Final Project: Leveraging Analytics for the Los Angeles County DHS

Introduction

The Los Angeles County Department of Health Services (DHS) serves over 10 million residents and is tasked with overseeing a comprehensive healthcare system across a diverse population. As one of the largest health systems in the United States, the department faces many challenges of managing and analyzing a significant amount of data, including health records and business operations. To address those challenges, the Los Angeles County DHS has turned to advanced analytics to improve patient care, optimize resource allocation, and forecast future healthcare needs. This approach has been invaluable in helping the department comply with federal regulations and to make more informed policy decisions.

For this project, I will be exploring how the Los Angeles County DHS can use advanced analytics to improve patient care, optimize resource allocation, and forecast healthcare needs for its diverse population. I will be implementing a combination of analytics models such as *Regression Trees*, *Optimization Models*, and *Time Series Forecasting*. This proposal will outline my ideas on how these three models can be combined and implemented in the analysis, which will allow the DHS to enhance operational efficiency and provide more effective healthcare solutions to its diverse population.

Three different Analytics models that might be used to create the solution

1. **Regression Trees**

Regression Trees can identify patterns in patient data to predict outcomes while highlighting key variables. For example, it could be used to predict the patient's length of stay in the hospital based on demographics, medical history, and reasons for admissions. By analyzing those key variables, the DHS can understand the patient groups who are more likely to require extended care or additional resources. This can support effective patient prioritization and resource allocation to ensure timely care for those who need it the most.

2. **Optimization Models**

Optimization models could help efficiently allocate resources by balancing the predicted demand with operational constraints. These models could optimize staff schedules to minimize overtime costs while also ensuring sufficient coverage during peak demand periods. Additionally, optimization could be applied to hospital bed allocation across multiple facilities based on the forecasted admission rates, which will improve overall resource utilization in the hospital and ensures that healthcare services are provided where they are needed the most. More efficient allocation of resources achieved by the optimization models can also improve patient satisfaction by reducing waiting time in receiving care, e.g.

3. **Time Series Forecasting**

Time series forecasting analyzes historical trends to predict future outcomes, such as hospital admission rates or seasonal fluctuations in outpatient appointments. For example, it could be used to forecast hospital admission rates during specific seasons (such as fall) or by health crises, such as increased hospitalizations during the flu season or the COVID-19 pandemic. By predicting these trends, DHS can proactively adjust clinic schedules and allocate their resources more efficiently, reducing strain during high-demand periods and improve the overall service delivery.

Data that might be needed to use the models:

1. **Regression Trees**

The data required for a regression tree model includes:

- Patient Demographics: Age, gender, socioeconomic factors
- Clinical Data: Diagnoses, test results, treatment types
- Hospital Operational Data: Admission records, discharge data, staffing levels.

Helps in predicting outcomes such as length of stay based on patient and operational characteristics.

2. **Optimization Models**

The data required for optimization models includes:

- Historical and Forecasted Admission Data: facility and department, to understand demand.
- Resource constraints: staff availability, budgets, equipment capacity, other operational limits.
- Cost Data: Staffing costs, resource costs, financial constraints affecting resource allocation.

3. **Time Series Forecasting**

The data needed for time series forecasting includes:

- Historical Admission Data: facility, department, patient types.
- External Data: Weather trends, public health alerts, socioeconomic factors influencing patient admission rates.
- Seasonal and Event Data: historical patterns on health crises such as flu season, or COVID-19 pandemics, impacting hospital admissions.

Combining the three models

1. **Data Preprocessing and Integration**

First, centralize the data from different sources such as electronic health records, operational systems, and external sources (e.g., CDC information about the onset of the flu season or pandemic for each location; average social economic status of the patients' zipcode from the Census or the American Community Survey), with appropriate linking using unique patient identifiers; and then standardize and clean the datasets to ensure consistency and compatibility across different models. This ensures that the data is ready for analysis, which will improve the accuracy and reliability of the models.

2. **Demand Forecasting**

Use **Time Series Forecasting** to analyze historical trends and predict future admission rates and facility utilization, which helps to identify peaks and troughs in demand. This allows the DHS to anticipate periods of high or low resource needs and to prepare accordingly.

3. **Resource Allocation**

Put the forecasted demand into **Optimization Models** to create efficient resource allocation plans to optimize the allocation of staff, beds, and other hospital equipment to ensure they are available where and when they are needed the most. As new data becomes available, these plans can be adjusted to reflect the changes in patient demand and operational constraints.

4. **Predictive Analytics for Patient Care**

Apply **Regression Trees** to predict the patient's length of stay in the hospital, readmission risks, or the likelihood of needing extended care. Regression trees can identify factors that contribute to extended stays or readmissions, which will provide us with actionable insights to prioritize high-risk patients and to implement early interventions. This helps to improve patient outcomes while also managing hospital resources effectively.

How the specific data might be collected

1. **Data Sources**

Data can be collected from both internal and external sources:

- Internal Sources: admission logs, staffing schedules, financial records, patient demographics, diagnoses, test results, etc.

- External sources: public health records, socioeconomic data from Census or the American Community Survey (ACS), weather patterns, CDC information about flu seasons onsets, etc.

2. **Data Refresh Frequency**

- Real-Time Data: Admission and discharge data should be refreshed daily to provide updated insights for managing patient flow.
- Weekly: patient demographics and clinical data can be updated weekly to capture new patient information.
- Monthly updates: Operational and financial data, including staff schedules, bed occupancy rates, etc., can be refreshed monthly to reflect long-term trends to ensure operational efficiency.

3. **Challenges and Solutions**

- Develop seamless data pipeline that extracts, transforms, and loads data into a centralized database.
- Leverage cloud-based systems to scale the data to allow real-time data to be accessed across multiple departments, ensuring that the system can handle increased demand during peak periods, such as flu season and public health emergencies such as the COVID-19 pandemic.

How often the data might need to be refreshed

1. **Regression Trees**

Data needs to be refreshed weekly to ensure that patient demographics, clinical data, and hospital operational data are up to date. Given the dynamic nature of the variables, updating the data regularly are critical for maintain the accuracy of the predictions related to patient outcomes, such as length of stay, etc.

2. **Optimization models**

Data needs to be refreshed monthly or on-demand. Routine updates are needed to align with budget cycles and staffing schedules to ensure that the resources are efficiently allocated. However, for unexpected surges in hospital admissions due to seasonal flu or public health events such as the COVID-19 pandemic, additional updates are needed to adjust the allocation of resources to allow the hospitals to plan accordingly.

3. **Time Series Forecasting**

Data needs to be refreshed daily to capture real-time information such as daily admissions, discharge rates, and hospital utilization. This frequent update allows the forecasting models to respond quickly to emerging trends, and to improve short-term predictions, such as predicting daily admission rates during seasonal fluctuations or health crises.

Conclusion

By combining Regression Trees, Optimization Models, and Time Series Forecasting, the Los Angeles County DHS can build a robust analytics framework. This integrated approach will enhance decision-making across patient care, resource allocation, and operational efficiency. Regular updates to the data and model retraining will ensure that the system remains flexible and responsive to evolving healthcare demands and external factors. As a result, the DHS will be better equipped to deliver high-quality care and to optimize resources, leading to improved service delivery for its diverse community.

References

"Analytics helps major public health system run efficient programs and improve patient care." SAS, SAS Institute Inc., https://www.sas.com/en_us/customers/los-angeles-county-department-of-health-services.html. Accessed 15 Nov. 2024