Agenda 2063 Progress Analysis

Hussam Hanafi

11/3/2019

# Introduction

[Agenda 2063](https://au.int/en/agenda2063/overview) is the African Unions’ Strategic framework for delivering inclusive and and sustainable development signed on May 2013.

As part of this agenda the AU has set over-arching goals and more granular targets to be achieved in the 50 years from 2013 to 2063.

The purpose of this report is to evaluate progress on two of those targets in the medium term [First Ten Year Implementation Plan](https://au.int/sites/default/files/documents/33126-doc-ten_year_implementation_book.pdf) by analysing Indicator data from the [World Bank Open Data](https://data.worldbank.org/) Initiative.

## First Ten Year Implementation Plan Targets

The Two Targets were selected partially arbitrarily in the hope that they would be indicative of each other; however, the selection was also driven by the completeness of the data sets as there is often large gaps in data sets dealing with progress indicators due to a number of factors mostly affecting the data collection stage.

For this Report the two targets/indicators chosen were as follows:

* As part of the first Goal: A High Standard of Libing, Quality of Life and Well Being for All; the FTYIP sets a target of reducing Unemployment relative 2013 by at least 25%
* As Part of the fourth Goal: Transfromed Economies and Job Creation; the FTYIP sets a target of 7% Annual growth

This Report will use those two indicators to attempt to analyse progress and attempt to predict the value of these indicators by 2023.

# Data Analysis

We will analyse each of these indicators Seperately using the following steps:

1. Accessing the data from the World Bank Open Data web-site
2. Filtering out AU Memeber States
3. Wrangling the data into tidy format
4. Initial Visualization
5. Wrangling the data to minimize the effect of missing data points
6. Visualizing the Data for Analysis Purposes
7. Subsetting a Test Set to be used with our models

## Unemployment Rate

### Data Wrangling

We Will start the with the Unemployment Rate, Loading the Data from the World Bank modelled estimate with the code below:

### Loading Unemployment Rate  
## Creating Raw Data Folder  
Raw\_Data\_Folder <- file.path(getwd(),"Raw\_Data")  
dir.create(Raw\_Data\_Folder)  
  
##Downloading and Unpacking the File  
#download.file("http://api.worldbank.org/v2/en/indicator/SL.UEM.TOTL.ZS?downloadformat=csv",   
# file.path(Raw\_Data\_Folder,"API\_SL.UEM.TOTL.ZS\_DS2\_en\_csv\_v2\_422140.zip"), mode = "wb")  
Unemployment\_zip <- "API\_SL.UEM.TOTL.ZS\_DS2\_en\_csv\_v2\_422140.zip"  
Unemployment\_csv <- "API\_SL.UEM.TOTL.ZS\_DS2\_en\_csv\_v2\_422140.csv"  
unzip(file.path(Raw\_Data\_Folder,Unemployment\_zip), exdir = file.path(Raw\_Data\_Folder,"Unemployment"))  
  
##Reading the File into R  
Unemployment <- read\_csv(file.path(Raw\_Data\_Folder,"Unemployment",Unemployment\_csv), skip = 3)

Wrangling our Data into tidy format with the follwing code:

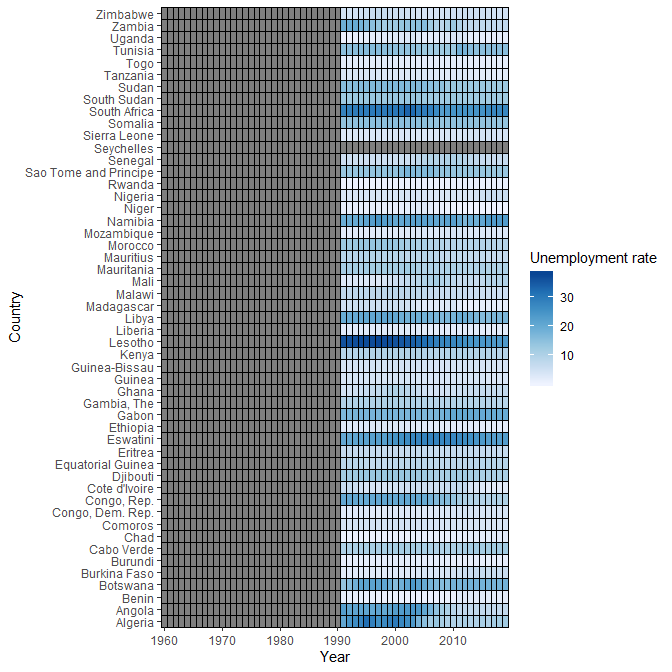
## Tidying the Data  
colnames(Unemployment) <- str\_replace(colnames(Unemployment), " ", "\_")   
Unemployment <- Unemployment %>%  
 filter(Country\_Name %in% AU) %>%   
 gather(key = "Year", value = "Unemployment", `1960`:`2019`, convert = TRUE) %>%   
 select(-c(X65,Indicator\_Code)) %>%   
 filter(Country\_Name %in% AU) %>%   
 select(-Country\_Code, -Indicator\_Name)

Below is a sample of the data frame:

Unemployment %>% sample\_n(4) %>% knitr::kable(col.names = c("Country","Year","Unemployment rate"))

|  |  |  |
| --- | --- | --- |
| Country | Year | Unemployment rate |
| Burkina Faso | 1961 | NA |
| Eswatini | 1962 | NA |
| Burkina Faso | 2019 | 6.15 |
| Malawi | 1999 | 8.43 |

Our Initial Visualization of the data shows some clear places where we have missing data which we’ll work to remove:



The below code achieves this:

## Removing years and countries with no data  
No\_Data <- Unemployment %>%  
 group\_by(Year) %>% summarize(No\_Data = sum(!is.na(Unemployment)))   
Unemployment <- Unemployment %>%   
 left\_join(No\_Data, by = "Year") %>% filter(No\_Data != 0) %>% select(-No\_Data)   
  
No\_Data <- Unemployment %>% group\_by(Country\_Name) %>% summarize(No\_Data = sum(!is.na(Unemployment)))   
Unemployment <- Unemployment %>%   
 left\_join(No\_Data, by = "Country\_Name") %>% filter(No\_Data != 0) %>% select(-No\_Data)

However since the stated goal is to reduce Unemployment by 25% relative to 2013 we will extract the 2013 values and use those as our Baseline; additionally, the below code will also set our test data and remove it from our training set:

## Setting Baseline and Test Sets  
# Baseline  
Unemployment %>% group\_by(Country\_Name) %>% filter(Year == c(2013)) %>%   
 summarize(Baseline = Unemployment) %>% filter(is.na(Baseline))

## # A tibble: 0 x 2  
## # ... with 2 variables: Country\_Name <chr>, Baseline <dbl>

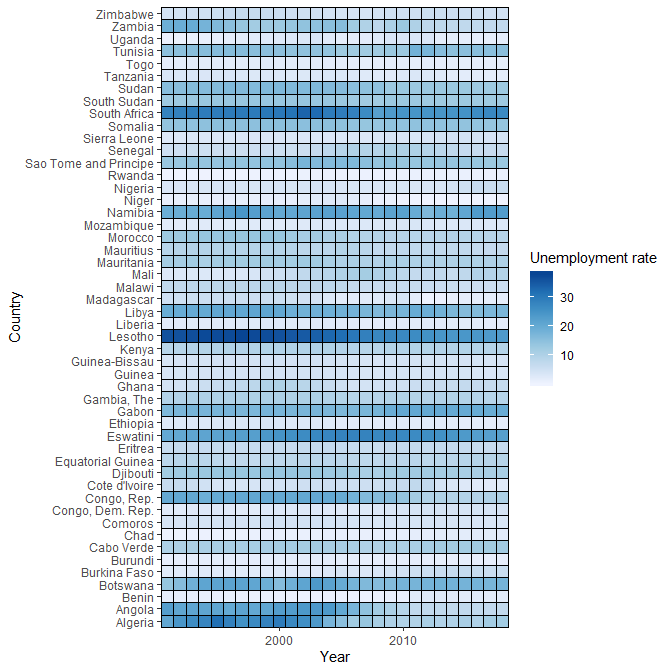
Baseline\_Unemployment <- Unemployment %>%   
 group\_by(Country\_Name) %>% filter(Year == (2013)) %>%   
 summarize(Baseline = Unemployment) %>% filter(!is.na(Baseline))  
  
# Tests Set  
Unemployment %>% group\_by(Country\_Name) %>% filter(Year == c(2019)) %>%   
 summarize(Baseline = Unemployment) %>% filter(is.na(Baseline))

## # A tibble: 0 x 2  
## # ... with 2 variables: Country\_Name <chr>, Baseline <dbl>

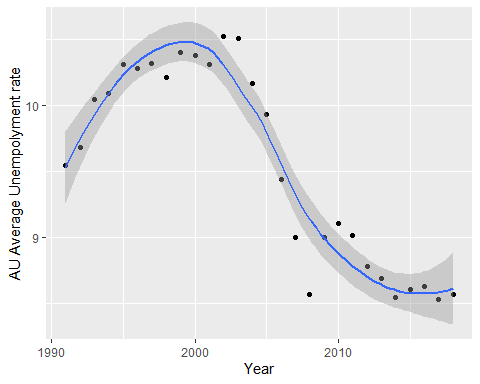
Unemployment\_test <- Unemployment %>% filter(Year == 2019)  
  
# Removing Test Data  
Unemployment <- Unemployment %>% filter(Year != 2019)

### Visualizations and Analysis

Now we are ready to Look at our data and extract some useful analysis. We repeat the initial visualization to confirm:

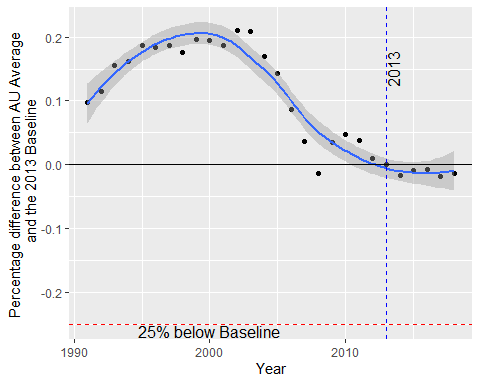


Next we Look at the Continent Wide Average Unemployment rate and try discern any trends:



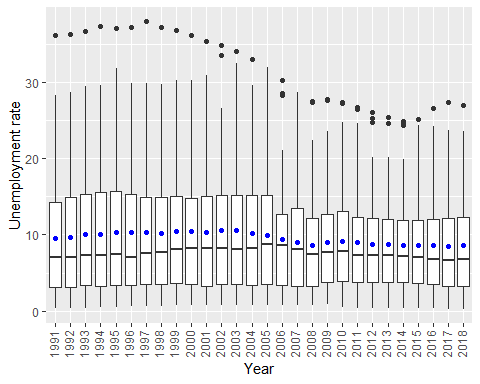
We can clearly see a downward trend with a potential leveling off, however this doesn’t provide a meaningful comparison to the stated target, instead we take a comparison between the average and the baseline we set:

## Warning: `data\_frame()` is deprecated, use `tibble()`.  
## This warning is displayed once per session.

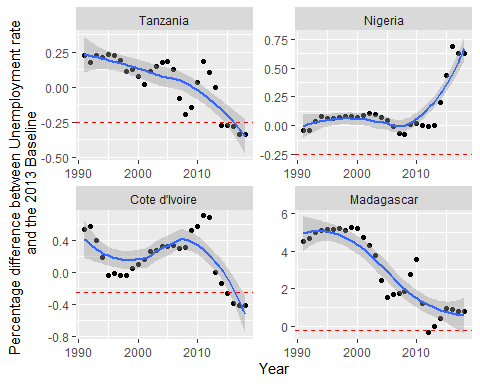


As expected we see the same trend, it is however clear that the continent wide performance leaves something to be desired with the reduction in Unemployment rate by 2019 not approaching the target.

A Boxplot showing the unemployment rate over the years paints a more nuanced picture showing large variations between Member States and a mean (in blue) that is significantly larger than the median implying that most member states are perfroming better than average:



To illustrate this variation we look at the two top and bottom performing countries in 2018



## Annual GDP Growth

We Repeat this process with data relating to GDP Growth, in this section of the report we will show less of the code as there are essential similarities in the process.

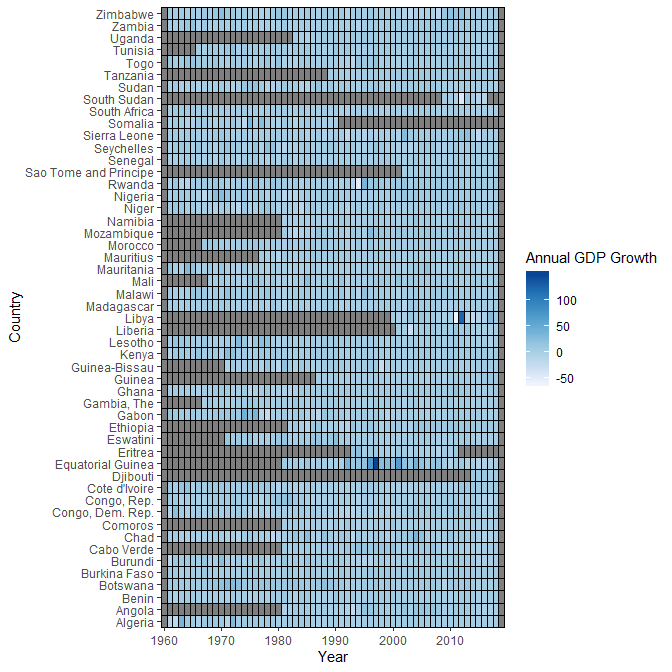
### Data Wrangling

Once we have loaded and wrangled the Annnual GDP Growth Data from the World Bank Open Data website

Below is a sample of our tidied data:

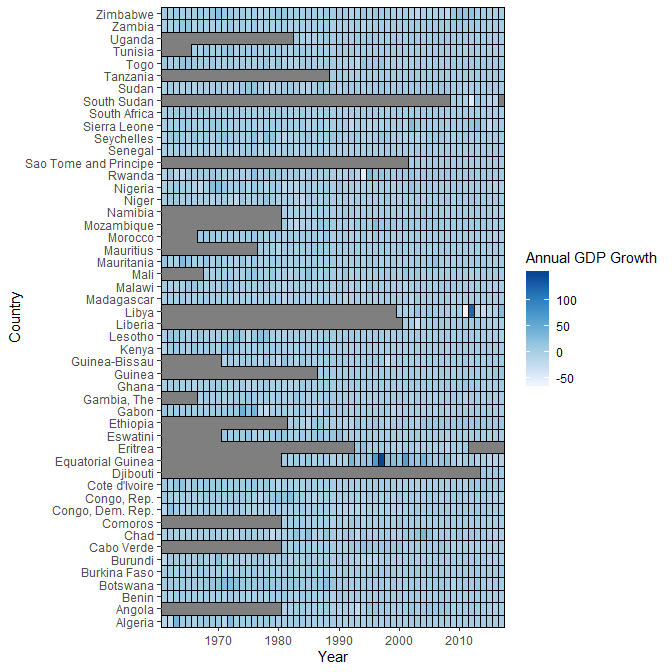
|  |  |  |
| --- | --- | --- |
| Country | Year | Unemployment rate |
| Niger | 2010 | 8.36 |
| Ghana | 1978 | 8.48 |
| Togo | 1984 | 5.56 |
| Senegal | 2018 | 6.77 |

While here we make our initial visualization



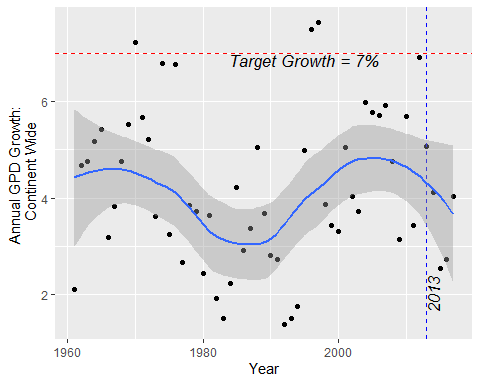
We can see that the missing data is more sporadic than with the Unemployment rate, we clean up the missing data opting to estimate missing data for our test set with data from the closest year. The exception to this is Somalia where the latest data is from 1990 thus we exclude it from this analysis.

We now visualise the data in our Wrangled data set after the exclusion of the Test data

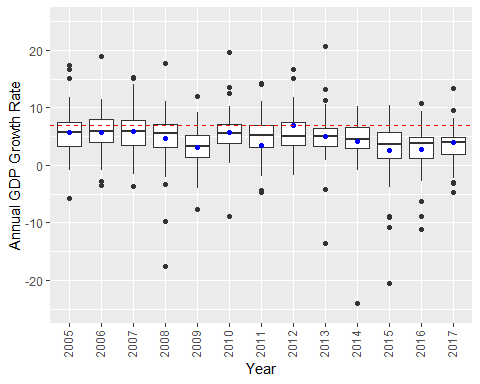


### Visualization and Analysis

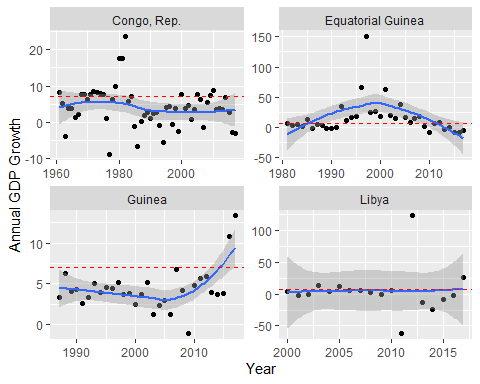
We start by looking for Continent wide trends:



The Scatter plot shows a massive variation from year to year with most of the years falling below the 7% Target. A Boxplot confirms this variation; please note that for clarity we only look at a subset of the years and do not show the largest outliers:

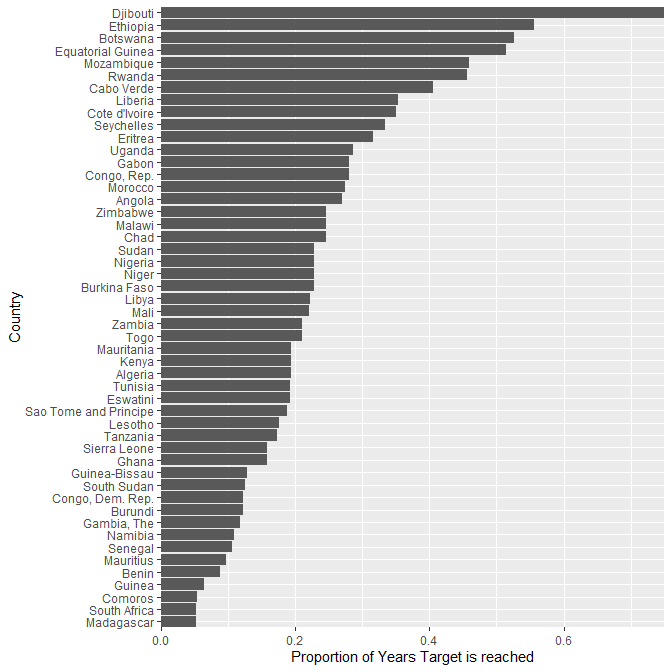


From here it is clear that the majority of Member states are not achieving the set target of 7%, a look at the countries with the highest and lowest growth rates provides a clearer picture of individual trends:



However it fails to give us an idea of what countries are performing since a consistent growth is just as if not important to a high growth; additionally, countries like Libya with high variation that suffer large GDP losses can skew the trend by making equally large gains.

We propose that looking at the proportion of years the target is reached is a more indicative metric:



If we narrow our view to the years since the launch of Agenda 2063 we get yet a different look at how Member States are performing:

