

# Population in Nigeria.

A STUDY OF DETERMINING THE  
BEST FORECAST METHOD.

## AUTHORS:

- EGERSON DANIEL
- NDIFREKE GODDEY
- OMOTAYO OLAWUYI
- ODIDI NAOMI TRACY
- EGBONWON OMOBOLANLE THESSY.

# The Federal Republic of Nigeria.

G A I N T   O F   A F R I C A .

---

Nigeria is the most populous country in Africa with over 200 million citizens and a growth rate of 2.58% as per the last census (NIG,2020). It is home to the biggest tech powerhouse and richest man in Africa. It is also the poverty capital of the world.



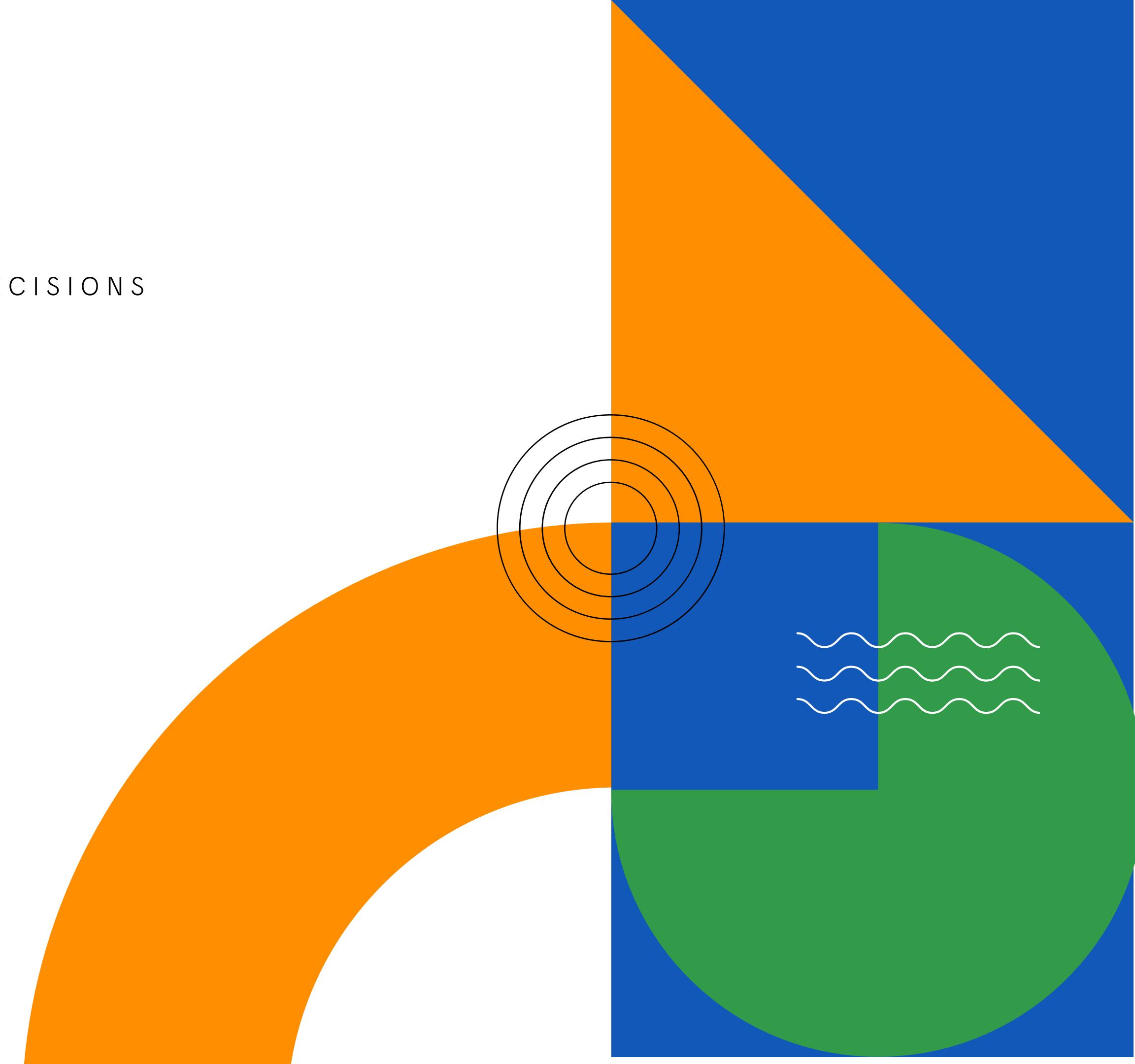
# Objectives

USING DATA TO MAKE STRATEGIC DECISIONS

---

We obtained the population data from 1950 to 2020 and the objective of this study is to:-

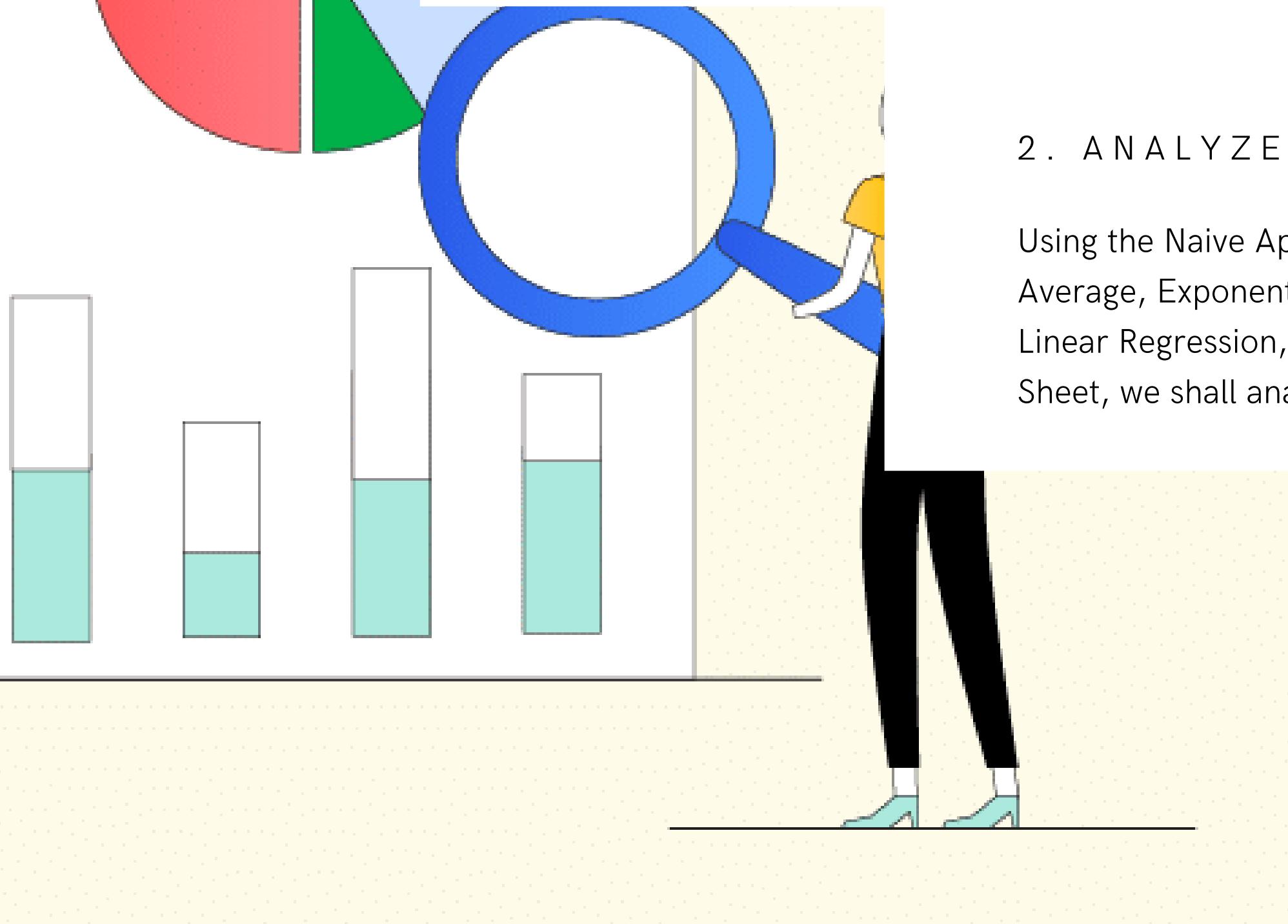
- analyze said data using different methods to forecast the population.
- determine which predictive method is the best fit for the data.



# Methodology

## 1. CLEAN DATA

We first ensure data is accurate and in the right format. To do this we filter duplicated and empty cells, retain the unique values etc.



## 2. ANALYZE DATA

Using the Naive Approach, Moving Average, Exponential Smoothing, Simple Linear Regression, Forecast, and Forecast Sheet, we shall analyze the data.

## 3. MAKE DEDUCTIONS .

We determine which approach is best based on the result of the concluded analysis.

# Definition of Terms.

## WHAT TO EXPECT



### MEAN ABSOLUTE DEVIATION (MAD).

This is the average of the absolute value of the error. It tells the story of if the error was underestimated or overestimated.

### MEAN SQUARED ERROR (MSE)

This is the average of the squared individual errors. The smaller the MSE values, the more stable the model. Interpreting the MSE can be misleading because it will always accentuate large error values.

### MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

It measures the (absolute) size of each error in percentage. It is regarded as a better error measurement as it does not accentuate large error values.

### SEASONALITY INDEX

It is an average used to compare an actual observation relative to what it would be if there were no seasonal variation. An index value is attached to each period of the time series within a year.

### ERROR & ABSOLUTE ERROR

Error is the difference between the actual value and the forecasted value.

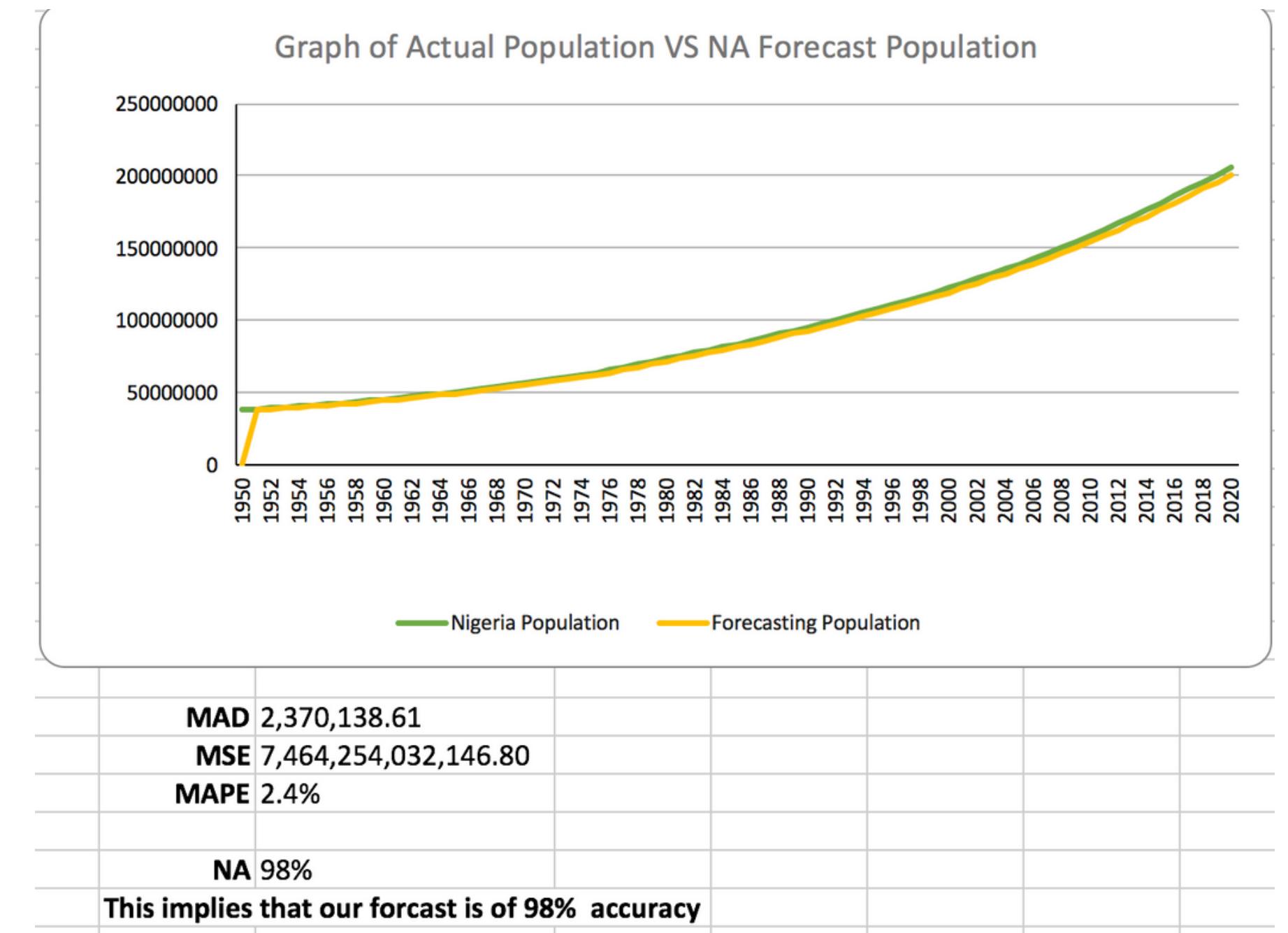
Absolute Error is the error value in positive values.

# Naive Approach

[CLICK HERE FOR THE EXCEL SHEET](#)

The naïve approach considers what happened in the previous period and predicts the same thing will happen again.

Using Excel, our data produced a 98% accuracy using the Naive Approach method. From the graph, we observe a positive relationship between the actual and forecasted populations such that as one increases the other increases. In-depth information as to how we arrived at this value is in the link above.



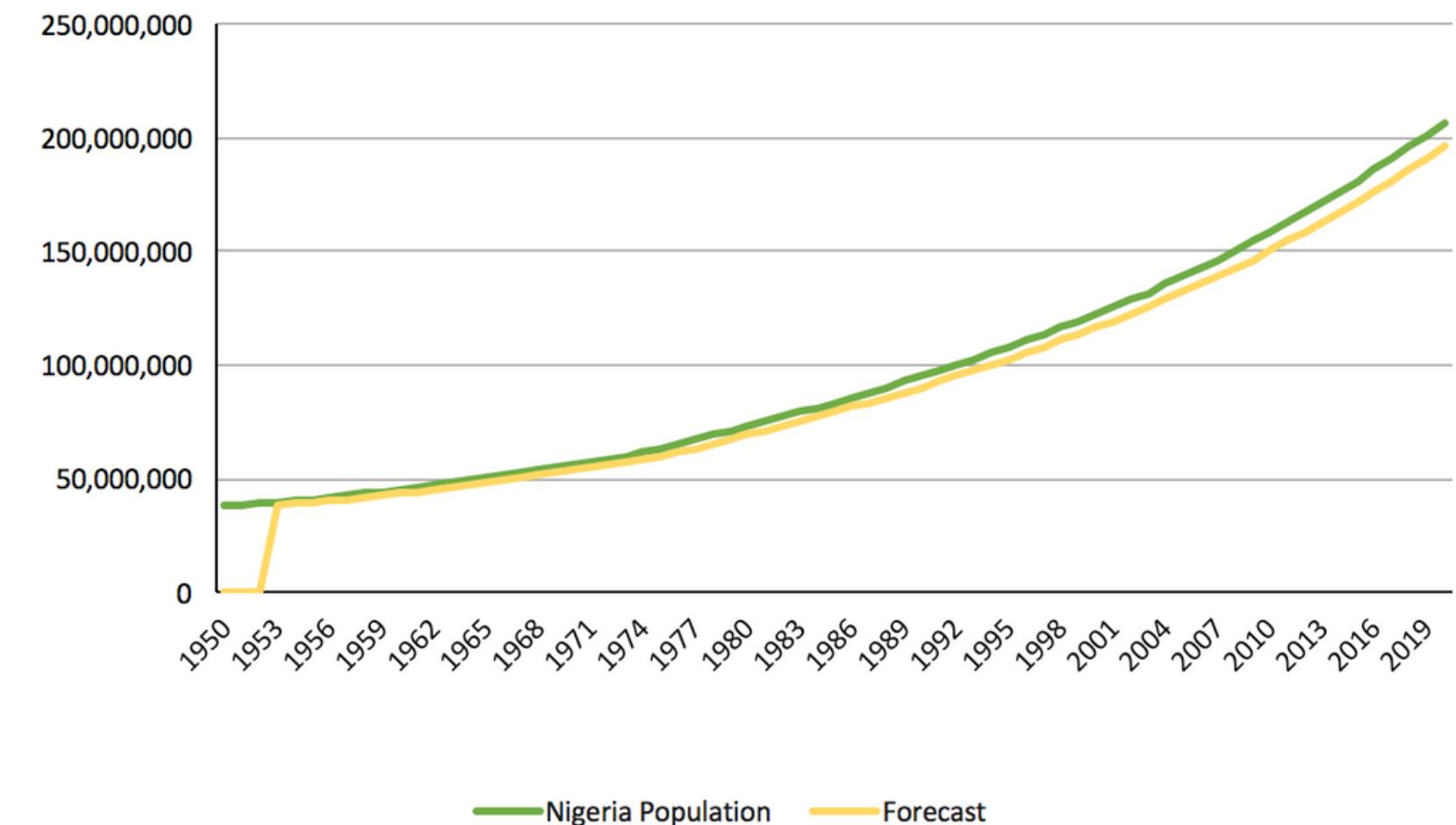
# Moving Average

[CLICK HERE FOR THE EXCEL SHEET](#)

The idea behind the moving average (MA) model is that the average performance of the recent past is a good predictor of future performance. The key element in using the moving average model is the proper selection of the number of past records. In this study, our period is 3.

Using Excel, our data produced a 95% accuracy using the Moving Average method. From the graph, we observe a positive relationship between the actual and forecasted populations such that as one increases the other increases. In-depth information as to how we arrived at this value is in the link above.

Graph of Actual Population VS MA Forecast Population



**MAD= 4,825,521**

**MSE= 30,092,029,713,772**

**MAPE= 5%**

**MA= 95%**

**This implies that our forecast is of 95% accuracy**

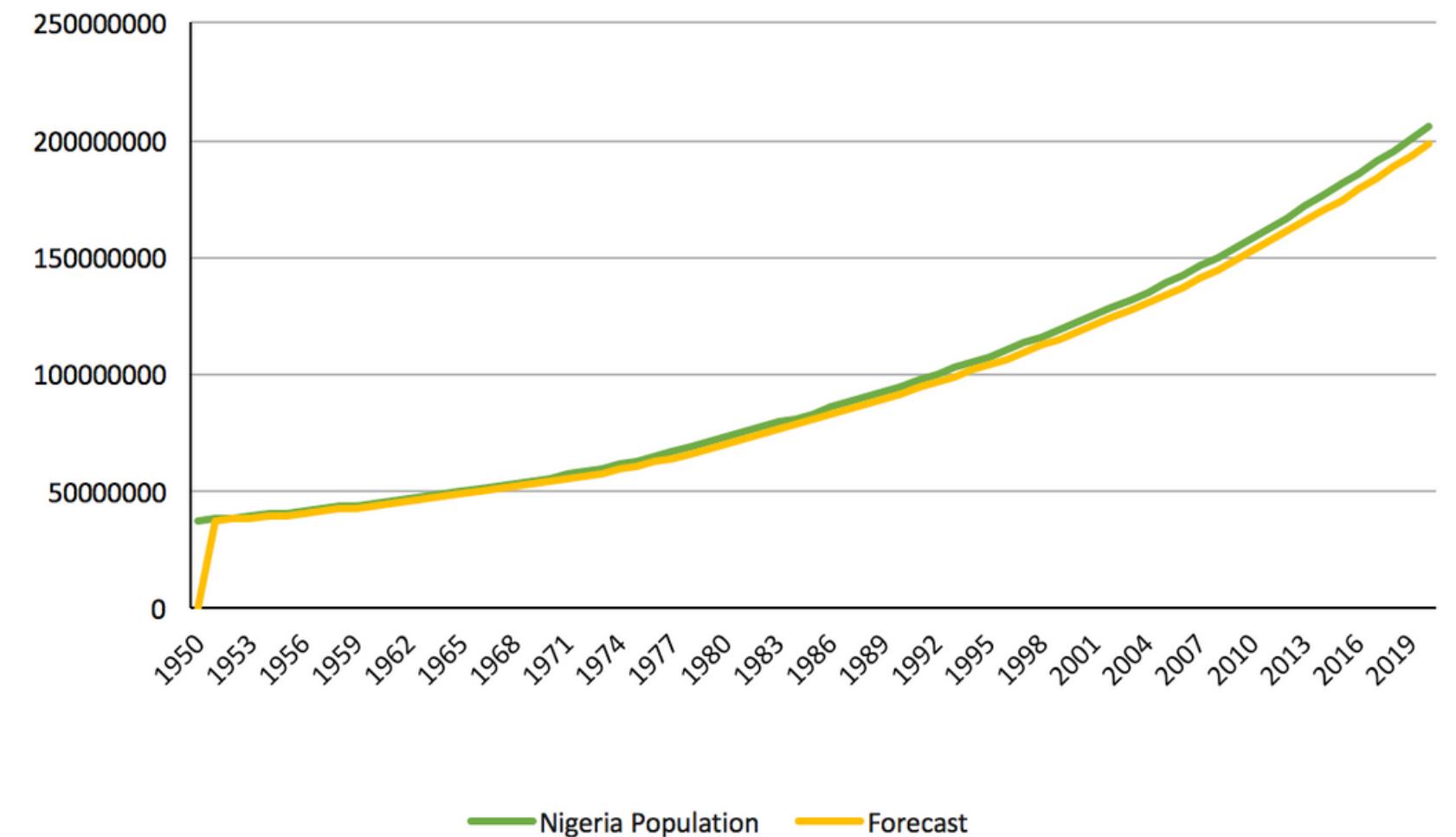
# Exponential Smoothing

[CLICK HERE FOR THE EXCEL SHEET](#)

Exponential smoothing also uses averages. Rather than just getting an average and using it as the next forecast, it decreasingly weighs exponents depending on outside factors, like season or age of a product. This is done to 'smooth' the averages and create a reliable forecast.

Using Excel, our data produced a 96.7% accuracy using the Exponential Smoothing method. From the graph, we observe a positive relationship between the actual and forecasted populations such that as one increases the other increases. In-depth information as to how we arrived at this value is in the link above.

Graph of Actual Population VS ES Forecast Population



**MAD=** 3,341,601.67

**MSE=** 14,897,382,523,947.00

**MAPE=** 3.31%

**ES= 96.7%**

This implies that our forecast is of 96.7% accuracy

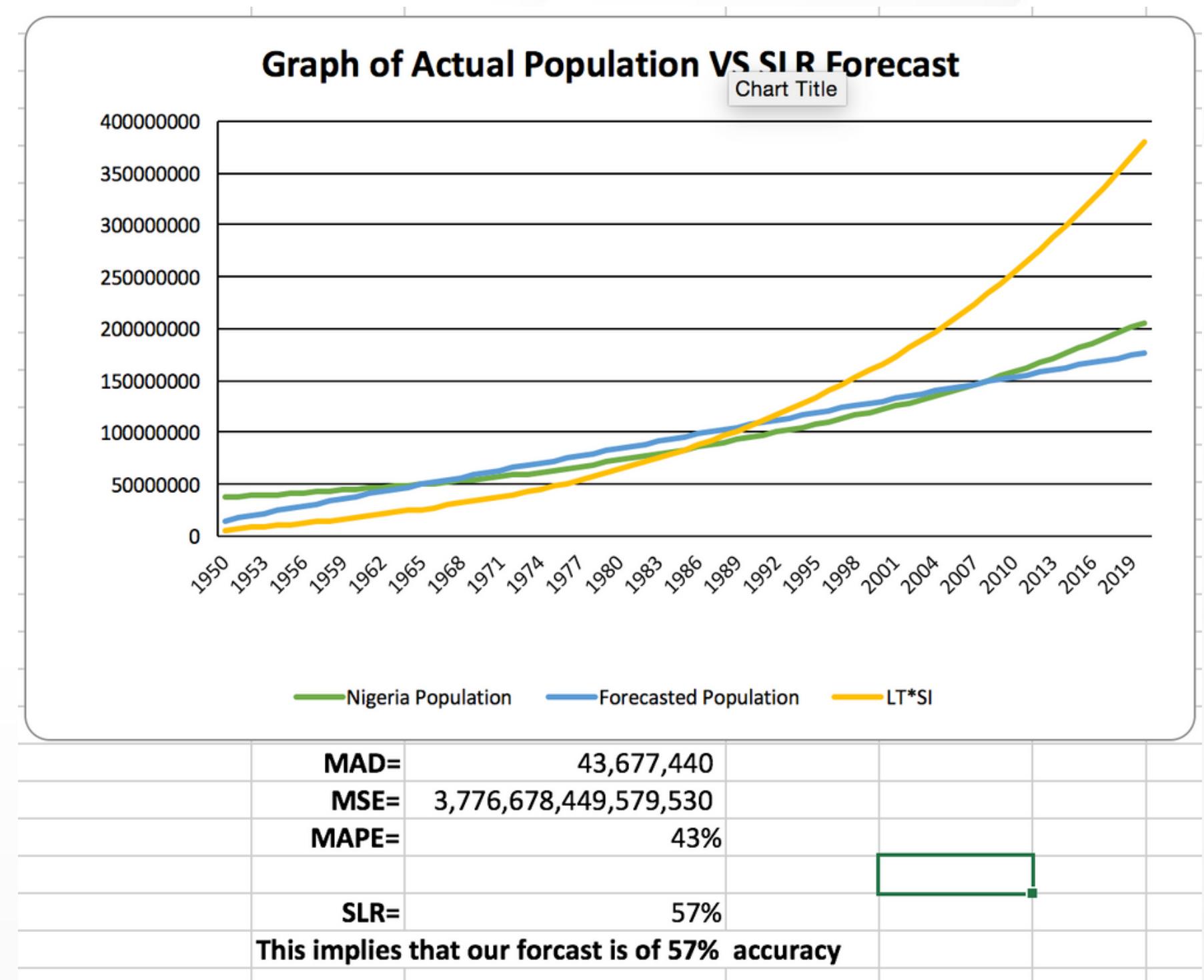
# Simple Linear Regression

[CLICK HERE FOR THE EXCEL SHEET](#)

Regression is a causal forecasting method that fits curves to the entire data set to minimize the forecasting errors. It's used to estimate the relationship between two quantitative variables.

We can understand the strength of their relationship using SLR.

Using Excel, our data produced a 57% accuracy using the Simple Linear Regression method. In-depth information as to how we arrived at this value is in the link above.

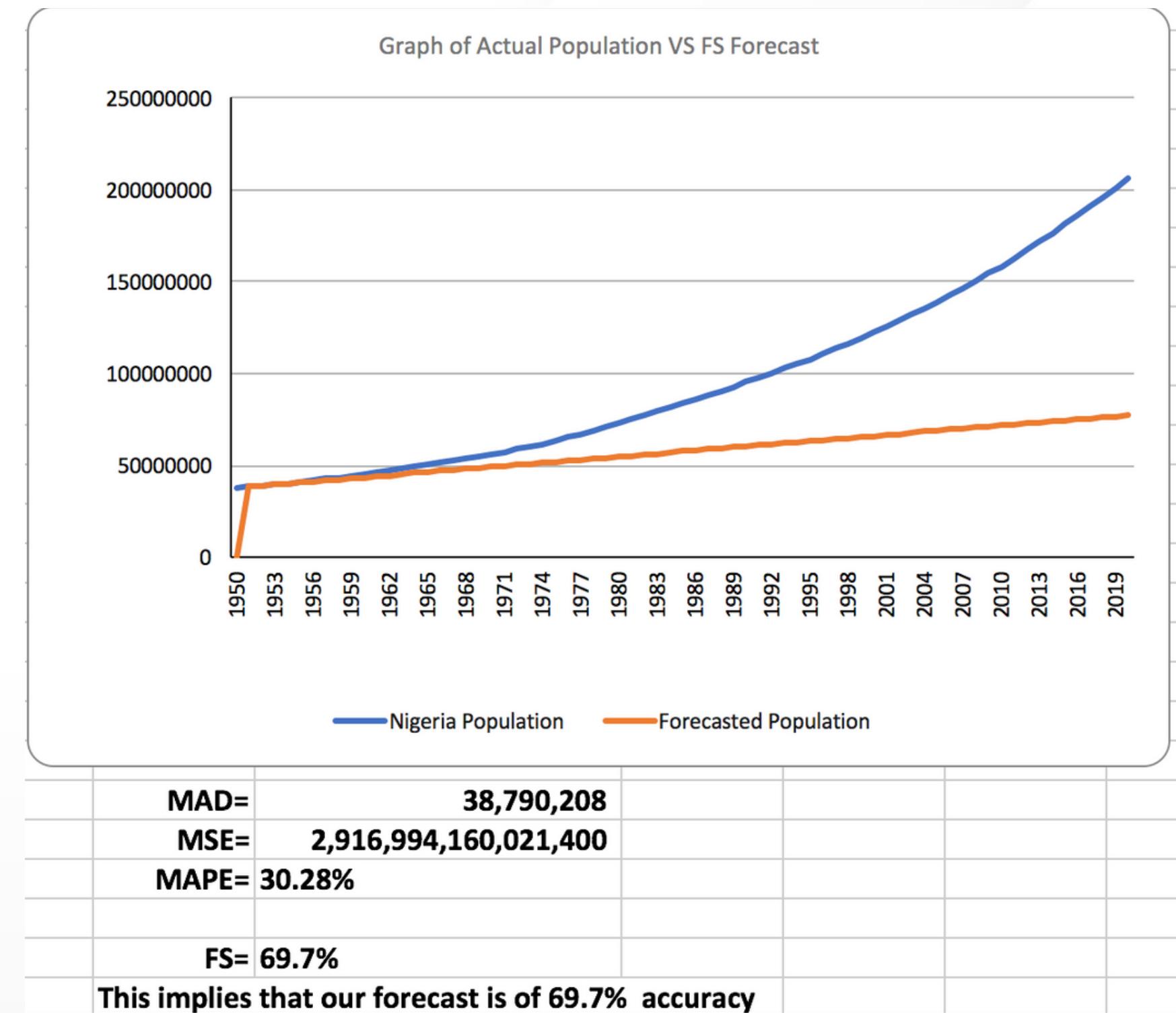


# Forecast & Forecast Sheets

[CLICK HERE FOR THE EXCEL SHEET](#)

Forecasting is a technique that uses historical data as inputs to make informed estimates that are predictive in determining the direction of future trends.

Using Excel, our data produced a 69.7% accuracy using the Forecast method. In-depth information as to how we arrived at this value is in the link above.



# Conclusion



The forecasting methods were compared against each other using MAD, MSE, and MAPE. Of the three accuracy measurements, MAPE provides the most accurate and fair comparison of the forecasting methods. Using the Naive Approach is the best fit predicting method for our data set.

	NA	MA	ES	SLR	FS
MAPE	98%	95%	96.7%	57%	69.7%
	98%	95%	96.7%	57%	69.7%

# Meet The Team





# The learning continues

We have only just began. Here's to the team with  
the best visuals (we know) 😊 🥂

