

# ANALYTICS FOR HOSPITAL AND HEALTHCARE DATA

TEAM ID : PNT2022TMID16345

## DATA ANALYSIS

```
Home Page - Select or create a notebook | final code ibm - Jupyter Notebook | Hospital Stay Prediction - EDA analysis | +
```

localhost:8888/notebooks/final%20code%20ibm.ipynb#

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)

```
In [27]: 1 df_num_test = test.select_dtypes([np.number])
2 df_num_test.head()
```

```
Out[27]:
```

|   | Visitors with Patient | Admission_Deposit |
|---|-----------------------|-------------------|
| 0 | 2                     | 3095.000000       |
| 1 | 4                     | 4018.000000       |
| 2 | 3                     | 4492.000000       |
| 3 | 3                     | 4173.000000       |
| 4 | 4                     | 4161.000000       |

```
In [28]: 1 df_cat_test = test.select_dtypes([np.object])
2 df_cat_test.head()
```

```
Out[28]:
```

|   | Hospital_code | Hospital_type_code | City_Code_Hospital | Hospital_region_code | Available<br>Extra<br>Rooms<br>in<br>Hospital | Department | Ward_Type | Ward_Facility_Code | Bed<br>Grade | City_Code_Patient |
|---|---------------|--------------------|--------------------|----------------------|---|------------|-----------|--------------------|--------------|-------------------|
| 0 | 21            | c                  | 3                  | Z                    | 3   | gynecology | S         | A                  | None         | No                |
| 1 | 29            | a                  | 4                  | X                    | 2   | gynecology | S         | F                  | None         | No                |
| 2 | 26            | b                  | 2                  | Y                    | 3   | gynecology | Q         | D                  | None         | No                |
| 3 | 6             | a                  | 6                  | X                    | 3   | gynecology | Q         | F                  | None         | No                |
| 4 | 28            | b                  | 11                 | X                    | 2   | gynecology | R         | F                  | None         | No                |

```
In [29]: 1 admission_encode = {'Trauma': 1, 'Urgent': 2, 'Emergency': 3 }
```

```
Home Page - Select or create a notebook | final code ibm - Jupyter Notebook | Hospital Stay Prediction - EDA analysis | +
```

localhost:8888/notebooks/final%20code%20ibm.ipynb#

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)

```
In [29]: 1 admission_encode = {'Trauma': 1, 'Urgent': 2, 'Emergency': 3 }
2 severity_encode = {'Minor': 1, 'Moderate': 2, 'Extreme': 3 }
3
4
5 df_cat_train['Type of Admission'] = df_cat_train['Type of Admission'].map(admission_encode)
6 df_cat_train['Severity of Illness'] = df_cat_train['Severity of Illness'].map(severity_encode)
7
8 df_cat_test['Type of Admission'] = df_cat_test['Type of Admission'].map(admission_encode)
9 df_cat_test['Severity of Illness'] = df_cat_test['Severity of Illness'].map(severity_encode)
```

```
In [30]: 1 df_cat_train['Stay'] = df_cat_train['Stay'].replace({'0-10':1, '11-20':2, '21-30':3, '31-40':4, '41-50':5, '51-60':6, '61-70':7,
2 '71-80':8, '81-90':9, '91-100':10, '100+':11})
3
4 df_cat_train['Age'] = df_cat_train['Age'].replace({'0-10':1, '11-20':2, '21-30':3, '31-40':4, '41-50':5, '51-60':6, '61-70':7,
5 '71-80':8, '81-90':9, '91-100':10})
6
7 df_cat_test['Age'] = df_cat_test['Age'].replace({'0-10':1, '11-20':2, '21-30':3, '31-40':4, '41-50':5, '51-60':6, '61-70':7,
8 '71-80':8, '81-90':9, '91-100':10})
9
```

```
In [31]: 1 df_cat_train['Stay'] = df_cat_train['Stay'].astype(int)
```

```
In [32]: 1 from sklearn.preprocessing import LabelEncoder
2 LE = LabelEncoder()
3
4 df_cat_train['Hospital_code'] = LE.fit_transform(df_cat_train['Hospital_code'])
5 df_cat_train['Hospital_type_code'] = LE.fit_transform(df_cat_train['Hospital_type_code'])
6 df_cat_train['City_Code_Hospital'] = LE.fit_transform(df_cat_train['City_Code_Hospital'])
7 df_cat_train['Hospital_region_code'] = LE.fit_transform(df_cat_train['Hospital_region_code'])
8 df_cat_train['Department'] = LE.fit_transform(df_cat_train['Department'])
9 df_cat_train['Ward_Type'] = LE.fit_transform(df_cat_train['Ward_Type'])
10 df_cat_train['Ward_Facility_Code'] = LE.fit_transform(df_cat_train['Ward_Facility_Code'])
11 df_cat_train['City_Code_Patient'] = LE.fit_transform(df_cat_train['City_Code_Patient'])
12 df_cat_train['Bed Grade'] = LE.fit_transform(df_cat_train['Bed Grade'])
13
```

Home Page - Select or create a notebook | final code ibm - Jupyter Notebook | Hospital Stay Prediction - EDA |

localhost:8888/notebooks/final%20code%20ibm.ipynb#

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)

```
14 df_cat_train.head()
```

Out[32]:

|   | Hospital_code | Hospital_type_code | City_Code_Hospital | Hospital_region_code | Available<br>Extra<br>Rooms<br>in<br>Hospital | Department | Ward_Type | Ward_Facility_Code | Bed<br>Grade | City_Code_Patie |
|---|---------------|--------------------|--------------------|----------------------|---|------------|-----------|--------------------|--------------|-----------------|
| 0 | 7             | 2                  | 2                  | 2                    | 3   | 3          | 2         | 5                  | 1            |                 |
| 1 | 1             | 2                  | 4                  | 2                    | 2   | 3          | 3         | 5                  | 1            |                 |
| 2 | 9             | 4                  | 0                  | 0                    | 2   | 1          | 3         | 4                  | 1            |                 |
| 3 | 25            | 1                  | 1                  | 1                    | 2   | 3          | 2         | 3                  | 1            |                 |
| 4 | 25            | 1                  | 1                  | 1                    | 2   | 3          | 3         | 3                  | 1            |                 |

```
In [33]: 1 from sklearn.preprocessing import LabelEncoder
2 LE=LabelEncoder()
3
4 df_cat_test['Hospital_code']=LE.fit_transform(df_cat_test['Hospital_code'])
5 df_cat_test['Hospital_type_code']=LE.fit_transform(df_cat_test['Hospital_type_code'])
6 df_cat_test['City_Code_Hospital']=LE.fit_transform(df_cat_test['City_Code_Hospital'])
7 df_cat_test['Hospital_region_code']=LE.fit_transform(df_cat_test['Hospital_region_code'])
8 df_cat_test['Department']=LE.fit_transform(df_cat_test['Department'])
9 df_cat_test['Ward_Type']=LE.fit_transform(df_cat_test['Ward_Type'])
10 df_cat_test['Ward_Facility_Code']=LE.fit_transform(df_cat_test['Ward_Facility_Code'])
11 df_cat_test['City_Code_Patient']=LE.fit_transform(df_cat_test['City_Code_Patient'])
12 df_cat_test['Bed Grade']=LE.fit_transform(df_cat_test['Bed Grade'])
13
14 df_cat_test.head()
```

Out[33]:

|  | Hospital_code | Hospital_type_code | City_Code_Hospital | Hospital_region_code | Available<br>Extra<br>Rooms<br>in<br>Hospital | Department | Ward_Type | Ward_Facility_Code | Bed<br>Grade | City_Code_Patie |
|--|---------------|--------------------|--------------------|----------------------|---|------------|-----------|--------------------|--------------|-----------------|
|--|---------------|--------------------|--------------------|----------------------|---|------------|-----------|--------------------|--------------|-----------------|

Home Page - Select or create a notebook | final code ibm - Jupyter Notebook | Hospital Stay Prediction - EDA |

localhost:8888/notebooks/final%20code%20ibm.ipynb#

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)

```
In [34]: 1 from sklearn.preprocessing import StandardScaler
2 sc = StandardScaler()
3
4 num_scaled = sc.fit_transform(df_num_train)
5
6 df_num_scaled = pd.DataFrame(num_scaled, columns = df_num_train.columns)
```

```
In [35]: 1
2 num_scaled_test = sc.fit_transform(df_num_test)
3
4 df_num_scaled_test = pd.DataFrame(num_scaled_test, columns = df_num_test.columns)
```

```
In [36]: 1 df_num_scaled.shape
```

Out[36]: (313793, 2)

```
In [37]: 1 df_cat_train = df_cat_train.reset_index(drop=True)
2 df_num_scaled = df_num_scaled.reset_index(drop=True)
3 df_cat_test = df_cat_test.reset_index(drop=True)
4 df_num_scaled_test = df_num_scaled_test.reset_index(drop=True)
```

```
In [38]: 1 df_cat_train.shape
```

Out[38]: (313793, 14)

```
In [39]: 1 df_full = pd.concat([df_num_scaled, df_cat_train],axis=1)
2 df_full_test = pd.concat([df_num_scaled_test, df_cat_test],axis=1)
```

```
In [40]: 1 df_full.shape
```

Out[40]: (313793, 16)

```
In [41]: 1 df_full.head()
```

Home Page - Select or create a notebook | final code ibm - Jupyter Notebook | Hospital Stay Prediction - EDA |

localhost:8888/notebooks/final%20code%20ibm.ipynb#

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 (ipykernel)

```
In [37]: 1 df_cat_train = df_cat_train.reset_index(drop=True)
2 df_num_scaled = df_num_scaled.reset_index(drop=True)
3 df_cat_test = df_cat_test.reset_index(drop=True)
4 df_num_scaled_test = df_num_scaled_test.reset_index(drop=True)
```

```
In [38]: 1 df_cat_train.shape
```

Out[38]: (313793, 14)

```
In [39]: 1 df_full = pd.concat([df_num_scaled, df_cat_train],axis=1)
2 df_full_test = pd.concat([df_num_scaled_test, df_cat_test],axis=1)
```

```
In [40]: 1 df_full.shape
```

Out[40]: (313793, 16)

```
In [41]: 1 df_full.head()
```

Out[41]:

| Visitors<br>with<br>Patient | Admission_Deposit | Hospital_code | Hospital_type_code | City_Code_Hospital | Hospital_region_code | Available<br>Extra<br>Rooms<br>in<br>Hospital | Department | Ward_Type | Ward_Facility |
|-----------------------------|-------------------|---------------|--------------------|--------------------|----------------------|---|------------|-----------|---------------|
|-----------------------------|-------------------|---------------|--------------------|--------------------|----------------------|---|------------|-----------|---------------|

