FINAL PORTFOLIO REPOART

Analysis of Healthcare System of India: A Multi-Problem Approach

1.0. Introduction

This report presents an extensive analysis of healthcare infrastructure across various states, utilizing a blend of datasets and analytical techniques. The primary goal of this project is to address disparities in healthcare access, evaluate utilization rates, and explore correlations between healthcare resources and outcomes. Through a systematic process of data cleaning, and visualization, the project aims to derive actionable insights and provide strategic recommendations for improving healthcare delivery.

2.0. Dataset Overview

The datasets used in this analysis include information on the number of healthcare facilities, the availability of hospital beds, and the distribution of healthcare professionals across different states. Additionally, demographic data and utilization rates of healthcare facilities were considered to provide a comprehensive view of the healthcare landscape.

3.0. Problem 1: Distribution of Healthcare Facilities

The first problem involved analysing the distribution of healthcare facilities across various states. The dataset provided information on the number of facilities categorized by type (e.g., primary healthcare centres, rural hospitals, urban hospitals). The analysis aimed to identify regions with a shortage of healthcare facilities and areas with an adequate or surplus of resources.

- **Data Cleaning and Wrangling:** The datasets were merged based on common columns such as state names. Missing values were handled by dropping rows with critical missing data, and columns were stripped of leading/trailing spaces to ensure consistency.
- **Insights:** The analysis revealed significant disparities in the distribution of healthcare facilities between rural and urban areas. Some states were identified as having a surplus of facilities in urban regions, while rural areas lagged.
- **Recommendations:** A reallocation of resources to underserved rural areas is recommended, along with the development of new facilities in high-need regions.

3.1. Problem 2: Availability of Hospital Beds

This problem focused on evaluating the availability of hospital beds across different regions. The goal was to assess whether the number of beds was proportional to the population served in each state.

- **Data Cleaning and Wrangling:** The data were aggregated to calculate the average population served per hospital bed, and any discrepancies in bed counts were resolved through data validation checks.

- **Insights:** The results indicated that some states had a significantly lower number of beds per capita, particularly in rural areas, leading to potential overcrowding in hospitals.
- **Recommendations:** Investment in hospital infrastructure is necessary for states with low bed availability, particularly in rural regions where healthcare access is already limited.

3.2. Problem 3: Healthcare Professional Distribution

In this problem, the focus shifted to analysing the distribution of healthcare professionals, including doctors, lab technicians, and pharmacists. The objective was to identify areas with a shortage of healthcare workers.

- **Data Cleaning and Wrangling:** The datasets were merged to calculate the total number of doctors per state, and any inconsistencies in the data were corrected.
- **Insights:** The analysis showed a disproportionate concentration of healthcare professionals in urban areas, leaving rural regions with inadequate medical support.
- **Recommendations:** Initiatives to incentivize healthcare professionals to work in rural areas are essential. This could include offering financial incentives, housing, and other benefits.

3.3. Problem 4: Correlation Between Facilities and Health Outcomes

This problem explored the relationship between the number of healthcare facilities and health outcomes, measured by a hypothetical health score.

- **Data Cleaning and Wrangling:** The datasets were combined, and a correlation analysis was performed to explore the relationship between healthcare infrastructure and health outcomes.
- **Insights:** A moderate positive correlation was found, indicating that regions with more healthcare facilities generally have better health outcomes.
- **Recommendations:** Strengthening healthcare infrastructure could lead to improved health outcomes, particularly in regions currently underserved.

3.4. Problem 5: Visualization of Healthcare Facilities on a Map

In this problem, the goal was to visualize the distribution of healthcare facilities across India using mapping tools.

- **Data Cleaning and Wrangling:** Latitude and longitude data, was unfortunately not usable, but states data together with total number of hospitals was plotted and a Geodata Frame was created for mapping.
- **Insights:** The visualization highlighted regions states with high number of healthcare facilities, as well as areas with few facilities.
- **Recommendations:** The map can serve as a tool for policymakers to identify and prioritize areas for healthcare development.

3.5. Problem 6: Facility Utilization Rates

This problem involved analysing the utilization rates of healthcare facilities to identify underused or overburdened resources.

- **Data Cleaning and Wrangling:** A hypothetical utilization dataset was created to assess the usage of facilities.
- **Insights:** The analysis identified several facilities with low utilization rates, suggesting potential inefficiencies in resource allocation.
- **Recommendations:** Resources should be reallocated to ensure that underused facilities are either repurposed or better integrated into the healthcare network.

3.6. Problem 7: Facility Type Distribution

This problem focused on analysing the types of healthcare facilities available in each region to identify any gaps in service offerings.

- **Data Cleaning and Wrangling:** The datasets were categorized by facility type, and a count of each type was performed.
- **Insights:** The results indicated a lack of specialized healthcare facilities in many rural areas.
- **Recommendations:** Investment in specialized healthcare services, such as community, maternity and paediatric care, is needed in underserved regions.

3.7. Problem 8: Correlation Between Utilization and Outcomes

This problem explored the relationship between facility utilization rates and health outcomes to determine if underused facilities negatively impact health outcomes.

- **Data Cleaning and Wrangling:** A correlation analysis was performed using the hypothetical utilization and outcome data.
- **Insights:** A weak correlation was found, suggesting that other factors, such as the quality of care, may play a more significant role in health outcomes.
- **Recommendations:** Further research is needed to explore the factors influencing health outcomes, beyond just facility utilization rates.

3.8. Problem 9: Strategic Recommendations

The final problem synthesized the insights gained from the previous analyses to provide strategic recommendations for improving healthcare infrastructure.

- **Data Cleaning and Wrangling:** The datasets were merged to provide a comprehensive view of healthcare infrastructure and outcomes.
- **Insights:** The cumulative analysis reinforced the need for targeted interventions in rural areas and a more equitable distribution of healthcare resources.

- **Recommendations:** Policymakers should focus on developing rural healthcare infrastructure, incentivizing healthcare professionals to work in underserved areas, and ensuring that resources are allocated efficiently to maximize health outcomes.

4.0. Conclusion

This report provides a detailed analysis of healthcare infrastructure, revealing significant disparities and areas for improvement. Through data-driven insights and strategic recommendations, this project aims to guide policymakers in making informed decisions to enhance healthcare access and quality across diverse regions. The use of Python for data processing and Power BI for visualization ensures that the findings are not only robust but also accessible for stakeholders to explore and act upon.