**Healthcare System Readiness Assessment Report**

**COMS5027A - Health Analytics – Lab 4**

**Name:** Masixole Boya

**Student number:** 1869204

**Introduction**

Healthcare system readiness is a crucial factor in effectively managing and reducing the impact of pandemics and healthcare crises. To assess the readiness of healthcare systems across different regions, an in-depth analysis was conducted, with the use of a dataset that takes into account a number of important aspects of hospital bed resources.

This report presents the findings of a comprehensive analysis, to a certain extent, of healthcare system readiness in various regions of the United States of America. In particular, the focus is on their preparedness to handle healthcare crises such as the COVID-19 pandemic. The analysis makes use of extensive data on hospital bed resources and employs a clustering technique known as the k-means algorithm, to identify distinct patterns among regions. This report provides critical insights into the current state of healthcare readiness. The usefulness of the project is briefly inferred later in the report.

**Methodology**

The analysis employed a rigorous workflow that included data preprocessing, exploratory data analysis (EDA), clustering, and interpretation of results. Principal component analysis (PCA) was used to identify key factors contributing to healthcare system readiness. The k-Means clustering algorithm was employed to categorize groups based on their readiness levels.

**Findings and Inferences**

First Principal Component:

The first Principal Component (PC1) has the most significant influence on the data's variance, and it reflects notable separation among the clusters. This component likely represents a crucial factor related to healthcare system readiness. Also, it could be indicative of the overall capacity and preparedness of healthcare systems within different regions. Regions with higher values along PC1 may have well-equipped healthcare systems - sufficient hospital beds, and resources, making them better prepared to handle pandemics like COVID-19. Conversely, regions with lower PC1 values may have limited healthcare resources and may require additional support to enhance their readiness.

Second Principal Component:

While the second principal component (PC2) does not exhibit as much separation among the clusters, it still contributes to the overall data variance. PC2's influence may be associated with factors that contribute to healthcare system readiness but are not as important or critical as those captured by PC1. This component could represent additional attributes, such as the availability of specialized healthcare facilities or specific healthcare policies and practices. Regions with higher PC2 values may have certain advantages or characteristics that contribute to their readiness, even if their overall capacity (as captured by PC1) is not significantly high.

Inference:

I infer that PC1 is the total number of beds that would be available at the hospital. Then PC2 is the population of elderly people living in that region. The clusters then are the regions or characteristic, thereof. The reason for this thinking is that the primary aspect that the dataset is concerned about is readiness in terms of bed capacity for the hospitals. As such the majority of the attributes in the dataset are about bed capacity. The data goes on to look at other characteristics of the regions that I consider to be secondary or even tertiary factors as they are neither under the control of the hospitals nor are the affecting the day to day running of the hospital. Such features are the adult population of the region – this would be anyone above the age of 18 in South Africa, and the population above the age of 65 years.

I consider such attributes to be secondary because the decisions of equipping hospitals with more resources, extending the building or adding more facilities would depend on the local health governing body such as the Department of Health, sometimes at the request of hospital management. Such bodies rarely make these decisions based on the dynamic statistics of the local population. Or better put, whenever they (the Governing bodies) appear in public to reason why they decided to equip a hospital with more facilities, they never mention being led by the community. In South Africa, for example, hospitals are more likely to be built, extended, or equipped as a tribute to some black hero of the times of struggle (apartheid). I do acknowledge, however, that these community attributes may have a long term effect on the region’s capacity in that: if the majority of the population in a region is above 65 years of age, it is likely that hospital visits (per some arbitrary unit time) are much higher that a community with a youthful population (between the ages of 18 and 45). This is because elderly citizens are more likely to get sick as age catches up with them. This would then result in reduced potential capacity for the hospital in the region with older people.

Now if a new region (a datapoint not seen in the dataset) comes and fits cluster 2, I would categorize it as having more bed capacity and therefore is more ready to face a pandemic. If it fits in cluster one, I would categorize it as having less bed capacity and therefore, less ready to face a pandemic.

**So What?**

The findings from this analysis have significant implications for healthcare policymakers and authorities. By identifying clusters of regions with varying levels of healthcare readiness, this project enables targeted interventions and resource allocation. In the end, it empowers decision-makers to enhance healthcare system preparedness, ensuring better responses to future healthcare challenges and crises.

Top of Form