CHANGE POINT ANALYSIS TO QUANTIFY THE IMPACT OF AFRICAN GOVERNMENT POLICY INTERVENTIONS TO SLOW THE SPREAD OF COVID-19.

BUSINESS OBJECTIVE

To Quantify the statistical significance of a public health policy introduced by african governments to slow the spread of COVID-19.

Task 1.1 DATA ANALYSIS WORKFLOW

The SIR/SEIR Model is one of the compartimental models which are commonly used to study the number of people having infectious disease e.g Covid-19 in a population. The model categorizes each individual in the population into three groups:

- Susceptible(S)- The number of susceptible people
- Infectious(I)- The number of infectious people
- R - The number of recovered(and immune) or the deceased

These variables represent the number of people in each compartment at a particular time. This simple model can be formulated in terms of ordinary differential equations which enable analytical treatment or fast evaluations, Thus this Model will help us in simulating the effects of different possible interventions implemented by Senegal government aimed at slowing the spread of Covid-19.
The model will also help us in estimating the actual effects of the measures taken and scenario forecasting for building control Strategies in mitigating the epidemic.
Due to the rapid spread of Covid-19,there is a need for time critical decisions on containment and also control strategies developed for mitigation of the disease.
During the initial stages of the epidemic, when the numbers are low, there is large statistical and systematic error and there is a delay when the policies are put in place and the impact of the policies. Prior knowledge is integrated into modelling efforts to reduce the uncertainties e.g the reporting Delay is D is estimated and integrated using Bayesian inference.
SEIR Model or susceptible-exposed-infectious-recovered model is a derivative of the basic SIR model with an incorporation of the latency or incubation period.

Distinct characteristics of Exponential functions

- Domain - The domain of exponential function is all real numbers (from infinity to infinity)
- Range - The range is from zero to infinity
- Exponential function is neither even nor odd hence No symmetry
- It is bounded below by 0 but no upper bound

- There is a horizontal asymptote of y=0 but there is no vertical asymptote.

The similarity between exponential growth,decay,geometric progression and logistic growth is that all of them are mathematically exponent.

Differences between Exponential Growth, Decay, Geometric progression and logistic growth.

-The population passes the carrying capacity while in logistic growth it seldom passes the carrying capacity.

-In Exponential growth, a steady phase or stationary state is seldom achieved while in logistic growth a steady state can be achieved.

-In Exponential growth, the values increase over time, in logistic growth the values at first increase but they decrease as the population approaches the carrying capacity and in exponential decay, the values decrease over time exponentially.

The number of susceptible people and infected people over time are nonlinear in nature which makes the Model to be a system of nonlinear ordinary differential equations. Which is during the initial stage of Covid-19 outbreak, the number of infections and recoveries are smaller compared to the susceptible people, therefore the differential equation for the infected reduces to a simple linear equation exhibiting exponential growth.

For one to get the whole picture of the spread of Covid-19 disease, it is important to focus on the initial phase of the outbreak of Covid-19 before any serious mitigation policies are implemented by the government. And from this, you can estimate the central epidemiological parameters to be used in SIR/SEIR models.

The prior and posterior distributions using Markov chain Monte Carlo sampling are used to model the rate parameters for the SIR/SEIR model, this is because at the initial stage of the outbreak of Covid-19, the parameters are unknown to us. We chose informative priors based on available knowledge and chose uninformative priors for the remaining parameters. We keep the informative prior random so that the data would constrain the parameters. Student T distribution is also used because of its ability to allow more outliers than normal distribution.

The results of the model are expected to be a reduction in the spreading rate which is a decline in the daily new cases which is related to the government interventions and the individual behavioural change. Predictions for future dates can be made by bayesian inference.

The objective of the project is to detect changes in the spread of Covid -19, the changes can be analysed to quantify their effectiveness in slowing the spread of the disease. These change points can be related to non-pharmaceuticals hence the right to say change point detection quantifies the effect of non-pharmaceuticals interventions.

Task 1.2 PUBLIC INTEREST IN CORRUPTION

Business objective: To Disentangle the relative interest of people in corruption issues .

Twitter is a platform where people share their interests, their emotions and other forms of expressions. The spread of news in Twitter platforms is faster because of the large network of communities and followers,following,retweeting connections, hence mathematical modelling can be applied to group twitter users according to their interest levels and reactions to news outburst.

At the beginning of February, reports emerge that an important politician was taking bribes,at the break of the news:

-Every twitter user is susceptible to developing an interest in the corruption issues and therefore we can group these into a Susceptible compartment.

- At the outbreak we have a number of people who already have their interest on corruption issues, We put them into the interested compartment.
- As the news spreads we have those who lose their interest on corruption issues, they will either not develop the interest again or they will develop if more information is given about corruption. Those who might develop the interest again they transit back to being Susceptible while those who won't develop the interest we can say they are removed.

From these we get S- susceptible, I - interested, S-susceptible or R- removed.

Assumptions.

1. We will assume that the rate of increase in interest is proportional to the tweets, retweets,hashtags and user mentions in corruption issues between the susceptibles.
2. We will assume that there is a constant rate with which people lose their interest on corruption issues.

To assess the effect of the news about the politician being convicted, we need to focus on the first phase and integrate our knowledge on the spread of news on twitter to estimate the parameters for short term forecasting and evaluation of the effect of the corruption issues.

Using prior knowledge we choose the central parameters to be, the tweeting rate,losing rate, dormant rate, the number of those initially with an interest in corruption issues.

We can perform bayesian inference on the central parameters for the SIR/SIS model using the Markov chain Monte Carlo sampling. To simulate the effect of the corruption issues, we model the effect of conviction of the politician on 15th August, as a change point in the tweeting rate of corruption issues.

Considering the SIS Model where after interested parties lose interest and they become susceptible again, and when triggered by other corruption issues, they become interested again. This means that in the long run, everyone will be interested and we will have a steady rise in corruption cases.

Task 2.

At a time like this when there is a pandemic, prompt decisions and measures are required by governments in order to mitigate the spread. At the initial stage of the pandemic, We have the whole population with everyone being susceptible to the disease, we have those already infected and those who will recover. This sets a class of compartments which can be modeled as [Susceptible-Infected-Recovered(SIR)] and the rates at which they change(susceptible people becoming infected and infected becoming recovered), hence the SIR model is best to model the Covid19 spread. During this initial, neither the central epidemiological parameters nor the effectiveness of the government interventions are known, we inferred the spreading rate, recovery rate, reporting Delay and the number of infected people in Senegal, then we applied the Bayesian inference based on Markov chain Monte Carlo sampling to a class of compartmental models SIR. We did analysis on the temporal change of the spreading rate and the effectiveness of the interventions that were put in place by the senegal government. It also shows us the potential change points.

In Senegal, public interventions to slow the spread of Covid19 spread were implemented in three steps in one month: (i) Around march 15th, Religious activities and gatherings were stopped (ii) Around 23rd march, nationwide curfew (iii) Around 31st march, Strong social distancing. From the observed number of cases, we quantified the effect of the interventions using change point analysis. At first the change point assumed at 15th march, the spreading rate decreased from 0.40 to 0.00 with credible intervals(CIs) of [0.15-1.04] and [0.00-0.75] respectively.At the second change point, assumed around 23rd march, the spreading rate increased to 0.01 with (CIs [0.00-1.14]). In the third change point, the spreading rate increased to 0.11 with (CIs[0.08-0.20]) as shown the disgram below. In the first change, the spreading rate slowed down the spread of the virus but still it implied an exponential growth. In the last two change points, there is a rise in the spreading rate, which explains the rise in the number of cases.

(i) Mild social distancing

The spreading rate decreases to 99% . Though the people reduced the number of contacts after these interventions, the total number of reported cases continues to grow. We also observe an exponential increase in the number of new infections after the intervention becomes effective because the growth rate is positive; spreading rate - recovery rate > 0.

(ii) Nationwide Curfew

The spreading rate increases to 5%. Although the government implements nationwide curfew, still there are contacts between people with the hours before curfew, hence the total number of reported cases continues to grow alongside scenario (i) for the time period of the reporting delay D. We observe an exponential decrease in the number of new infections after the intervention becomes effective since the growth rate becomes negative; spreading rate - recovery rate < 0.

(iii) Strong social distancing

Here the spreading rate increases to 0.11 which can be explained by the assumption that even when close contact buns were implemented, still there was contact withins homes. After the delay plus the change duration, two weeks, there is a decrease in daily new infections and also a negative growth rate.

LIMITATIONS OF THE ANALYSIS

- Reliable short and long term forecasts were difficult because of the concurrent change in interventions and individuals hence change in spreading rates.

-Complication between the inference and subsequent forecasts by delay of 2 weeks between intervention and first spreading rate estimate. Because of this delay, any uncertainty in the magnitude of an intervention in the previous 2 weeks can have a major impact in the case number in the subsequent two weeks.

RECOMMENDATIONS

-During the initial stages of an outbreak, the governments need to be diligent in keeping and recording the statistics of the cases numbers and other information which is instrumental in decision making on mitigation of the epidemic.

-In measuring the effect of the interventions, it is important to take into account the reporting delay. The reporting delay D, together with the time required to implement the interventions, means that the changes in our behaviours can only be detected in confirmed cases in two weeks' time. This delay combined with the changes in spreading rates indicates that careful planning for future measures is essential.