

## Sprint 4

The image displays two screenshots of a Jupyter Notebook interface, likely running on a local host (localhost:8888). The notebook is titled "Naalaiyathiran" and shows the following code and output:

**Cell 18:**

```
df.info()
```

**Output:**

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 137057 entries, 0 to 137056
Data columns (total 17 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   case_id                             137057 non-null  int64
 1   Hospital_code                       137057 non-null  int64
 2   Hospital_type_code                  137057 non-null  object
 3   City_Code_Hospital                 137057 non-null  int64
 4   Hospital_region_code               137057 non-null  object
 5   Available Extra Rooms in Hospital  137057 non-null  int64
 6   Department                         137057 non-null  object
 7   Ward_Type                          137057 non-null  object
 8   Ward_Facility_Code                 137057 non-null  object
 9   Bed_Grade                          137022 non-null  float64
10   patientid                          137057 non-null  int64
11   City_Code_Patient                  134900 non-null  float64
12   Type of Admission                  137057 non-null  object
13   Severity of Illness                 137057 non-null  object
14   Visitors with Patient               137057 non-null  int64
15   Age                                137057 non-null  object
16   Admission_Deposit                  137057 non-null  int64
dtypes: float64(2), int64(7), object(8)
memory usage: 17.8+ MB
```

**Cell 19:**

```
df = df.dropna()
```

**Cell 57:**

```
df.replace({'Type of Admission':{'Emergency':3,'Trauma':1,'Urgent':2}}, inplace = True)
df.replace({'Department':{'gynecology':3,'anesthesia':1,'radiotherapy':2,'TB & Chest disease':4,'surgery':5}}, inplace = True)
```

**Cell 64:**

```
df.replace({'Severity of Illness':{'Extreme':3,'Moderate':2,'Minor':1}}, inplace = True)
y = df['Severity of Illness']
```

**Cell 65:**

```
x = df[['Type of Admission', 'Department']]
```

**Cell 66:**

```
from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(x,y,test_size = 0.3)
```

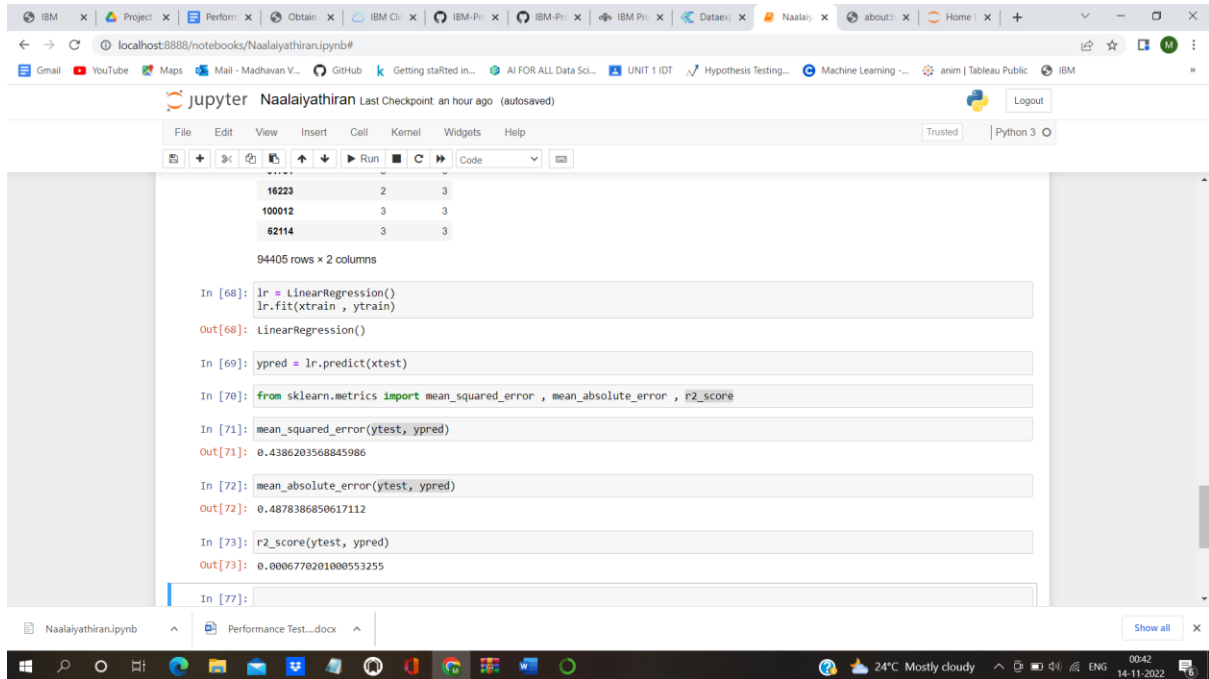
**Cell 67:**

```
from sklearn.linear_model import LinearRegression
xtrain
```

**Output:**

	Type of Admission	Department
44116	4	3
96276	3	3
68106	3	3
3965	4	3
82842	3	3
...	...	...
24393	3	3

## Sprint 4



```
16223    2    3
100012    3    3
62114    3    3

94405 rows x 2 columns

In [68]: lr = LinearRegression()
lr.fit(xtrain, ytrain)

Out[68]: LinearRegression()

In [69]: ypred = lr.predict(xtest)

In [70]: from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score

In [71]: mean_squared_error(ytest, ypred)

Out[71]: 0.4386203568845986

In [72]: mean_absolute_error(ytest, ypred)

Out[72]: 0.4878386850617112

In [73]: r2_score(ytest, ypred)

Out[73]: 0.0006770201000553255

In [77]:
```

A prediction model was developed using Linear Regression technique for predicting severity of illness based on inputs