SPRINT-1

TEAM ID	PNT2022TMID35169
PROJECT NAME	Analytics for Hospitals Health-Care Data

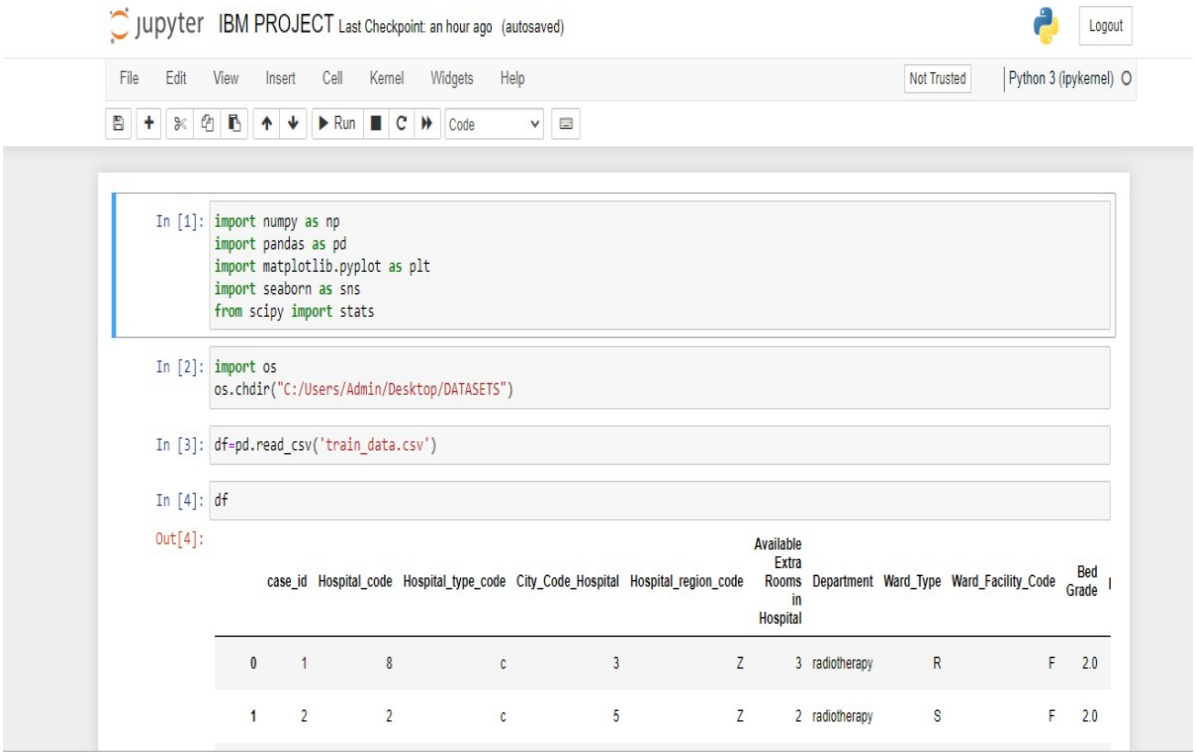
Steps to be done

- Collection of data
- Data preprocessing
- Upload the dataset

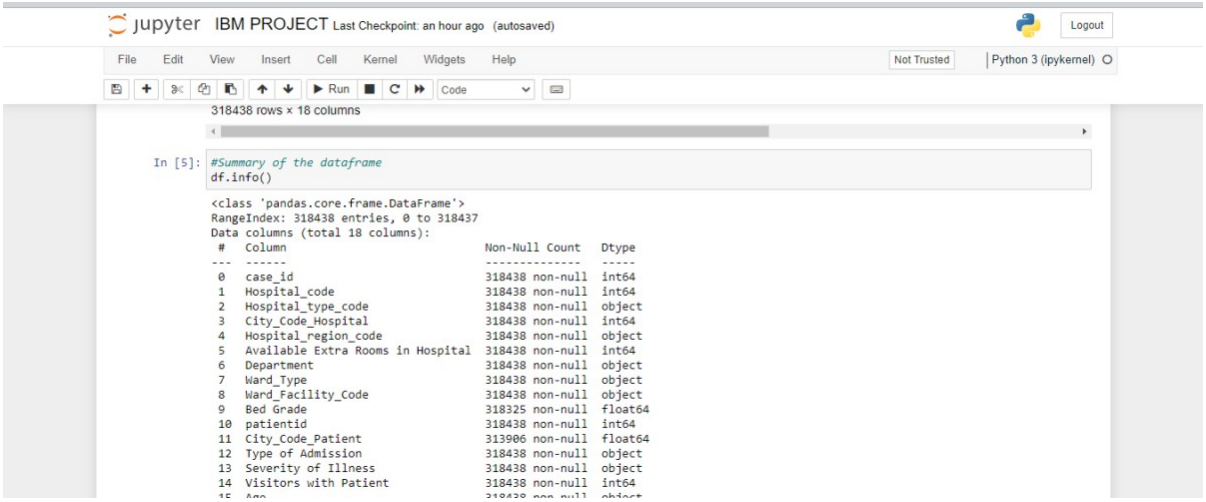
COLLECTION OF DATA:

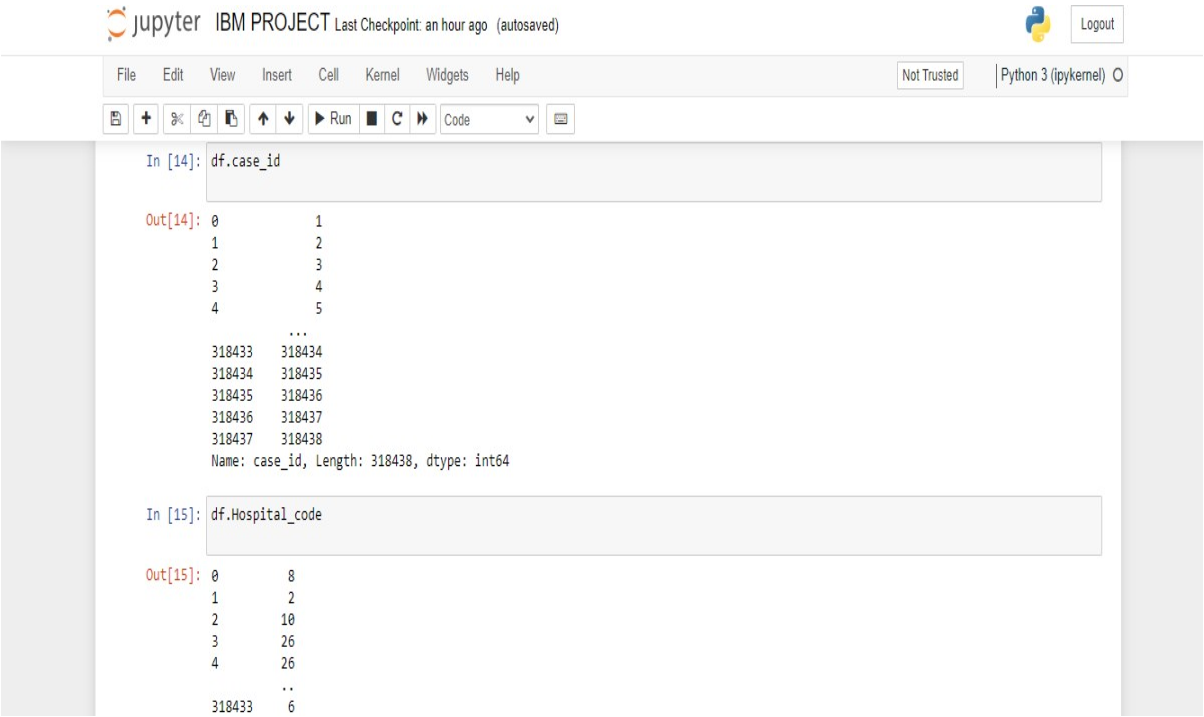
https://www.dropbox.com/s/eqe4g6nvg6qcsmu/Healthcare_Data.zip?dl=0

DATA PREPOCESSING:



	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0





Use Jupyter notebook to remove the null values

jupyter IBM PROJECT Last Checkpoint: an hour ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3 (ipykernel)

Run Code

In [22]:
```
df.Ward_Type
```

Out[22]:
```
0         R
1         S
2         S
3         R
4         S
         ..
318433    Q
318434    Q
318435    R
318436    Q
318437    Q
Name: Ward_Type, Length: 318438, dtype: object
```

In [23]:
```
df.Ward_Facility_Code
```

Out[23]:
```
0         F
1         F
2         E
3         D
4         D
         ..
318433    F
318434    E
```

jupyter IBM PROJECT Last Checkpoint: an hour ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3 (ipykernel)

Run Code

In [42]:
```
df ['Bed Grade'].unique()
```

Out[42]:
```
array([ 2.,  3.,  4.,  1., nan])
```

In [43]:
```
df['Visitors with Patient'].unique()
```

Out[43]:
```
array([ 2,  4,  3,  8,  6,  7, 13,  5,  1, 10, 15, 11, 12,  9, 24, 16, 14,
       20,  0, 19, 18, 17, 23, 21, 32, 30, 22, 25], dtype=int64)
```

In [44]:
```
df['Severity of Illness']
```

Out[44]:
```
0         Extreme
1         Extreme
2         Extreme
3         Extreme
4         Extreme
            ...
318433    Moderate
318434    Moderate
318435       Minor
318436       Minor
318437       Minor
Name: Severity of Illness, Length: 318438, dtype: object
```

Upload the dataset: