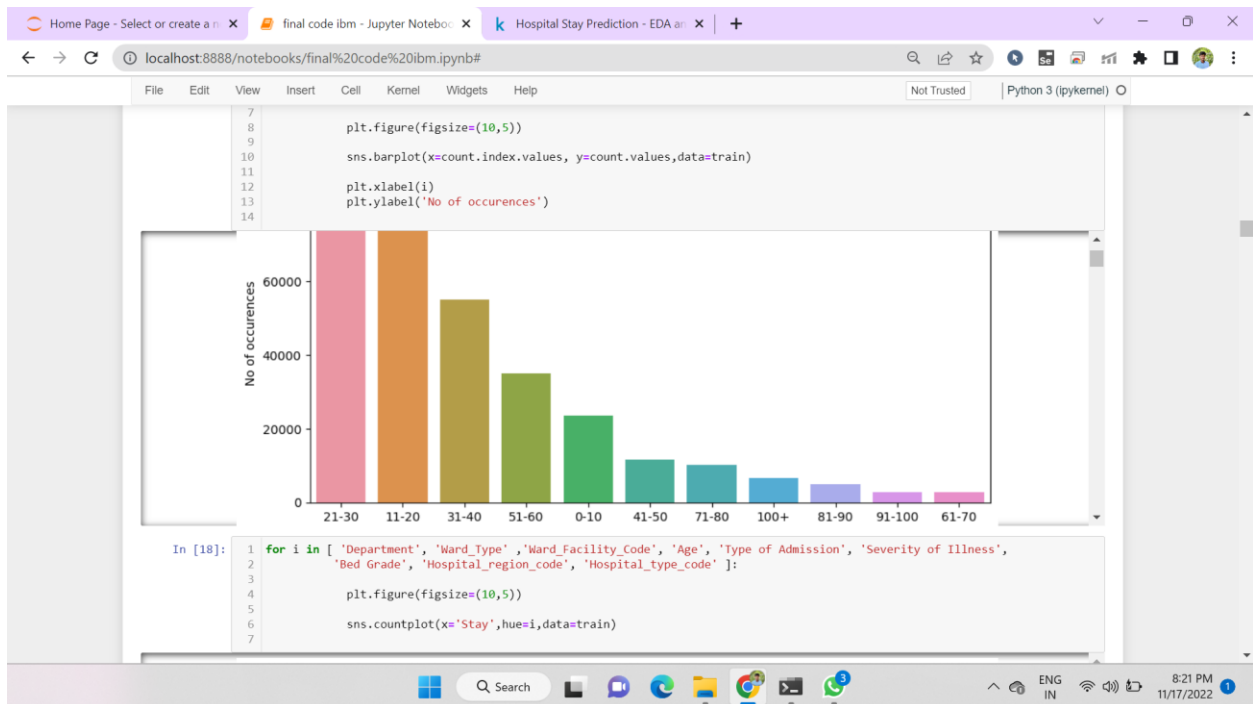
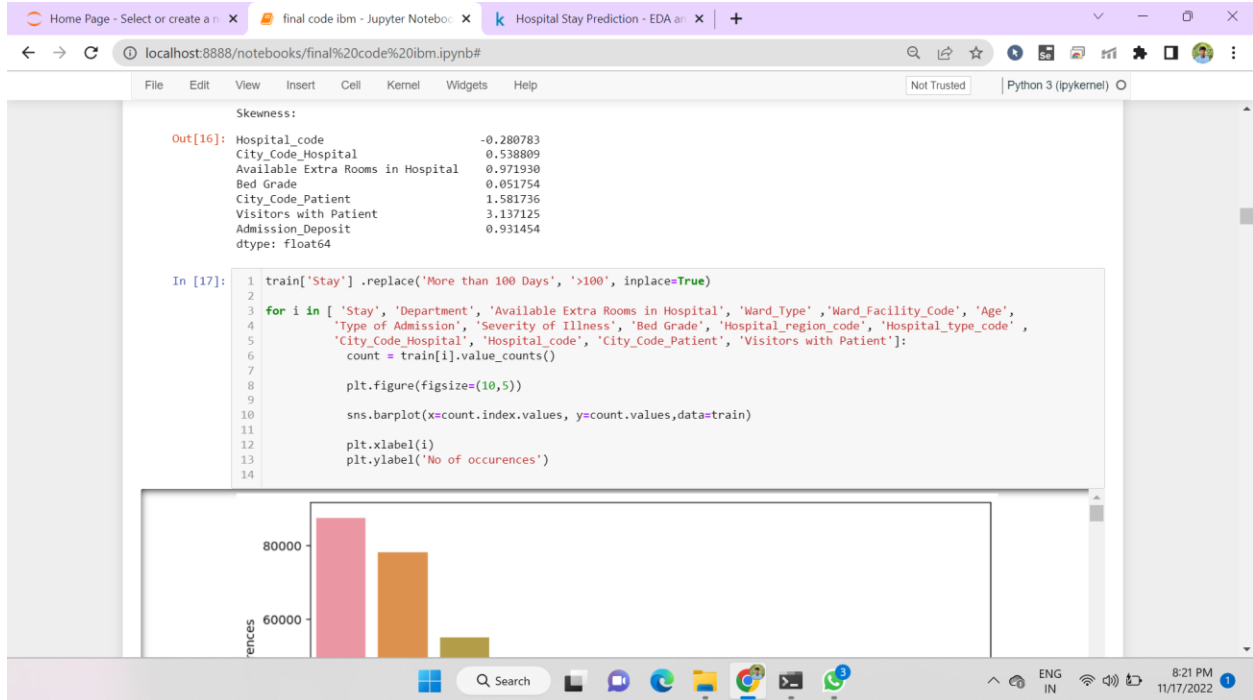


# Analytics for Hospitals Health-Care Data

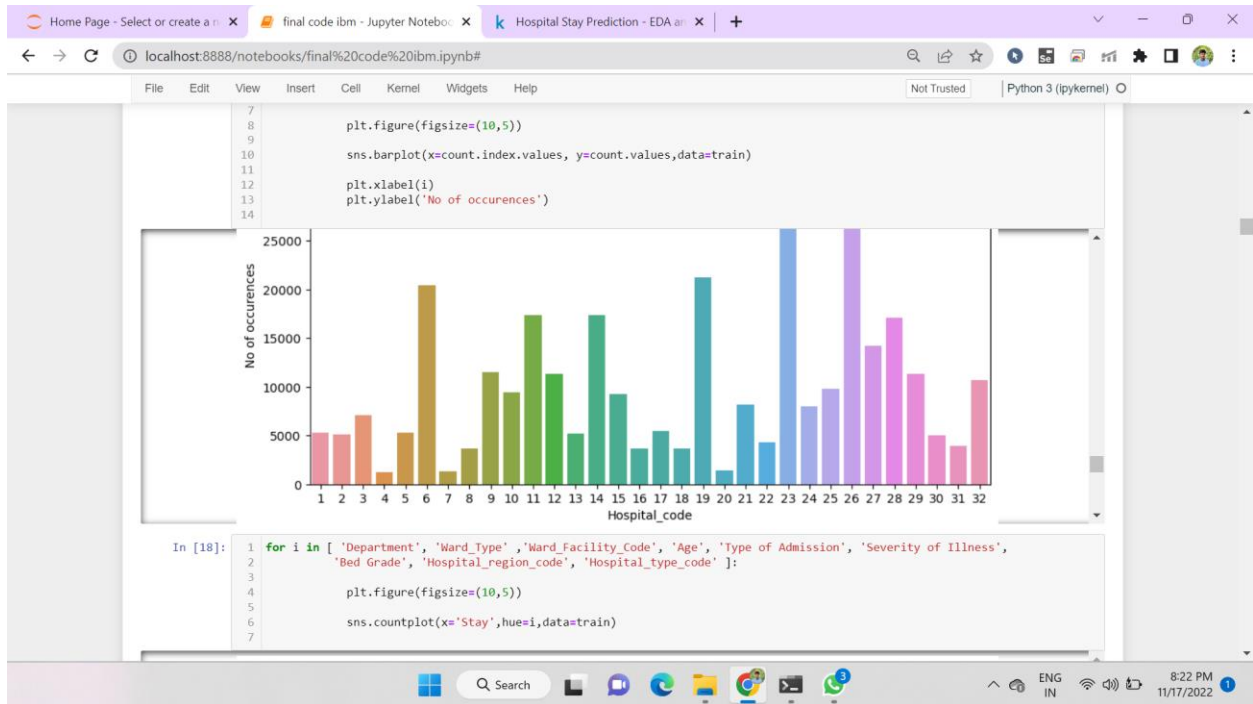
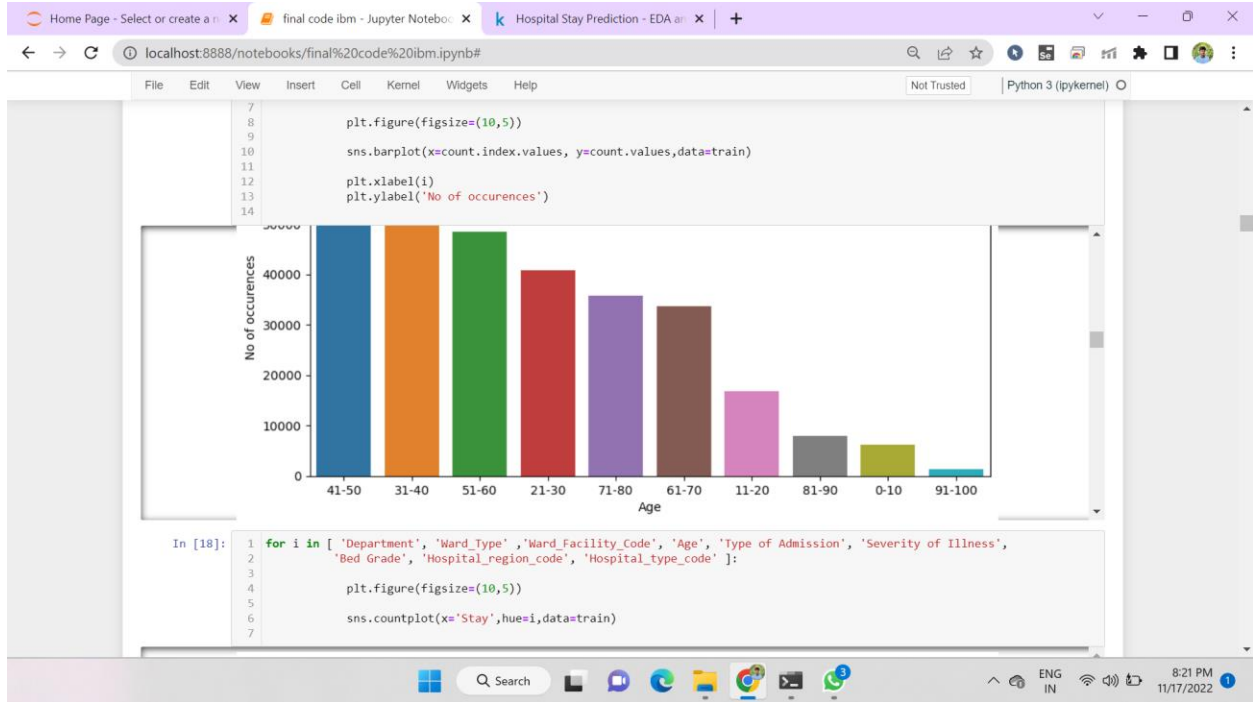
TEAM ID: PNT2022TMID16292

## EDA AND FLOW CHART



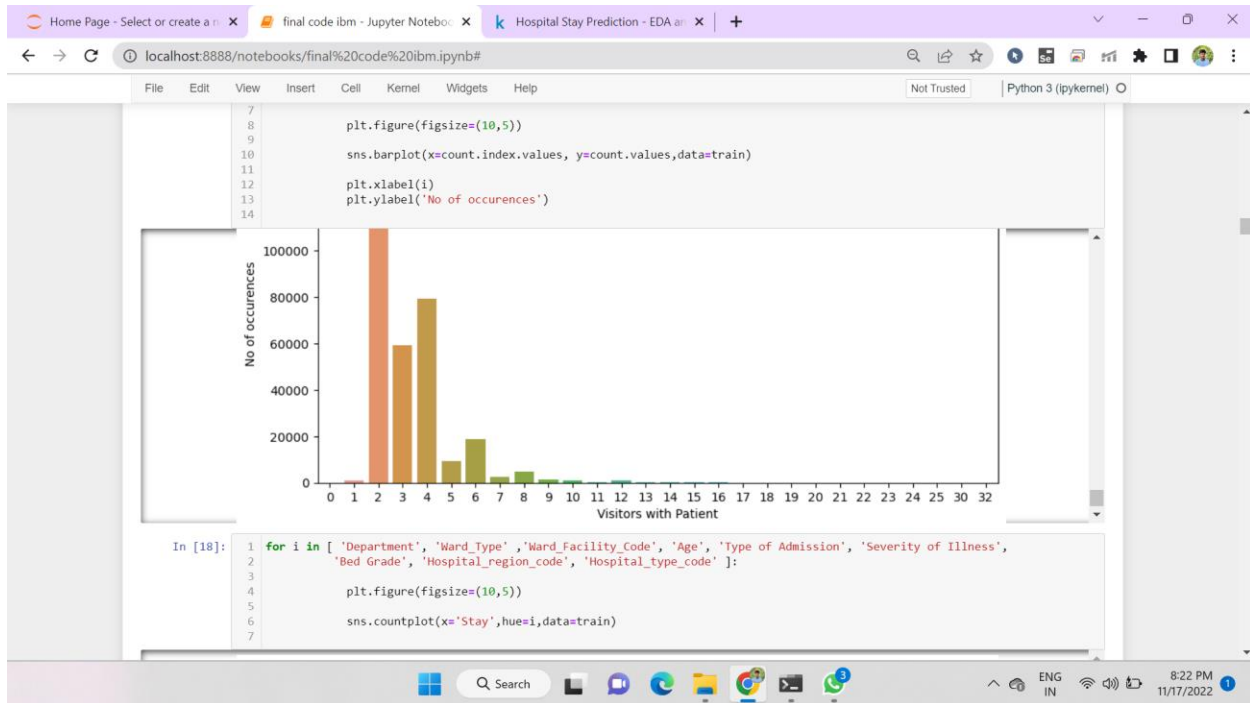
# Analytics for Hospitals Health-Care Data

TEAM ID: PNT2022TMID16292

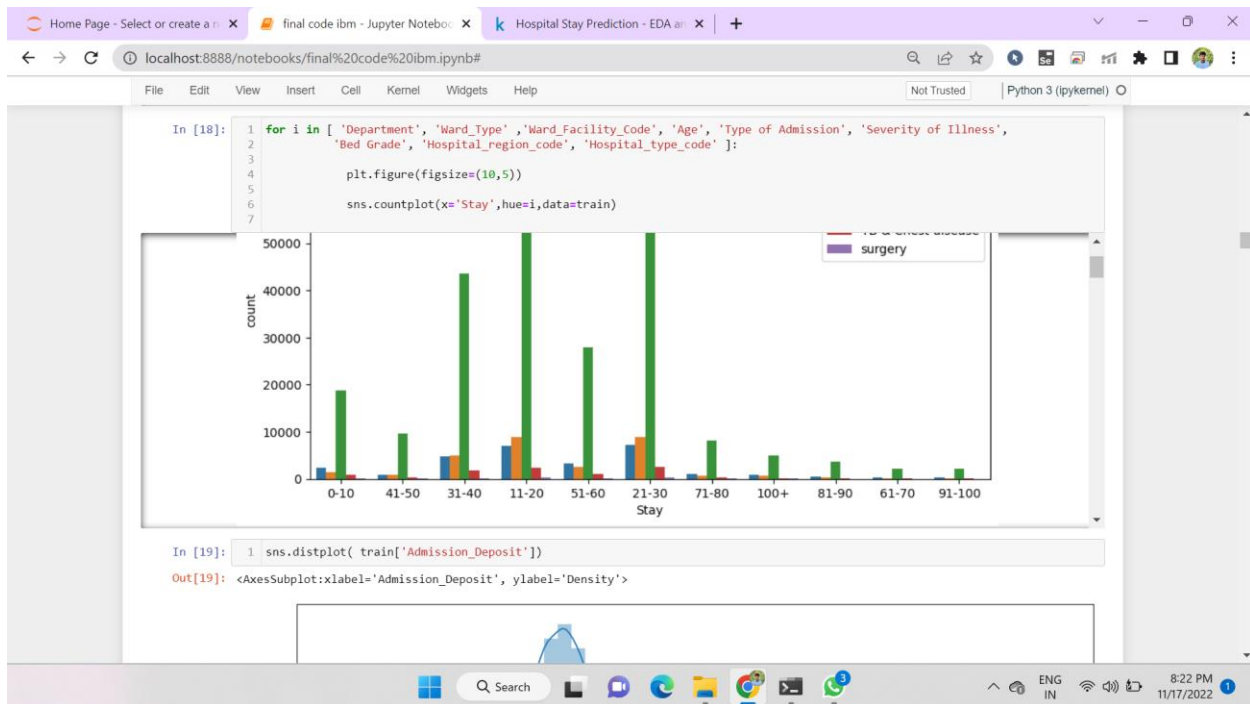


# Analytics for Hospitals Health-Care Data

TEAM ID: PNT2022TMID16292

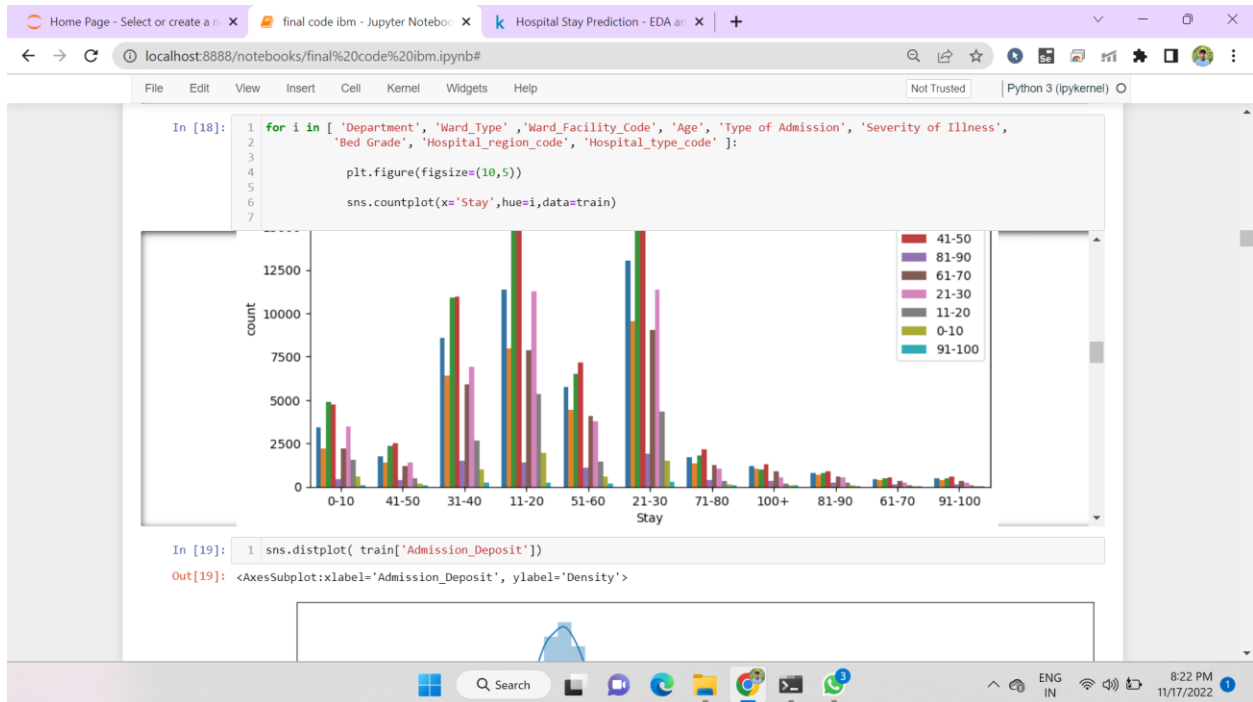
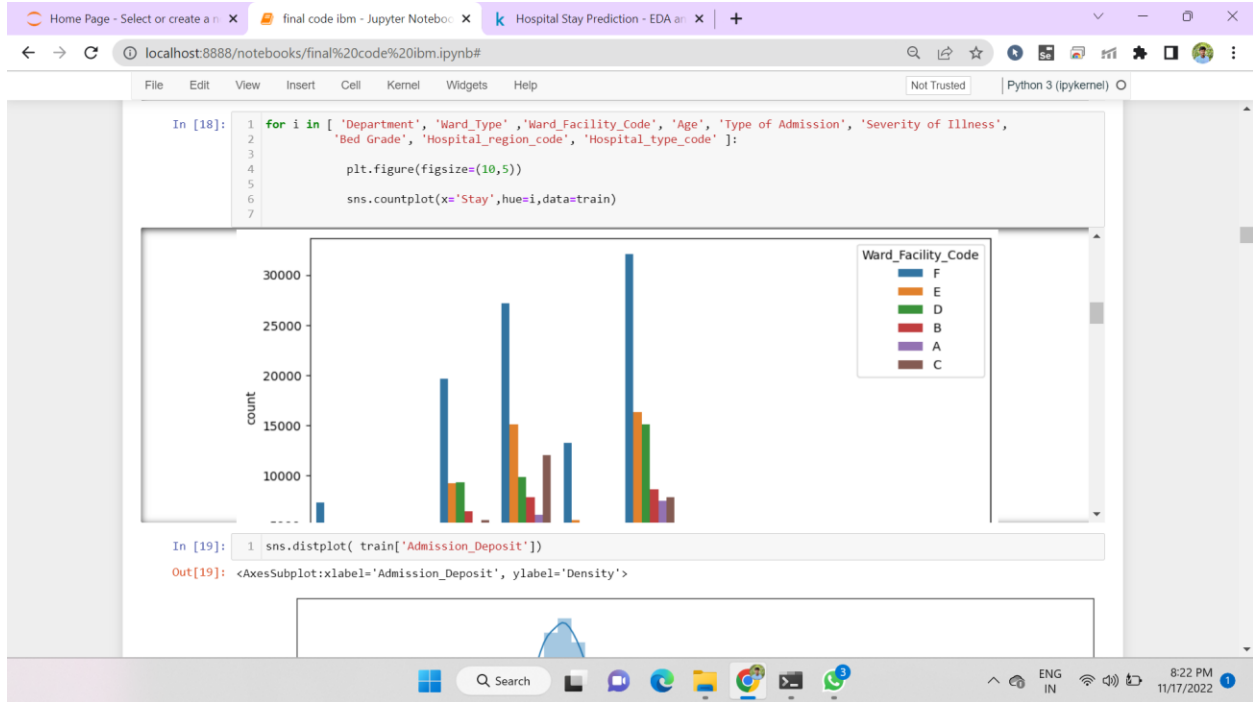


## BIVARIATE ANALYSIS



# Analytics for Hospitals Health-Care Data

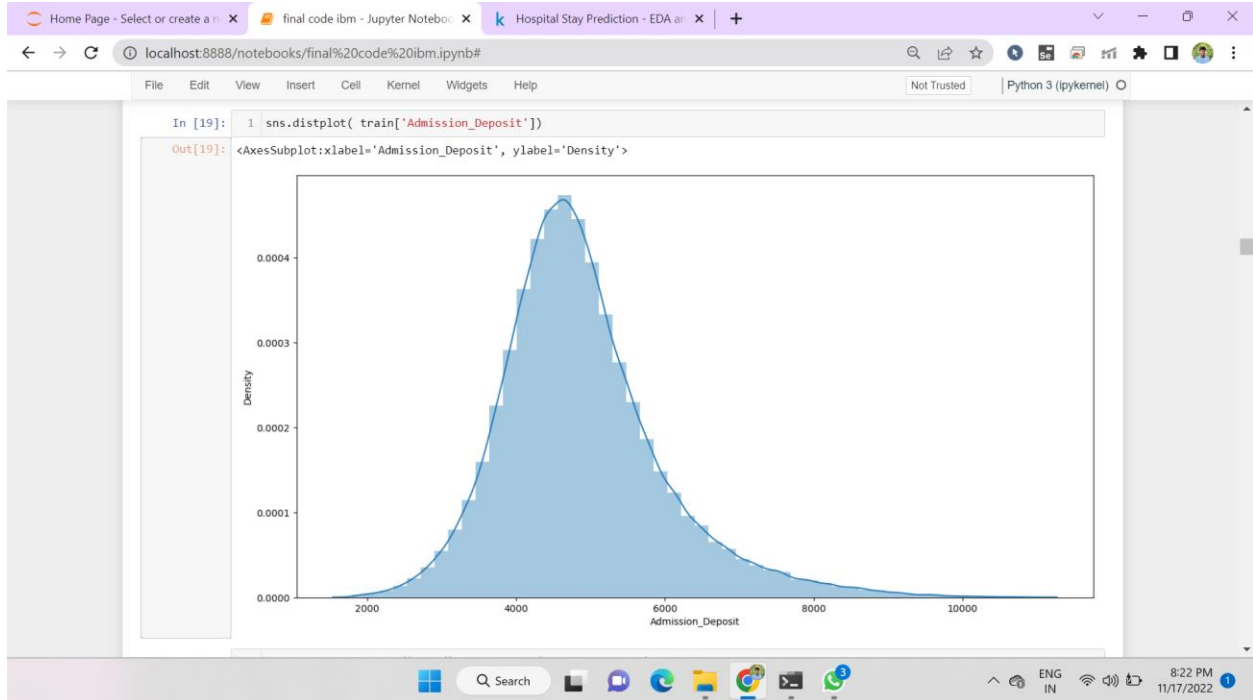
TEAM ID: PNT2022TMID16292



# Analytics for Hospitals Health-Care Data

TEAM ID: PNT2022TMID16292

## NUMERICAL FEATURES



## MISSING VALUE TREATMENT

The code in the Jupyter Notebook is as follows:

```
In [20]: 1 Total = train.isnull().sum().sort_values(ascending=False)
2
3 Percent = (train.isnull().sum()*100/train.isnull().count()).sort_values(ascending=False)
4
5 missing_data = pd.concat([Total, Percent], axis = 1, keys = ['Total', 'Percentage of Missing Values'])
6 missing_data
```

The output is a table with the following data:

	Total	Percentage of Missing Values
City_Code_Patient	4532	1.423197
Bed Grade	113	0.035486
Hospital_code	0	0.000000
Hospital_type_code	0	0.000000
City_Code_Hospital	0	0.000000
Hospital_region_code	0	0.000000
Available Extra Rooms in Hospital	0	0.000000
Department	0	0.000000
Ward_Type	0	0.000000
Ward_Facility_Code	0	0.000000
Type of Admission	0	0.000000
Severity of Illness	0	0.000000
Visitors with Patient	0	0.000000
Age	0	0.000000
Admission_Deposit	0	0.000000
Stay	0	0.000000

# Analytics for Hospitals Health-Care Data

TEAM ID: PNT2022TMID16292

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)

Stay 0 0.000000

```
In [21]: 1 Total = train.isnull().sum().sort_values(ascending=False)
2
3 Percent = (train.isnull().sum()*100/train.isnull().count()).sort_values(ascending=False)
4
5 missing_data = pd.concat([Total, Percent], axis = 1, keys = ['Total', 'Percentage of Missing Values'])
6 missing_data
```

Out[21]:

	Total	Percentage of Missing Values
City_Code_Patient	4532	1.423197
Bed Grade	113	0.035486
Hospital_code	0	0.000000
Hospital_type_code	0	0.000000
City_Code_Hospital	0	0.000000
Hospital_region_code	0	0.000000
Available Extra Rooms in Hospital	0	0.000000
Department	0	0.000000
Ward_Type	0	0.000000
Ward_Facility_Code	0	0.000000
Type of Admission	0	0.000000
Severity of Illness	0	0.000000
Visitors with Patient	0	0.000000
Age	0	0.000000
Admission_Deposit	0	0.000000
Stay	0	0.000000

Windows Search | ENG IN | 8:22 PM 11/17/2022

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 [22]: 1 Total = test.isnull().sum().sort_values(ascending=False)
2
3 Percent = (test.isnull().sum()*100/test.isnull().count()).sort_values(ascending=False)
4
5 missing_data = pd.concat([Total, Percent], axis = 1, keys = ['Total', 'Percentage of Missing Values'])
6 missing_data
```

Out[22]:

	Total	Percentage of Missing Values
City_Code_Patient	2157	1.573798
Bed Grade	35	0.025537
Hospital_code	0	0.000000
Hospital_type_code	0	0.000000
City_Code_Hospital	0	0.000000
Hospital_region_code	0	0.000000
Available Extra Rooms in Hospital	0	0.000000
Department	0	0.000000
Ward_Type	0	0.000000
Ward_Facility_Code	0	0.000000
Type of Admission	0	0.000000
Severity of Illness	0	0.000000
Visitors with Patient	0	0.000000
Age	0	0.000000
Admission_Deposit	0	0.000000

```
In [23]: 1 train.dropna(subset = ['Bed Grade', 'City_Code_Patient'], inplace = True)
```

```
In [24]: 1 test['Bed Grade'] = test['Bed Grade'].fillna(test['Bed Grade'].mode()[0], inplace = True)
```

Windows Search | ENG IN | 8:23 PM 11/17/2022

# Analytics for Hospitals Health-Care Data

TEAM ID: PNT2022TMID16292

DROP UNWANTED COLUMNS

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 [24]: 1 test['Bed_Grade'] = test['Bed_Grade'].fillna(test['Bed_Grade'].mode()[0], inplace = True)
        2 test['City_Code_Patient'] = test['City_Code_Patient'].fillna(test['City_Code_Patient'].mode()[0], inplace = True)

In [25]: 1 df_num_train = train.select_dtypes([np.number])
        2 df_num_train.head()
```

Out[25]:

	Visitors with Patient	Admission_Deposit
0	2	4911.000000
1	2	5954.000000
2	2	4745.000000
3	2	7272.000000
4	2	5558.000000

```
In [26]: 1 df_cat_train = train.select_dtypes([np.object])
        2 df_cat_train.head()
```

Out[26]:

	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	City_Code_Patient
0	8	c	3	Z	3	radiotherapy	R	F	2.000000	7.000000
1	2	c	5	Z	2	radiotherapy	S	F	2.000000	7.000000
2	10	e	1	X	2	anesthesia	S	E	2.000000	7.000000
3	26	b	2	Y	2	radiotherapy	R	D	2.000000	7.000000
4	26	b	2	Y	2	radiotherapy	S	D	2.000000	7.000000

Windows Taskbar: Search, ENG IN, 8:23 PM 11/17/2022