COUNTY COVID-19 PREPAREDNESS

TEAM MEMBERS:

Amina Abdi

Brenda Cheyech

Brian Kariuki

Christian Juma

Elizabeth Josephine

Elena Njeri Mwangi

Erick Onyango

Festus Raini Mogaka

Mercy Gatuguta

Gladwel Wanjau

Introduction

COVID‑19 is a new disease, distinct from other diseases caused by coronaviruses, such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS). The virus spreads rapidly, and outbreaks can grow at an exponential rate. At present, there are no therapeutics or vaccines proven to treat or prevent COVID‑19, although national governments, WHO and partners are working urgently to coordinate the rapid development of medical countermeasures. The coronavirus disease 2019 (COVID‑19) pandemic is exacting a huge toll on individuals, families, communities, and societies across the world . Daily lives have been profoundly changed, economies have fallen into recession, and many of the traditional social, economic, and public health safety nets that many people rely on in times of hardship have been put under unprecedented strain.

The severity of the pandemic has prompted countries to put in place mechanisms to monitor the spread of the disease at both national and international levels. Kenya has not been left behind. Everyday the government through the ministry of health releases data on the number of samples tested, new cases identified , death cases encountered and recoveries. All these add to the national tally and eventually into the international tally.

In the project below our focus is mainly on counties in Kenya and its preparedness toward this global pandemic , despite other current measure in place for our analysis we have got some dataset that will take us through As the kenya faces a stressful time with the rise of the Coronavirus

The main objective of this project is to rank counties in Kenya on COVID-19 preparedness.

The ranking was characterized by the number of available beds to population ratio. To answer our analysis question, the following factors will be investigated:
The CRISP-DM was used in this project to help come up with the objectives and give a step by step process of how   to deploy the dataset selected inorder to understand county preparedness for COVID 19

As the world faces a stressful time with the rise of the Coronavirus

STAGE 1: Business understanding

The main project objective is to rank counties in Kenya on COVID -19 preparedness.

Business questions that will answered in this project are:

1. Get difference between recommended isolation beds and bed available (,Create rank variable based on bed preparedness - a larger value signifies larger deficit in number of beds hence lower rank)

2. Calculate bed to population ratio

3. Rank counties based on hospital count,(a higher hospital count signifies better capacity handling hence higher rank)

4. Rank counties based on cumulative cases,(a higher number signifies a higher probability of infections hence lower rank in preparedness)

5. Rank counties based on density[number of people per sq km],(a higher number signifies a higher probability of infections hence lower rank in preparedness)

6. Rank counties based on bed to population ratio,(a higher ratio means fewer people are assigned to one bed hence better preparedness hence higher rank)

7. Compute aggregate ranking of counties ,(Order counties by aggregate ranking)
   a. The top 5 prepared counties
   b. The top 5 worst prepared counties
   c. Among worst prepared counties, which county has the lowest bed to population ratio?
   d. Which county has the lowest bed to population ratio overall

Problem Statement

The project aims to examine the level of Kenyan counties' preparedness for Covid-19. The focus is to examine how the counties have achieved the metrics used by the Ministry of Health (MOH) to determine the level of preparedness to deal with the rising Covid-19 cases. The metrics we focused on in this project are:Focusing on this problem area help us to identify answer question stated above

- Population per county

- Kenya health facilities per county

- Covid 19 cases per county

- Icu bed per county

Current response to COVID 19

Counties in kenya   have introduced the following measures

- widespread population-level physical distancing measures and movement restrictions in order to slow spread Physical distancing measures and movement restrictions, often referred to as "shut downs" and "lock downs," can slow COVID-19 transmission by limiting contact between people.

- rapid case identification, rapid testing and isolation of cases, comprehensive contact tracing and quarantine of contacts have also been placed   by county healthcare

- Maintaining hand hygiene through wash u hands regularly with soap and water

- Develop safe and effective vaccines and therapeutics that can be delivered at scale and that are accessible based on need

- respiratory etiquette --Cover nose/mouth with a tissue coughing or sneezing. Wearing a mask . Use disposable tissues and discard after use.

Produce Project Plan

- Define initial process plan -identify number of hospital in every county, identifying the ratio of normal isolation and ventilators to the entire population per county ,identifying high risk counties based on the number of infection cases ,

- Put identified goals which is to rank counties in Kenya on COVID-19 preparedness. and selected techniques into a coherent procedure.

- Estimate effort and resources ,-we use python 3.8.5 together with its in built libraries eg numpy , pandas , seaborn and matplotlib.pyplot

- Identify critical steps eg first look for the dataset needed for analysis , understanding the data , preparing the data for analysis ,do the actual analysis on the data , giving recommendation and finally evaluate the final result and identifying the next step

STAGE 2 : Data understanding

Data source

We have collected 4 dataset for this project the will be helpful for our analysis the first is

1. Population per county dataset -- availed by
   https://data.humdata.org/dataset/26801642-7088-4149-9605-6ae282897937/resource/bfe3
   0344-5fb4-45dd-b1d1-565690aba5fc/download/2019_knbs_kenya_population_census_ad
   mins_17dec2019.xlsx

2. Kenya health facilities -- availed by kaggle hospital list
   https://storage.googleapis.com/kaggle-forum-message-attachments/154760/5739/KEN%2
   0Health_Facilities.csv

3. Icu bed per county -- availed by
   https://data.humdata.org/dataset/e231c2f5-11c6-4006-b0f8-f6cb41410b30/resource/57af4
   fb4-1141-41aa-a5f8-a92ccd5663c9/download/number-of-icu-beds-per-county.xlsx

4. Cases per county -- availed by
   https://www.health.go.ke/wp-content/uploads/2020/07/Kenya-COVID-19-SITREP-127-2
   2-Jul-2020.pdf

Data description

There are 4 data files consisting of information as follows:

Census 2019 Population County dataset --The dataset comprises of the following columns:

1. County_pcode
2. County_name
3. 2019_population……………...total population per county
4. Male
5. Female
6. Total households
7. Density(persons per sq km)

The data type mainly compose of object and float

Kenya health facilities ---The dataset comprises of the following columns:

1. Facility no
2. Facility name

3. HMIS

4. Province

5. District

6. Division

7. Location

8. County

9. Constituency

10. Sub-location

11. Spatial reference

12. Facility type

13. Agency

14. Geolocation

15. Facility type no

The data type mainly compose of object and integer

ICU beds per county---The dataset comprises of the following columns:

1. Cluster

2. County

3. Facility name

4. Unnamed column

5. Regular isolation beds available

6. Regular isolation beds recommended

7. Available ICU/critical care bed for isolation

8. Recommended ICU/ critical care bed for isolation

9. Ventilators available for isolation

10. Ventilators recommended for isolation

11. ICU beds for general patients(public and private)

The data type mainly compose of object and float

Cases per county(as at 23rd July)

The dataset comprises of the following columns:

1. S/no

2. County

3. Population(2019census)

4. Cumulative cases

5. Cases per 100,000 population

The data type mainly compose of object and float

Data quality

The main aim of this project is to rank counties in Kenya on COVID-19 preparedness. Important columns in each data file are among the useful information necessary for this project objective to be achieved .under the population per county we will need County_name, 2019_population, male, female only , under kenya health facility we will need, facility name, county , under icu bed per county We will drop just 2 columns in this dataset i.e. cluster and unnamed column and use the rest of the column in our analysis , under cases per county We drop just the S/no column

<div align="center">

STAGE 3 : Data preparation

</div>

Data cleaning

 This task involved raising the data quality to the level required by the analysis techniques that you've selected. The first actions we took to address data quality problems is

1. Drop all the irrelevant column

2. Under the county population dataset we are gonna drop 'County_PCODE', 'Total_households' , 'Density (Persons Per Sq. Km'un columns

3. Under health facilities dataset we gonna drop 'Facility Number', 'HMIS', 'Province', 'District', 'Division', 'LOCATION', 'Constituency', 'Sub-Location', 'Spatial Reference Method', 'Facility Type', 'Agency', 'Geolocation', 'Facility Type NAME' columns

4. Under the icu bed dataset we are gonna drop clusters and unnamed columns.

5. Under the county cases dataset we are gonna drop 'S/N o' column

6. Drop rows with the following in the county column 'Regional Subtotal','Regional Sub Total','Regional Sub-total'

7. The 4 data files were merged on a key attribute which is county.

STAGE 4 : Data analysis

This data will be analyzed using python and the libraries that will be used here will be pandas ,numpy ,seaborn and matplotlib.pyplot.

The notebook containing the analysis for this project as well as the data sets are accessible through the following link to the GitHub repository:

Analysis steps

1. Get difference between recommended isolation beds and bed available (,Create rank variable based on bed preparedness - a larger value signifies larger deficit in number of beds hence lower rank)

2. Calculate bed to population ratio

3. Rank counties based on hospital count,(a higher hospital count signifies better capacity handling hence higher rank)

4. Rank counties based on cumulative cases,(a higher number signifies a higher probability of infections hence lower rank in preparedness)

5. Rank counties based on ranking [number of people per sq km],(a higher number signifies a higher probability of infections hence lower rank in preparedness)

6. Rank counties based on bed to population ratio,(a higher ratio means fewer people are assigned to one bed hence better preparedness hence higher rank)

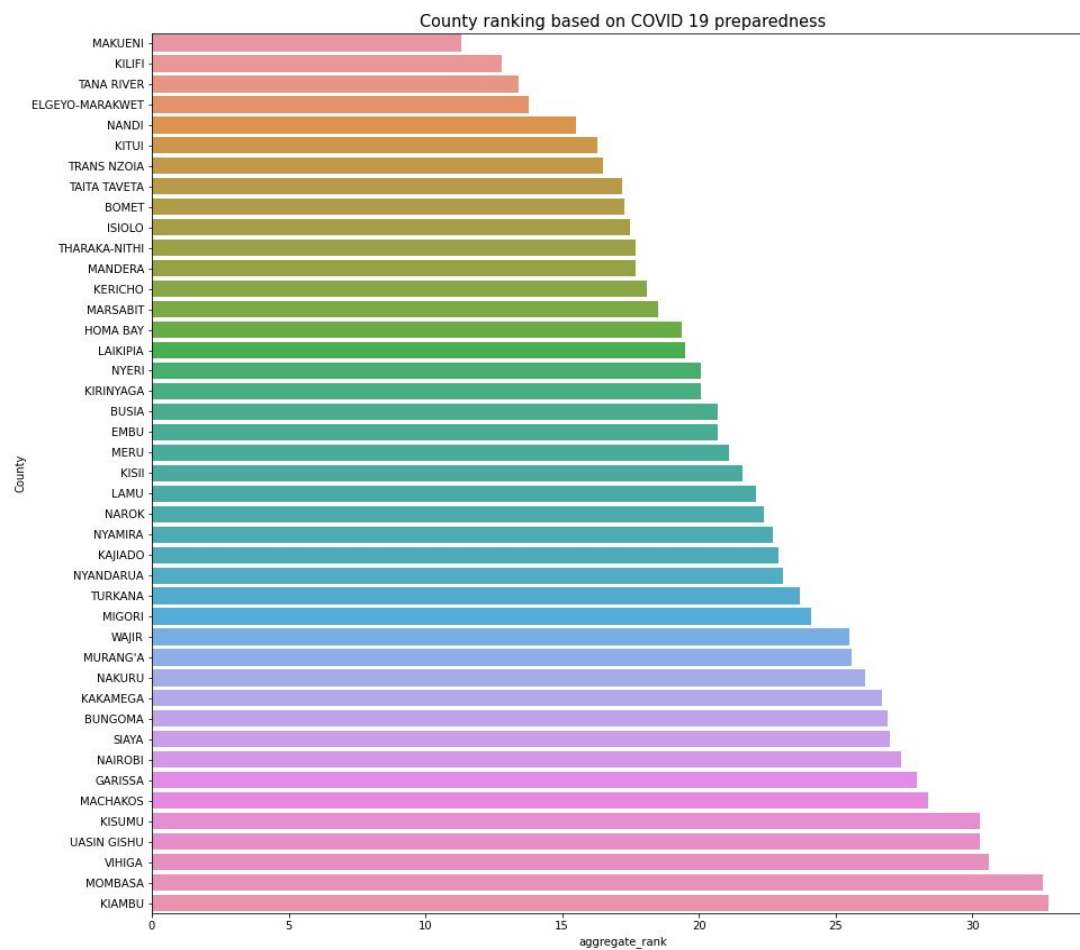7. Compute aggregate ranking of counties ,(Order counties by aggregate ranking)

Findings

County ranking based on Covid-19 Preparedness

County ranking based on COVID 19 preparedness

1.   The top 5 prepared counties

RANK	COUNTY
1	MAKUENI
2	KILIFI
3	TANA RIVER
4	ELGEYO-MARKET
5	NANDI

2.   The top 5 worst prepared counties

RANK	COUNTY
1	KIAMBU
2	MOMBASA
3	VIHIGA
4	UASIN GISHU
5	KISUMU

3.   VIHIGA County has the lowest bed to population ratio with a deficit of 34 beds

STAGE 5: Recommendation

The following are the recommendations made from the findings:

1. The least five prepared counties in terms of isolating beds to COVID 19 patients are:Kiambu,Kajiando,Busia,Mombasa and Nairobi they should consider allocating more beds capacity in preparation for the rise of COVID 19 cases.

2. The most prepared counties in relation to COVID 19 are:Makueni,Kilifi,Tana River,Elgeyo Marakwet and Nandi in that order.We arrived at this by looking at the bed to population ratio. These Counties stand a better chance of managing Covid 19 cases

3.   The least Country prepared to tackle the COVID 19 epidemic is VIHIGA as it has the lowest bed to population ratioThey should consider improving their preparedness by increasing COVID 19 isolation beds in line to their respective populations.

N/B The data used is dynamic as counties are putting more isolation beds to COVID 19 patients and the number of confirmed cases are rising dramatically across the country.

STAGE 6 : Evaluation

We have answered all posed business questions, now the data mining project is considered a success because the business objectives were met. The recommendations above provide better understanding based on   ranking counties in Kenya on COVID-19 preparedness.