# Modeling ICU Occupancy

A Project Report Submitted

In Fulfilment of the Requirements

For the Course of

**Internship** with **Celebal Technologies**

In

Fourth year – seventh Semester of

**Bachelor of Technology**

Specialization

In

Artificial Intelligence and Machine Learning

Under

**Ms. Poonam Kainthuria**

**Assistant Professor, Department of Informatics**

**School of Computer Science**

By

500068188 60 Naman Jain

500067512 61 Naman Jajoo

500067534 81 Saurav Verma

500065751 82 Shashank Panthari

# 

DEPARTMENT OF INFORMATICS

SCHOOL OF COMPUTER SCIENCE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES, BIDHOLI, DEHRADUN, UTTRAKHAND, INDIA

June-July 2021

1. **ABSTRACT**

Predicting the bed occupancy of intensive care unit (ICU) is a dynamic task. The uncertainty associated with critically ill patient and random arrival patient, severity of new patient leads to the bed capacity problem, and we need to take productive measure to improve the condition. In this project, we work toward creating predictive model using regression model and improving the prediction with minimizing loss.

1. **Introduction**

In the current era there is increase in monitoring patient’s details in digital form with the improvement in the technology we can further improve the hospital facilities by predicting ICU beds availability. Predicting free beds availability can be difficult task because of the unexpected arrival of the patients and dynamic reasons. The development of an automated model will assist physician in these matter and there will be a clear understanding of the beds availability.

In this section, we stimulate a mathematical mode to predict the ICU bed occupancy.

Even though the model can be applied to any unit, but we believe that it is reasonable to focus more on ICU beds availability as ICU is more critical since it is expensive and its important block of hospitals.

This section is divided into 2 sections in which first we introduce the concept of linear regression on which the model is based and second is to retrieve the data and process the hospitals records then formulating linear regression model to predict ICU occupancies.

**Linear regression**

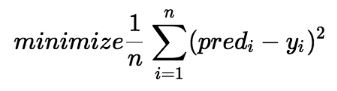
Linear regression is a ML algorithm user for supervised learning. Linear regression perform task to predict a dependent variable (Target) based on the given independent variable. So this regression technique is find out a linear relation between dependent and independent variable



**Hypothesis function for Linear Regression :**



**How to update θ1 and θ2 values to get the best fit line ?**



**Cost Function (J):**



1. **Problem Description**

In the study we address scheduling problem of beds, which hospitals faced during unavailability of beds However, in real-life hospitals, not all patients are of the same medical specialty and not all patients of the same medical specialty take the same path. Hence, we deal with a wide variety of different patient paths and different LOS in each unit.

# Therefore, this in particular is a mathematical model for simulating daily bed occupancy in an intensive care unit.

1. **Methodology**

In this section we will discuss how we acquire data set and how is data processed. Next, we proceed by introducing algorithm we used in our work.

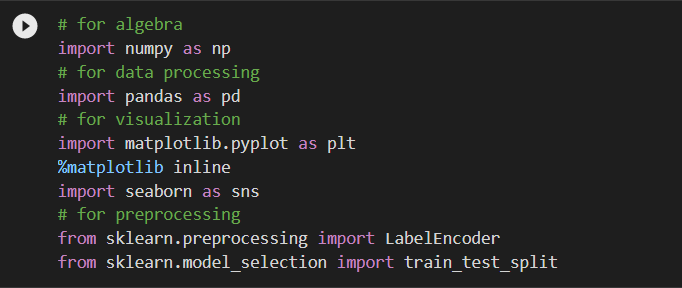
**Data Extraction**

The data concern all adults’ patients and we have nearly 4000 entries which comprises data from hospitals from multiple countries like Los Angeles, Contra Costa, Alameda, Fresno, Imperial, Inyo and more these records are from year 2005 to 2015.

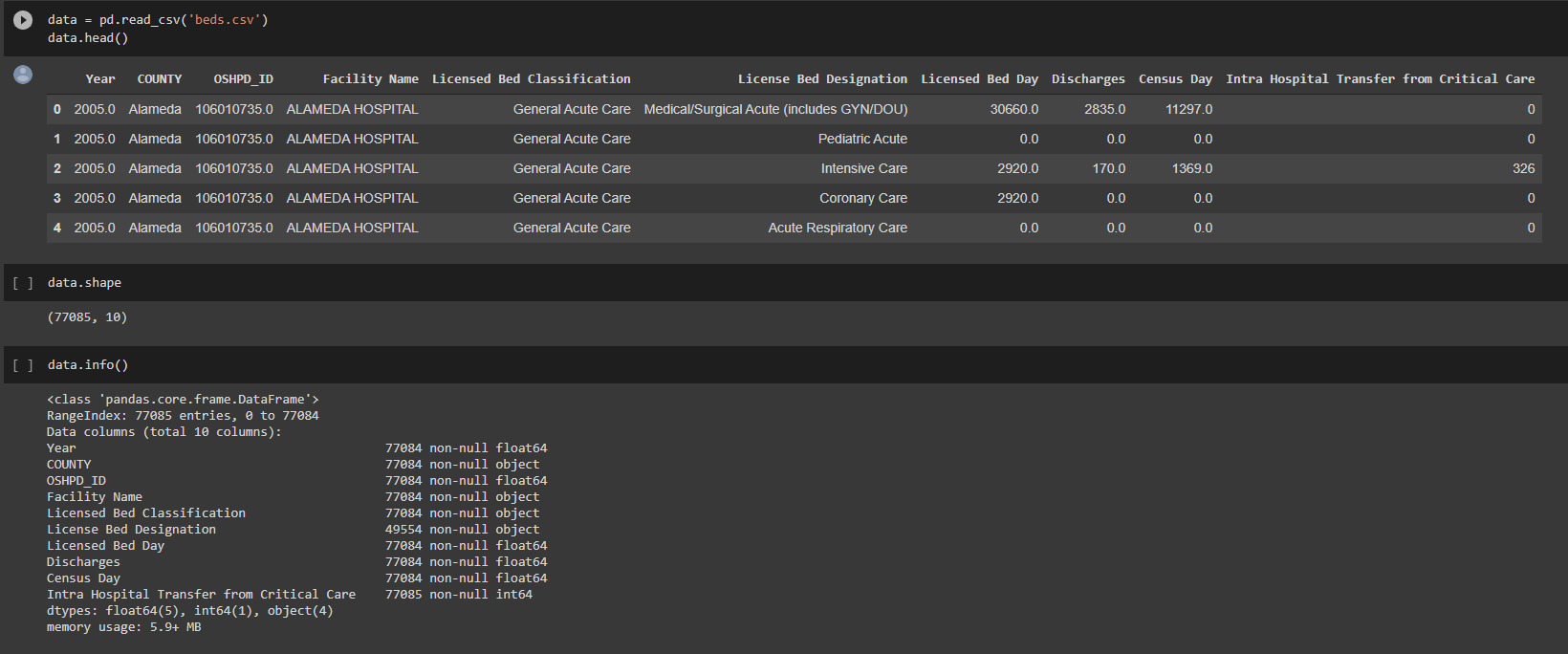
The data comprises of following features like, year, COUNTY, OSHPD\_ID, Facility Name, Licensed Bed Classification, License Bed Designation, Licensed Bed Day, Discharges, Census Day, Intra Hospital Transfer from Critical Care, available beds, County labels, Facility Label, Beds Required

### **Data pre-processing for extraction of features and labels**

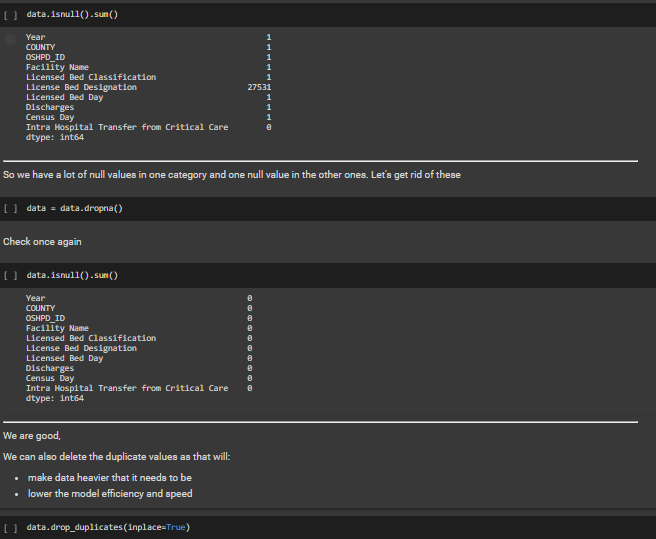
1. We will look for null values and we will manage the null values either by dropping complete row or by filling null values with mean as required.
2. Now we will remove columns, which have no effect on prediction and will remove those columns.
3. We will identify target class and visualization is performed as required.
4. Further, with this, we will be having a trend of patients and according to that after splitting data, we will further move to the implementation part of model.
5. At last, we will predict the availability of model with the help of our stimulation model.
6. **Implementation**
7. Loading important libraries.



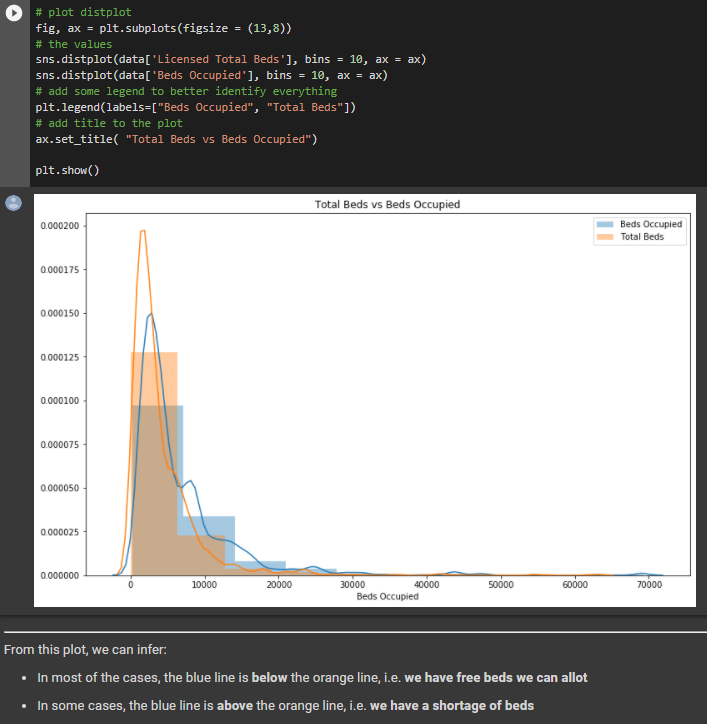
1. Loading and understanding data.



1. Working on null and duplicate values.

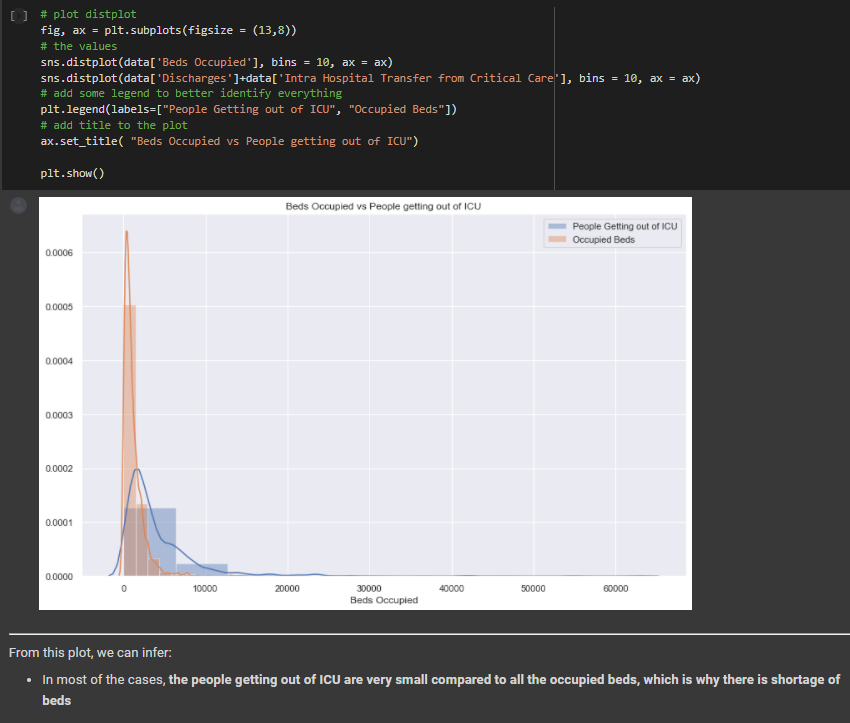


1. Visualisation.



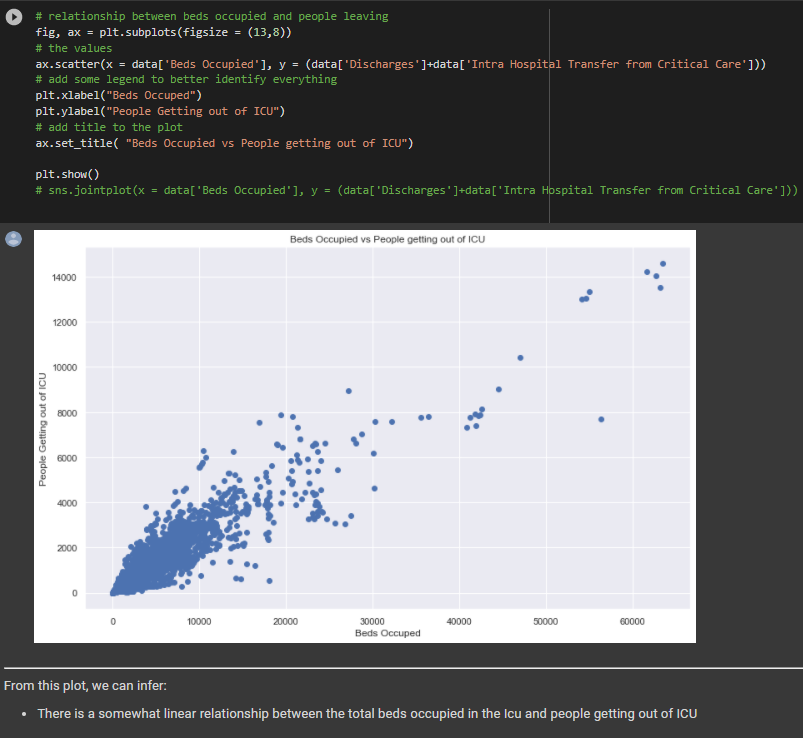
From this plot, we can infer:

* In most of the cases, the blue line is **below** the orange line, i.e. **we have free beds we can allot**
* In some cases, the blue line is **above** the orange line, i.e. **we have a shortage of beds**



From this plot, we can infer:

* In most of the cases, **the people getting out of ICU are very small compared to all the occupied beds, which is why there is shortage of beds**

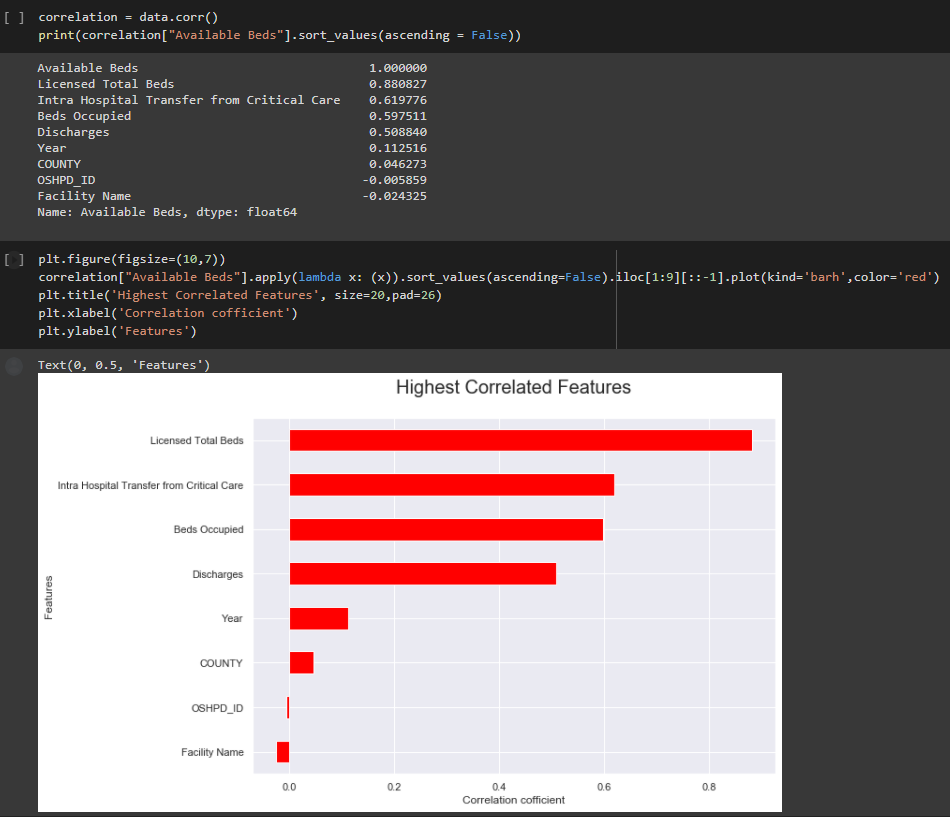


From this plot, we can infer:

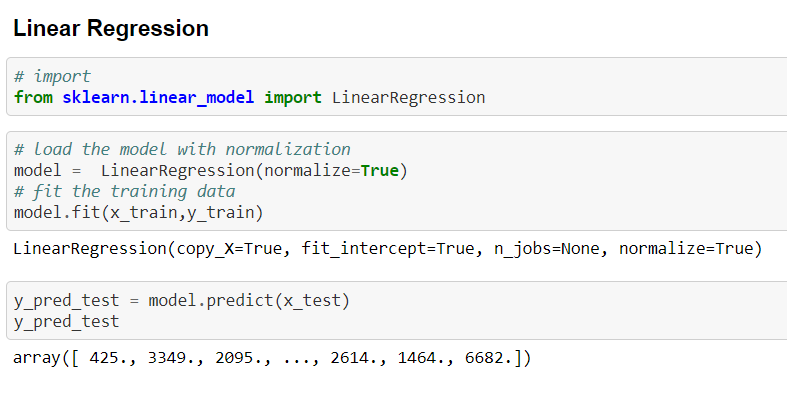
* There is a somewhat linear relationship between the total beds occupied in the ICU and people getting out of ICU

1. Making data training ready.

* **Step-1**: Get the object columns encoded so we can find out the correlation between them. Correlation is only calculated between continuous variables and not object type.
* **Step-2**: Get rid of excess/useless columns that do not help, and pick the best to start training. For this we can use correlation of these columns with the target variable to fish out the good attributes



1. Implementing Linear Regression.



1. **Conclusion**

We have predicted the number of Intensive care units present in the hospital using linear regression model. There by using simple mathematical operations we have estimated the number of beds available at the end of each day and trained our model based on that data for prediction of the number of beds that will be available in future.

1. **References**

* Barado, Julio, et al. “A Mathematical Model for Simulating Daily Bed Occupancy in an Intensive Care Unit.” Critical Care Medicine, vol. 40, no. 4, Apr. 2012, pp. 1098–104. PubMed, doi:10.1097/CCM.0b013e3182374828.
* Schiele, Julian, et al. “Predicting Intensive Care Unit Bed Occupancy for Integrated Operating Room Scheduling via Neural Networks.” Naval Research Logistics (NRL), vol. 68, no. 1, 2021, pp. 65–88. Wiley Online Library, doi:10.1002/nav.21929.
* McCabe, Ruth, et al. “Modelling Intensive Care Unit Capacity under Different Epidemiological Scenarios of the COVID-19 Pandemic in Three Western European Countries.” International Journal of Epidemiology, vol. 50, no. 3, June 2021, pp. 753–67. Silverchair, doi:10.1093/ije/dyab034.
* Ruyssinck, Joeri, et al. “Random Survival Forests for Predicting the Bed Occupancy in the Intensive Care Unit.” Computational and Mathematical Methods in Medicine, vol. 2016, Oct. 2016, p. e7087053. www.hindawi.com, doi:10.1155/2016/7087053.
* Baas, Stef, et al. “Real-Time Forecasting of COVID-19 Bed Occupancy in Wards and Intensive Care Units.” Health Care Management Science, vol. 24, no. 2, June 2021, pp. 402–19. Springer Link, doi: 10.1007/s10729-021-09553-5.