**Forecasting the Future: Solving the Emergency Department Patient Prediction Puzzle**

In the heart of a bustling city, a dedicated team of healthcare professionals is on a mission to provide immediate care and support to patients in need. But they face an ongoing and relentless challenge - the unpredictable influx of patients into their Emergency Department (ED). This is not a tale of the past; it's a current and pressing problem that demands a solution.

## **Introduction: Addressing the ED's Unpredictable Challenge**

The Emergency Department (ED) is a crucial part of the healthcare system, where every second counts. Yet, the team faces an unpredictable and ever-changing stream of patients. The task at hand is to predict the number of patients who will arrive at the ED in the next seven days. The aim is to provide healthcare providers with the necessary tools to optimize resource allocation, staff management, and overall operational efficiency. This, in turn, enhances patient care and ensures the ED's ability to meet the growing demand for immediate medical attention.

## **Data Source: A Wealth of Information**

To tackle this ongoing challenge, we used the MIMIC-IV-ED database, a remarkable de-identified dataset of ED admissions from the Beth Israel Deaconess Medical Center, spanning from 2011 to 2019. This data treasure trove is not a thing of the past; it's a contemporary source of invaluable insights. It complies with the Health Information Portability and Accountability Act (HIPAA) Safe Harbor provisions, designed for educational initiatives and research studies.

**Data Description: Uncovering Key Insights**

We delved deep into the MIMIC-IV-ED database, mining crucial information from key tables:

1. **Edstays Table**: This table tracks patient stays in the ED, offering a comprehensive view of patient demographics, admission and discharge times, and their outcomes. This is a reflection of what's happening in the ED right now.

2. **Diagnosis Table:** Coded diagnoses using the International Classification of Diseases (ICD) system, providing critical information for billing and understanding patient health in real-time.

3. **Medrecon Table:** Detailed medication reconciliation for each patient, revealing what medications patients were on before their ED visit, providing critical insights for current care.

4. **Pyxis Table:** Insights into medications dispensed by the BD Pyxis MedStation, an automated medication dispensing system in the ED, showing the current state of medication use.

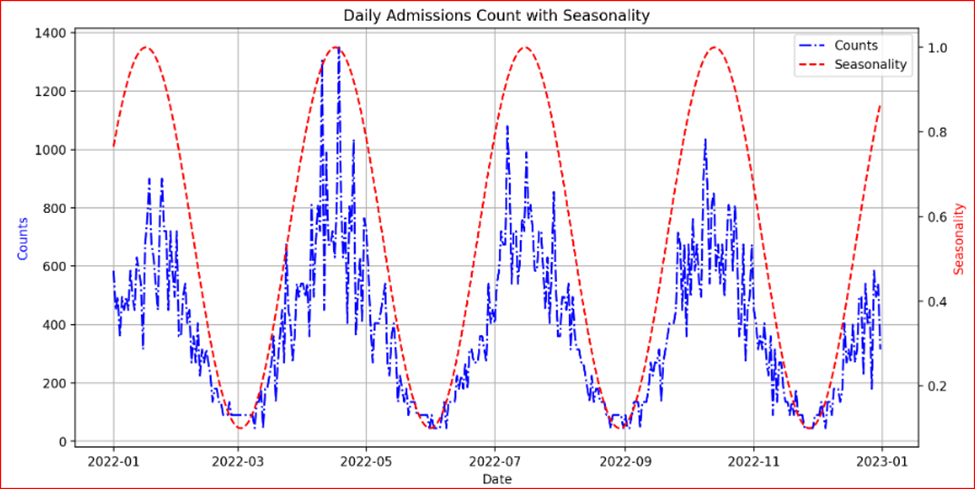
5. **Triage Table:** Records collected during the triage process, helping assess patient health and the reasons for their ED visit, offering insights into the current patient mix.

6. **Vitalsign Table:** A treasure trove of vital signs documented for patients during their ED stay, providing current health information.

In these tables, columns like admission date and time, patient ID, gender, mode of arrival, disposition, stay duration, and weekends/weekdays, were used for crafting the predictive model right now.

To distill meaningful insights, the team compressed a decade's worth of data into a single year, introducing seasonality and trend components. This transformation aimed to simulate real-world variations, ensuring that the predictive models would be robust and applicable to the dynamic nature of healthcare.

**Introduction of Seasonality to the Raw Scaled Data**



**RAW DATA EDA**

<https://github.com/abhishekk-fs/factihealth/blob/main/Version_1_1/ED/Codes/Report/report_edstays.html>

## 

## 

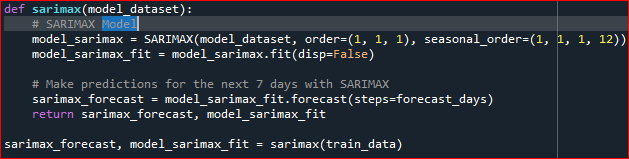
## 

## 

## **Modeling: Real-Time Insights**

Armed with current data, our healthcare professionals embarked on a journey into the world of predictive modeling, offering real-time solutions:

1**. SARIMAX Model:** A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) model was chosen for predicting patient counts. It's capable of capturing current time-dependent patterns in patient inflow, influenced by factors like the day of the week and seasonal variations.



2. **Prophet Model:** While its specific methodology remains shrouded in mystery, the Prophet model's flexibility in handling current time series data with seasonal and holiday effects makes it an invaluable partner in the quest for prediction.

## 

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **SARIMAX** | **Prophet** | **Preferred Model** |
| Mean Absolute Error | 112.16 | 143.39 | SARIMAX |
| Mean Squared Error | 18769.84 | 28795.71 | SARIMAX |
| Root Mean Squared Error | 137.00 | 169.69 | SARIMAX |
| Mean Absolute Percentage Error | 37.49% | 34.52% | Prophet |

### **SARIMAX vs. Prophet Performance Comparison:**

* + SARIMAX demonstrates superior performance over Prophet across key metrics, including mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE).
  + Prophet excels in mean absolute percentage error (MAPE).

### **Considerations for Model Selection:**

* + SARIMAX is favored for applications prioritizing lower absolute errors and minimizing deviations from observed values.
  + Prophet is advantageous when minimizing percentage errors is paramount, especially when this metric holds significant weight in decision-making.

### **Model Selection Insight:**

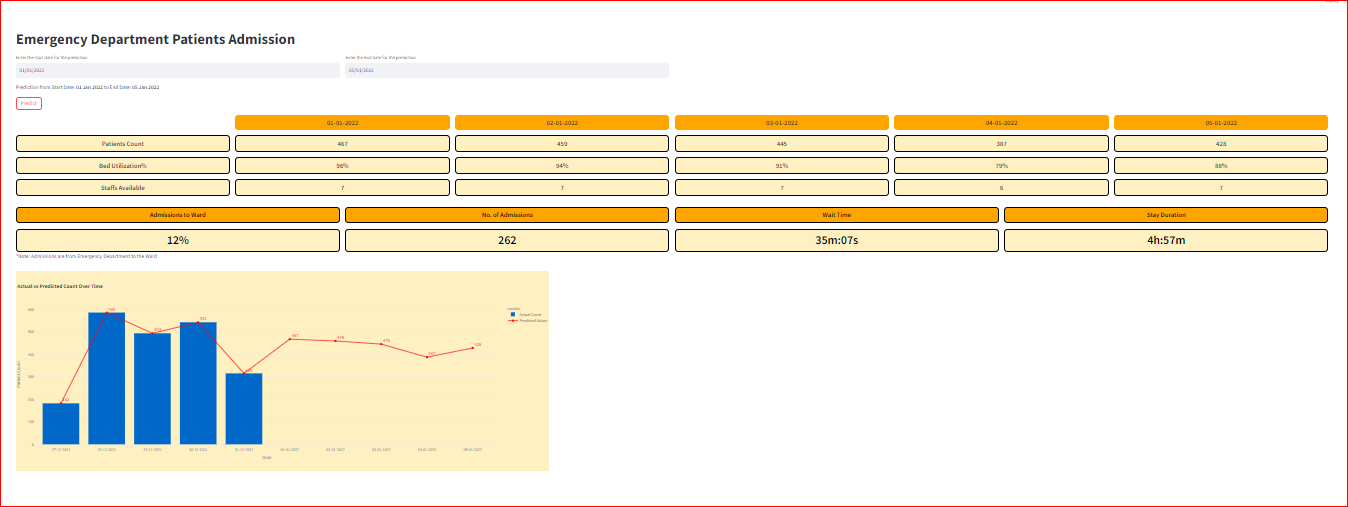
* + The choice between SARIMAX and Prophet hinges on the specific requirements of the application and the importance assigned to different error metrics within the given context. Opting for SARIMAX is prudent when minimizing absolute errors is critical, whereas Prophet may be the preferred choice when minimizing percentage errors is of utmost importance.

### **Current Implementation Decision:**

* + Presently, with the utilization of Scaled Seasonality Data, SARIMAX is the model of choice. This decision aligns with the objective of optimizing absolute error metrics, ensuring a tailored approach that suits the specific needs of the current application.

**The Power of Streamlit:**

To democratize access to these predictive models, the team turned to Streamlit, a powerful framework that transforms data scripts into shareable web applications. This decision not only made the models accessible to healthcare professionals but also facilitated seamless integration into existing hospital workflows.



**Key Performance Indicators: Measuring Success Now**

To navigate their journey and gauge success, our healthcare professionals established a set of Key Performance Indicators (KPIs):

1. **Patient Counts:** The predicted count of incoming patients to the ED, the core goal of this endeavor for the present moment.

2. **Stay Duration:** Calculated as the time between Discharge Time and Admission Time, a vital metric for efficient resource allocation and care right now.

3. **Bed Utilization:** A measure of bed efficiency, ensuring that the facility makes the most of its resources at this very moment. The formula is as follows:

Bed Utilization=(Average Patients×Average Length of Stay in Hours)(Total Beds×24)×100

4. **Staff Availability:** An estimate based on predicted patient-staff ratios and occupancy rates, critical for maintaining the highest standard of care today. Ratio is considered as 1:10(staffs to patients)

5. **Admission to Ward:** Indicates the number of patients transferred from the ED to the ward(Admission Rate)/total number of admissions from the ED.

6. **No. of Admissions**: Total no of admissions converted to the Ward from the Emergency Department.

7. **Wait Time:** It is the minimum time that the patient takes to get admitted in the Emergency Department. The base time is considered as 10 min followed by the Occupancy Rate on that Day. It is given by:

*wait\_time = base\_wait\_time + occupancy\_rate\*base\_wait\_time*

# **Files to Consider:**

1. MimicDatasetEDA.ipynb: EDA of all the mimic dataset.
2. DataSeasonalityTrendsIntroduction.ipynb: Sample data is generated with the help of this NB. Introduction of trends in the synthetic data generated.
3. DataSetExploreDailyScaled.ipynb: Since the data is distributed across 100 years. This dataset generates sample data considering the scaling factor.
4. host\_ed.py: This is a streamlit application to trigger the model.
5. Data, Model, Report: These folders consist of files, model and EDA report respectively. Data is downloaded from <https://physionet.org/content/mimiciv/2.2/>

# **Conclusion**:

The integration of advanced data science techniques with the intricacies of healthcare, as demonstrated in our patient prediction model for emergency department scenarios using the MIMIC-IV-ED dataset, holds profound implications for real-world applications. Our meticulous data transformation process and the strategic application of cutting-edge time series models, including PROPHET and SARIMAX, mark a pivotal advancement in predicting patient in the Emergency Department

These models are not mere intellectual exercises; they represent tangible tools that can revolutionize how hospitals operate. By accurately forecasting patient flow, healthcare institutions can proactively allocate resources, optimize staff deployment, and streamline emergency department workflows. The integration of Streamlit further democratizes access to these predictive insights, ensuring that healthcare professionals at all levels can leverage the power of data-driven decision-making.

In the real world, this translates to improved patient outcomes, reduced wait times, and enhanced overall healthcare delivery. Hospitals equipped with these predictive tools gain a strategic edge in managing patient influx, thereby fostering a more efficient and responsive healthcare system. As we navigate the evolving landscape of healthcare, these models stand not just as technological achievements but as beacons of progress, poised to make a meaningful impact on the frontline of patient care