distribution of Pvoters, SAS, output

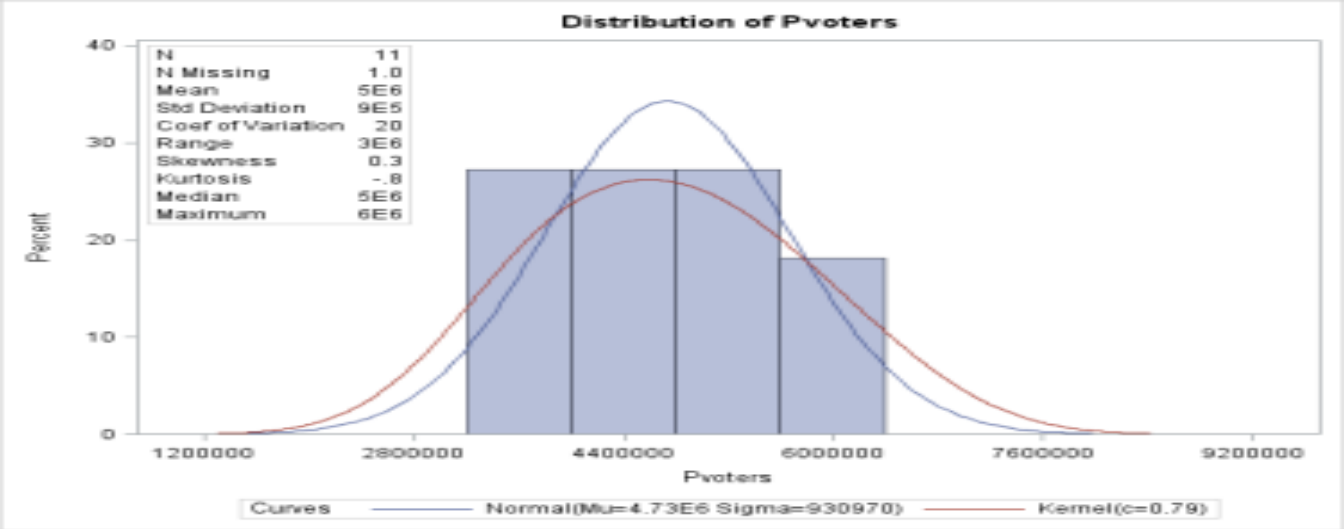


Figure 12 Distribution of Pndm, SAS output

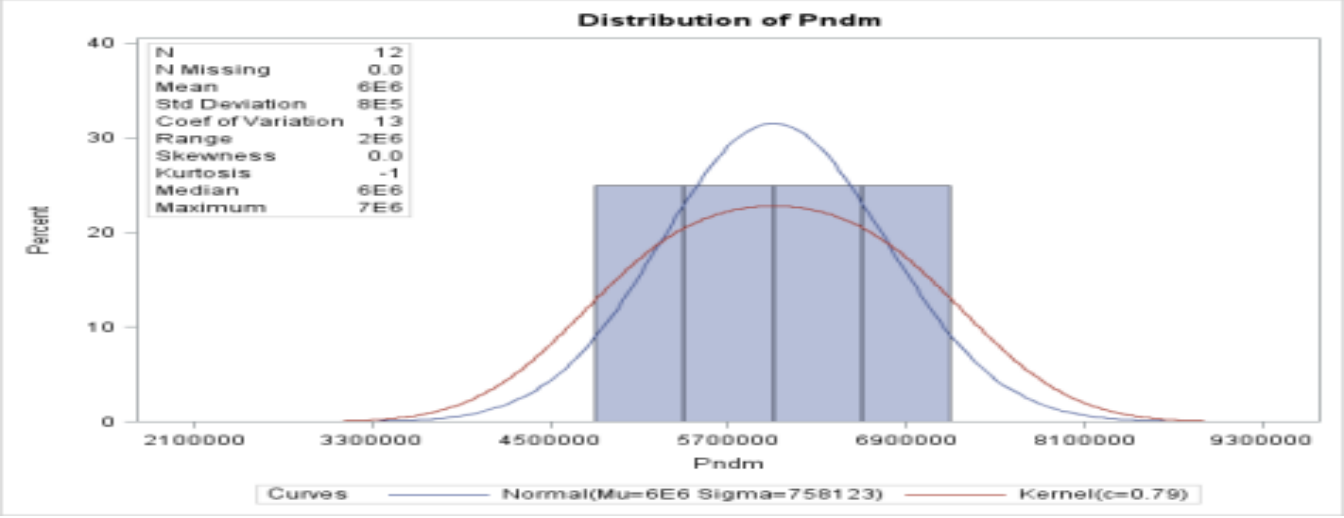


Figure 13 SGPlot for Pndm vs Pdm with year as imposing line pressure

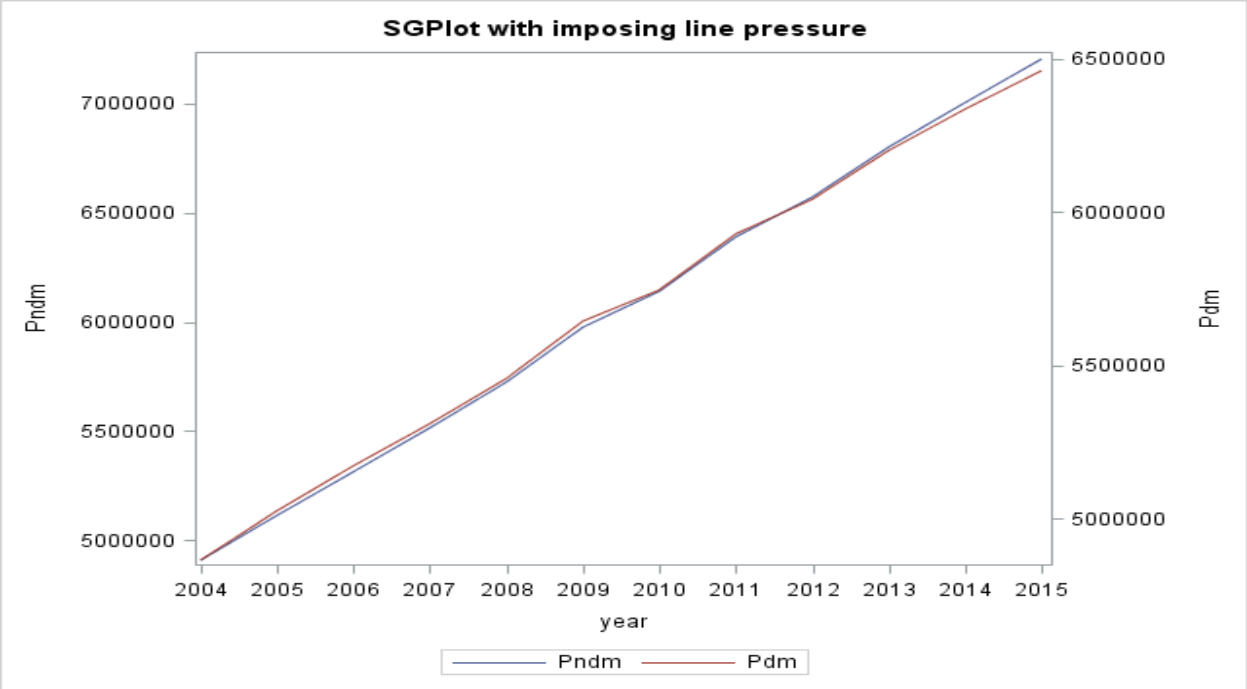
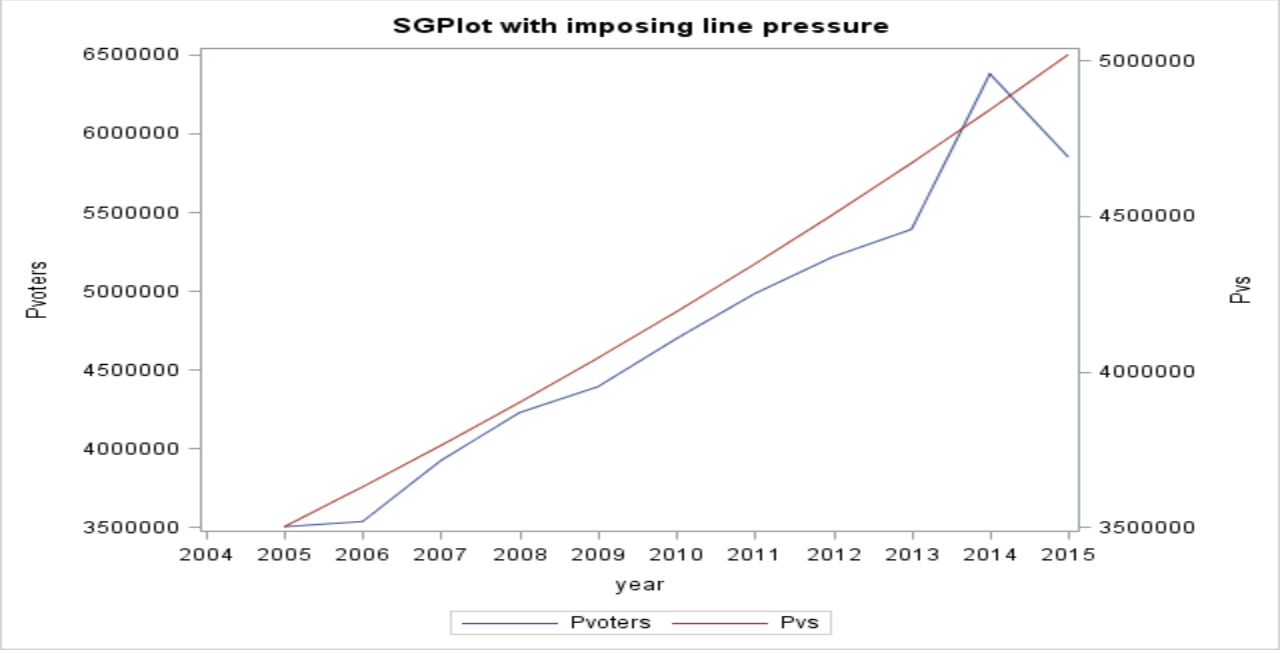


Figure 14 SGPlot for Pvoters vs Pvs with year as imposing line pressure



Appendix- C: Path Analysis and linear model comparison results.

lavaan 0.6.15 ended normally after 33 iterations	
Estimator	ML
Optimization method	NLMINB
Number of model parameters	59
Number of observations	27
Model Test User Model:	
Test statistic	19.802
Degrees of freedom	19
P-value (Chi-square)	0.407

lavaan 0.6.15 ended normally after 33 iterations	
User Model versus Baseline Model:	
Comparative Fit Index (CFI)	0.948
Tucker-Lewis Index (TLI)	0.819

lavaan 0.6.15 ended normally after 33 iterations	
Root Mean Square Error of Approximation:	
RMSEA	0.040
90 Percent confidence interval - lower	0.000
90 Percent confidence interval - upper	0.175
P-value H_0: RMSEA <= 0.050	0.486
P-value H_0: RMSEA >= 0.080	0.400

```
lavaan 0.6.15 ended normally after 33 iterations

Standardized Root Mean Square Residual:

  SRMR                                            0.076

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured
```

```
Linenar Model results
Call:
lm(formula = RegisteredVoters ~ Pop18 + Pop18d + Pop18m + Pop18dm +
    Pn_Pop18 + Pba + Pop18_d + Pop18_m + Pop18_dm + Votes + Turnout)

Residuals:
    Min      1Q  Median      3Q     Max
-347446  -48578       0   41011  353797

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.441e+06  1.137e+06   2.148 0.048501 *
Pop18       -7.628e-01  4.314e-01  -1.768 0.097353 .
Pop18d       1.150e+03  2.360e+02   4.875 0.000202 ***
Pop18m      -5.910e+00  2.668e+00  -2.215 0.042657 *
Pop18dm     -1.148e+03  2.355e+02  -4.874 0.000202 ***
Pn_Pop18     1.204e+00  4.451e-01   2.706 0.016270 *
Pba         -2.049e-01  2.781e-01  -0.737 0.472610
Pop18_d     -1.163e+03  2.920e+02  -3.985 0.001196 **
Pop18_m      7.928e+00  2.265e+00   3.500 0.003223 **
Pop18_dm     1.159e+03  2.927e+02   3.960 0.001258 **
Votes        1.849e+00  2.919e-01   6.333 1.34e-05 ***
Turnout     -6.929e+06  1.027e+06  -6.745 6.58e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 181000 on 15 degrees of freedom
Multiple Adjusted R-squared:  0.963
F-statistic: 62.44 on 11 and 15 DF,  p-value: 1.783e-10
```

BIBLIOGRAPHY

Jacques Demezier

Candidate for the Degree of

Master of Science Mathematics

Thesis: UNDERSTAND THE PROBLEM OF THE HAITIAN DATA MANAGEMENT USING A DIFFERENCE EQUATION : A CASE STUDY ON THE CALCULATION METHOD USED BY THE NATIONAL OFFICE OF IDENTIFICATION (ONI) TO GENERATE THE LIST OF VOTERS FOR THE GENERAL ELECTIONS OF 2015 AND 2016.

Major Field: Mathematics

Biographical: Data management Consultant and middle school mathematics teacher at Memorial Middle School.

Personal Data: Resides in Orlando, Florida. Single with four children.

Education: (prior degrees) 2016 MA in Modeling and Simulation from the University of Central Florida.

Completed the requirements for the Master of Science in Mathematics, Portsmouth, Ohio in July 2023

July 21, 2023

ADVISER'S APPROVAL: Type Adviser's Name Here