Exploratory Data Analysis:

Required libraries:

```python
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline
```

```python
data = pd.read_csv("../Source/train/healthcare/healthcare_train_data.csv")
```

```python
df
```


318438 rows × 18 columns

```python
df.head()
```



```python
df.tail()
```



```python
df.info()
```



```python
df.dtypes
```



```python
df.shape
```

(318438, 18)

Before Null Values checking :

```python
df.isnull().sum().sum()
```

```python
df.isnull()
```


318438 rows × 18 columns

```python
df.describe()
```



```python
df.isnull().sum()
```



```python
df.corr()
```



```python
df.isnull().sum()
```

Work With Null Values :

```python
df['Bed Grade'].fillna(df['Bed Grade'].median(),inplace=True)
df['City_Code_Patient'].isnull().sum()
```

```python
df.isnull().sum()
```



```python
df['City_Code_Patient'].fillna(method='ffill',inplace=True)
```

```python
df['City_Code_Patient'].isnull().sum()
```

After Cleaning Process :

Total Null Values Checking :

```python
df.isnull().sum()
```


Total Null Values :

```python
df.isnull().sum().sum()
```

```python
df.corr()
```



```python
sns.heatmap(df.corr(),annot=True)
plt.title("Correlation Matrix")
plt.show()
```


correlation Matrix

```python
df['Admission_Deposit'].hist(bins=30)
plt.title("Histogram for Admission_Deposit")
plt.show()
```


Histogram for Admission_Deposit

```python
df['Ward_Type'].hist()
plt.title("Histogram for Ward_Type")
plt.show()
```


Histogram for Ward_Type

```python
df['patientid'].hist(bins=30)
plt.title("Histogram for patientid")
plt.show()
```


Histogram for patientid