



Project Title	Drug safety and pharmacovigilance
Tools	Python, ML, SQL, Excel
Technologies	Data Analyst & Data scientist
Project Difficulties level	intermediate

Dataset : Dataset is available in the given link. You can download it at your convenience.

[Click here to download data set](#)

## About Dataset

### Context

Since as a beginner in machine learning it would be a great opportunity to try some techniques to predict the outcome of the drugs that might be accurate for the patient.

### Content

The target feature is

- Drug type

The feature sets are:

- Age

- Sex
- Blood Pressure Levels (BP)
- Cholesterol Levels
- Na to Potassium Ration

## Inspiration

The main problem here is not just the feature sets and target sets but also the approach that is taken in solving these types of problems as a beginner. So best of luck.

- Drug safety and pharmacovigilance are critical aspects of the pharmaceutical industry that focus on monitoring, assessing, and preventing adverse effects or any other drug-related problems. Python, as a versatile programming language, is widely used in this field for various tasks, including data analysis, signal detection, and automation of pharmacovigilance processes. Some ways Python is utilized in drug safety and pharmacovigilance:
  - Data Collection and Integration: Python can be used to collect and integrate data from various sources such as electronic health records (EHRs), clinical trials, and spontaneous reporting systems. Libraries like Pandas and NumPy are commonly employed for data manipulation and analysis.
  - Signal Detection and Data Mining: Python's machine learning and data mining libraries, such as scikit-learn and TensorFlow, can be applied to identify potential safety signals from large datasets. Algorithms like decision trees, random forests, and neural networks can help in recognizing patterns and trends in adverse event data.
  - Text Mining and Natural Language Processing (NLP): NLP tools in Python, like NLTK (Natural Language Toolkit) and spaCy, are used for mining information from unstructured data sources such as medical literature, social media, and clinical narratives. This helps in identifying and extracting relevant information related to drug safety.
  - Database Management: Python is often used to interact with databases (e.g., SQL databases) to manage and query pharmacovigilance data efficiently. SQLAlchemy is a popular Python library for database management.
  - Automation of Workflows: Pharmacovigilance processes involve numerous repetitive tasks. Python can be employed to automate these workflows, improving efficiency and reducing the risk of human error. Tools like Apache Airflow can be used for workflow automation.
  - Data Visualization: Python libraries such as Matplotlib and Seaborn enable the creation of visualizations for better understanding and presentation of pharmacovigilance data. Visualizations help in identifying patterns and trends that may not be apparent in raw data.
  - Web Scraping: Python's web scraping tools, such as BeautifulSoup and Scrapy, can be used to extract information from websites, forums, and social media platforms, providing additional data sources for pharmacovigilance activities.

- Regulatory Reporting: Python can be utilized to generate reports required for regulatory compliance. Frameworks like ReportLab can be used to create PDF reports summarizing safety data for submission to regulatory authorities.
- Data Quality Assurance: Python scripts can be developed to perform data validation and quality checks to ensure that the collected data is accurate and meets regulatory standards.
- Adverse Event Case Processing: Python can be used to develop tools for the processing of adverse event cases, including data entry, validation, and reporting.

#### NOTE :

1. this project is only for your guidance, not exactly the same you have to create. Here I am trying to show the way or idea of what steps you can follow and how your projects look. Some projects are very advanced (because it will be made with the help of flask, nlp, advance ai, advance DL and some advanced things ) which you can not understand .
2. You can make or analyze your project with yourself, with your idea, make it more creative from where we can get some information and understand about our business. make sure what overall things you have created all things you understand very well.

## **Example**

what steps you should have to follow

Here's a step-by-step guide for a beginner to work on a data analytics project focusing on Drug Safety and Pharmacovigilance using the provided columns:

### **Project Overview**

The aim of this project is to analyze the

40

You said:

data analytics project of Drug safety and pharmacovigilance for this column Age Sex Blood Pressure Levels (BP) Cholesterol Levels Na to Potassium Ratio for the beginner

For a beginner working on a data analytics project related to Drug Safety and Pharmacovigilance using the columns **Age, Sex, Blood Pressure Levels (BP), Cholesterol Levels**, and **Na to Potassium Ratio**, here's a structured approach:

**Project Title:**

**Exploratory Analysis of Patient Health Indicators in Drug Safety and Pharmacovigilance**

### **1. Objective**

The objective of this project is to analyze patient health data to understand how age, sex,

40

**You said:**

data analytics project of Drug safety and pharmacovigilance for these coloum Age Sex Blood Pressure Levels (BP) Cholesterol Levels Na to Potassium Ration for the biggner with code and output

**ChatGPT said:**

ChatGPT

Here's a step-by-step guide for a beginner to work on a data analytics project related to Drug Safety and Pharmacovigilance using the provided columns, along with sample Python code and expected output.

**Project Title:**

**Exploratory Data Analysis (EDA) of Patient Health Indicators in Drug Safety and**

# Pharmacovigilance

## 1. Objective

The goal is to explore the relationships between patient demographics and health indicators like blood pressure, cholesterol levels, and Na to potassium ratio. This exploration will help identify any patterns or correlations that could be important for drug safety and pharmacovigilance.

## 2. Dataset Overview

Let's assume you have a dataset with the following columns:

- **Age:** The age of the patient.
- **Sex:** The gender of the patient (e.g., Male/Female).
- **Blood Pressure Levels (BP):** Numerical or categorical values representing blood pressure.
- **Cholesterol Levels:** Numerical values representing cholesterol levels.
- **Na to Potassium Ratio:** The ratio of sodium to potassium in the body.

## 3. Python Code for Data Analysis

### Step 1: Import Required Libraries

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Enable inline plotting for Jupyter Notebooks
%matplotlib inline
```

## Step 2: Load the Dataset

```
# Assuming you have the data in a CSV file named  
'drug_safety_data.csv'  
df = pd.read_csv('drug_safety_data.csv')  
  
# Display the first few rows of the dataset  
df.head()
```

### Expected Output:

This will show the first few rows of your dataset, which should look something like this:

Age	Sex	BP	Cholesterol	Sodium to Potassium Ratio
45	Male	130/85	180	35.2
50	Female	140/90	220	30.1
...	...	...	...	...

## Step 3: Data Cleaning (if necessary)

```
# Check for missing values  
df.isnull().sum()
```

```
# If there are missing values, handle them accordingly
df = df.dropna() # Drop rows with missing values (or use
another method to handle them)
```

#### Step 4: Exploratory Data Analysis (EDA)

##### 1. Distribution of Age and Sex:

```
# Age distribution
plt.figure(figsize=(10,5))
sns.histplot(df['Age'], kde=True, bins=10)
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
```

```
# Sex distribution
plt.figure(figsize=(5,5))
sns.countplot(x='Sex', data=df)
plt.title('Sex Distribution')
plt.xlabel('Sex')
plt.ylabel('Count')
plt.show()
```

#### Expected Output:

- **Age Distribution:** A histogram showing how patient ages are distributed.

- **Sex Distribution:** A bar chart showing the count of males and females in the dataset.

## 2. Blood Pressure Analysis:

python

Copy code

```
# Convert BP into categorical if not already done
def categorize_bp(bp):
    sys, dia = map(int, bp.split('/'))
    if sys >= 140 or dia >= 90:
        return 'High'
    elif sys < 120 and dia < 80:
        return 'Normal'
    else:
        return 'Elevated'

df['BP Category'] = df['BP'].apply(categorize_bp)

# Plot BP categories
plt.figure(figsize=(7,5))
sns.countplot(x='BP Category', data=df)
plt.title('Blood Pressure Categories')
plt.xlabel('BP Category')
plt.ylabel('Count')
plt.show()
```



### Expected Output:

- A bar chart showing the distribution of patients across different blood pressure categories (e.g., High, Normal, Elevated).

### 3. Correlation Analysis:

```
# Correlation between numerical features
corr_matrix = df[['Age', 'Cholesterol', 'Na to Potassium
Ratio']].corr()

plt.figure(figsize=(8,6))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

### Expected Output:

- A heatmap showing the correlation between age, cholesterol levels, and Na to Potassium ratio.

### 4. Boxplots to Analyze Distribution by Sex:

```
# Cholesterol Levels by Sex
plt.figure(figsize=(8,6))
sns.boxplot(x='Sex', y='Cholesterol', data=df)
plt.title('Cholesterol Levels by Sex')
plt.xlabel('Sex')
```

```
plt.ylabel('Cholesterol')
plt.show()

# Na to Potassium Ratio by Sex
plt.figure(figsize=(8,6))
sns.boxplot(x='Sex', y='Na to Potassium Ratio', data=df)
plt.title('Na to Potassium Ratio by Sex')
plt.xlabel('Sex')
plt.ylabel('Na to Potassium Ratio')
plt.show()
```

### Expected Output:

- Boxplots showing the distribution of cholesterol levels and Na to potassium ratio by sex.

### 4. Interpretation of Results

- **Age Distribution:** An understanding of which age groups are most represented in the dataset.
- **Sex Distribution:** The ratio of male to female patients.
- **Blood Pressure Categories:** Insights into how many patients fall into high, normal, or elevated blood pressure categories.
- **Correlation Matrix:** Whether there's a significant correlation between the variables.
- **Boxplots:** Differences in cholesterol levels and Na to Potassium ratio between males and females.

### 5. Conclusion

Based on the exploratory analysis, you can summarize key findings such as:

- Whether certain age groups are more prone to high blood pressure or high cholesterol.
- How sodium and potassium balance varies with age and sex.
- Initial insights into how these variables might impact drug safety.

This project provides a solid foundation for understanding basic data analysis concepts in the context of Drug Safety and Pharmacovigilance.

## **Sample code**

[Reference link](#) for this code

```
import pandas as pd
```

In [3]:

```
df=pd.read_csv('./drug200.csv')
```

In [4]:

```
df.head() #for first 5 rows
```

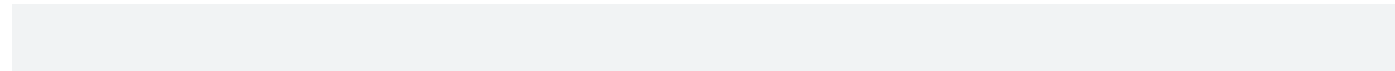
Out[4]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug

0	23	F	HIGH	HIGH	25.355	DrugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	DrugY

In [5]:

```
df.tail() #for last 5 rows
```



Out[5]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX

198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

In [6]:

```
df.info() #to check null,dtype,column names
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 200 entries, 0 to 199
```

```
Data columns (total 6 columns):
```

```

#      Column      Non-Null Count  Dtype
---  -
0     Age         200 non-null    int64
1     Sex         200 non-null    object
2     BP          200 non-null    object
3     Cholesterol  200 non-null    object
4     Na_to_K     200 non-null    float64
5     Drug        200 non-null    object

```

```
dtypes: float64(1), int64(1), object(4)
```

```
memory usage: 9.5+ KB
```

In [7]:

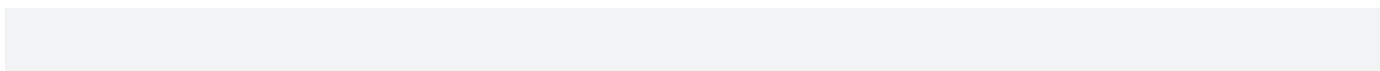
```
df.describe() #to apply math on numerical columns
```

Out[7]:

	Age	Na_to_K
count	200.000000	200.000000
mean	44.315000	16.084485
std	16.544315	7.223956
min	15.000000	6.269000
25%	31.000000	10.445500
50%	45.000000	13.936500
75%	58.000000	19.380000
max	74.000000	38.247000

In [8]:

```
df[['Age', 'BP']] #to extract specific columns
```



Out[8]:

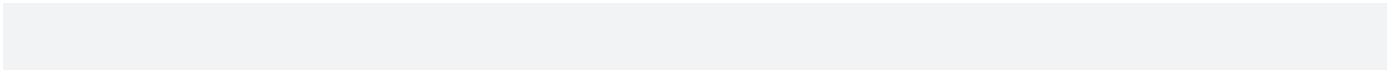
	Age	BP
0	23	HIGH
1	47	LOW
2	47	LOW
3	28	NORMAL
4	61	LOW
...	...	...
195	56	LOW
196	16	LOW
197	52	NORMAL

198	23	NORMAL
199	40	LOW

200 rows × 2 columns

In [9]:

```
df['Age']>40 #condition based extraction
```



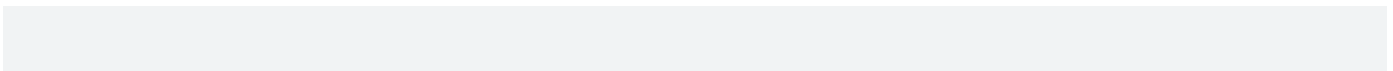
Out[9]:

```
0      False
1       True
2       True
3      False
4       True
...
195     True
196    False
197     True
198    False
199    False
```

Name: Age, Length: 200, dtype: bool

In [10]:

```
df[df['Age']>40]
```



Out[10]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
--	-----	-----	----	-------------	---------	------



1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
4	61	F	LOW	HIGH	18.043	DrugY
6	49	F	NORMAL	HIGH	16.275	DrugY
7	41	M	LOW	HIGH	11.037	drugC
...	...	...	...	...	...	...
192	72	M	LOW	HIGH	16.310	DrugY
193	72	M	LOW	HIGH	6.769	drugC
194	46	F	HIGH	HIGH	34.686	DrugY
195	56	F	LOW	HIGH	11.567	drugC
197	52	M	NORMAL	HIGH	9.894	drugX

113 rows × 6 columns

In [11]:

```
df[(df['Age']>40)&(df['BP']=='HIGH')]
```

Out[11]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
17	43	M	HIGH	HIGH	13.972	drugA
31	74	M	HIGH	HIGH	9.567	drugB
33	65	F	HIGH	NORMAL	31.876	DrugY
41	58	F	HIGH	NORMAL	14.239	drugB
50	58	F	HIGH	HIGH	19.416	DrugY
54	68	F	HIGH	NORMAL	10.189	drugB
56	65	M	HIGH	NORMAL	11.340	drugB

64	60	F	HIGH	HIGH	13.303	drugB
70	70	M	HIGH	HIGH	13.967	drugB
80	60	M	HIGH	HIGH	13.934	drugB
85	59	M	HIGH	HIGH	13.935	drugB
91	41	M	HIGH	NORMAL	15.156	DrugY
97	56	F	HIGH	HIGH	25.395	DrugY
101	45	F	HIGH	HIGH	12.854	drugA
108	72	M	HIGH	NORMAL	9.677	drugB
110	50	M	HIGH	HIGH	7.490	drugA
115	51	M	HIGH	HIGH	18.295	DrugY
119	61	F	HIGH	HIGH	25.475	DrugY

124	53	F	HIGH	NORMAL	12.495	drugB
126	66	M	HIGH	HIGH	16.347	DrugY
134	42	F	HIGH	HIGH	21.036	DrugY
136	55	F	HIGH	HIGH	10.977	drugB
138	51	M	HIGH	NORMAL	11.343	drugB
140	49	M	HIGH	NORMAL	6.269	drugA
142	60	M	HIGH	NORMAL	8.621	drugB
143	74	M	HIGH	NORMAL	15.436	DrugY
150	49	M	HIGH	NORMAL	8.700	drugA
151	68	M	HIGH	HIGH	11.009	drugB
161	57	F	HIGH	NORMAL	9.945	drugB

174	42	M	HIGH	NORMAL	12.766	drugA
175	73	F	HIGH	HIGH	18.348	DrugY
176	48	M	HIGH	NORMAL	10.446	drugA
186	70	M	HIGH	HIGH	9.849	drugB
187	47	M	HIGH	HIGH	10.403	drugA
188	65	M	HIGH	NORMAL	34.997	DrugY
189	64	M	HIGH	NORMAL	20.932	DrugY
190	58	M	HIGH	HIGH	18.991	DrugY
194	46	F	HIGH	HIGH	34.686	DrugY

In [12]:

```
df.groupby('BP').count() #to count by specific group
```

Out[12]:

	Age	Sex	Cholesterol	Na_to_K	Drug
BP					
HIGH	77	77	77	77	77
LOW	64	64	64	64	64
NORMAL	59	59	59	59	59

In [13]:

```
df.shape #rows columns
```

Out[13]:

```
(200, 6)
```

In [14]:

```
df.columns #columns
```

Out[14]:

```
Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='object')
```

In [15]:

```
df.dtypes #column data types
```

Out[15]:

Age int64  
Sex object  
BP object  
Cholesterol object  
Na\_to\_K float64  
Drug object

dtype: object

In [16]:

```
df.groupby(['Age', 'Drug']).count()
```

Out[16]:

		Sex	BP	Cholesterol	Na_to_K
Age	Drug				
15	DrugY	2	2	2	2
	drugX	1	1	1	1
16	DrugY	2	2	2	2
	drugC	1	1	1	1

17	drugX	1	1	1	1
...	...	...	...	...	...
72	drugX	1	1	1	1
73	DrugY	2	2	2	2
74	DrugY	2	2	2	2
	drugB	1	1	1	1
	drugX	1	1	1	1

135 rows × 4 columns

In [17]:

```
df.groupby('Age').agg({'Na_to_K':'mean', 'Cholesterol':'max'}) #specific column
extraction based on specific calculation
```



Out[17]:

	Na_to_K	Cholesterol
--	---------	-------------



Age		
15	14.338333	NORMAL
16	15.509667	NORMAL
17	10.832000	NORMAL
18	23.404667	NORMAL
19	19.641000	NORMAL
20	16.967000	NORMAL
21	28.632000	NORMAL
22	15.964600	NORMAL
23	16.497143	NORMAL
24	16.080750	NORMAL

25	19.011000	HIGH
26	16.634250	NORMAL
28	17.021143	NORMAL
29	21.153000	HIGH
30	10.443000	HIGH
31	17.633250	NORMAL
32	12.290000	NORMAL
33	33.486000	HIGH
34	18.320250	NORMAL
35	9.969667	NORMAL
36	13.716250	NORMAL

37	14.776000	NORMAL
38	19.832000	NORMAL
39	14.867000	NORMAL
40	16.426000	NORMAL
41	16.959250	NORMAL
42	20.771500	NORMAL
43	15.393750	NORMAL
45	12.298000	NORMAL
46	20.985500	NORMAL
47	15.779625	NORMAL
48	12.741000	NORMAL

49	10.824857	NORMAL
50	13.097800	NORMAL
51	16.559500	NORMAL
52	21.408000	NORMAL
53	16.530333	NORMAL
54	24.658000	HIGH
55	9.119000	NORMAL
56	15.235750	HIGH
57	17.295500	NORMAL
58	23.507600	NORMAL
59	12.754333	HIGH

60	12.224000	NORMAL
61	15.075250	NORMAL
62	21.888500	NORMAL
63	25.917000	HIGH
64	18.144667	NORMAL
65	22.995500	NORMAL
66	12.227000	NORMAL
67	14.249000	NORMAL
68	14.634750	NORMAL
69	12.332667	NORMAL
70	14.768333	HIGH

72	11.849500	NORMAL
73	18.784500	HIGH
74	14.471000	NORMAL

In [18]:

```
df.isna().sum() #find null
```

Out[18]:

```
Age          0
Sex          0
BP           0
Cholesterol  0
Na_to_K      0
Drug         0
```

```
dtype: int64
```

In [19]:

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

In [20]:

```
# df=df.fillna(369)
```

In [21]:

```
df=df.rename(columns={'Na_to_K':'Sodium to Potassium Ratio'})
```

In [22]:

```
df.head()
```

Out[22]:

	Age	Sex	BP	Cholesterol	Sodium to Potassium Ratio	Drug
0	23	F	HIGH	HIGH	25.355	DrugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	DrugY

In [23]:

```
# df=df.drop_duplicates() to remove duplicate rows
```

In [24]:

```
df['Age2']=df['Age'].apply(lambda x : x * 1)
```

In [25]:

```
df.head()
```

Out[25]:

	Age	Sex	BP	Cholesterol	Sodium to Potassium Ratio	Drug	Age2
0	23	F	HIGH	HIGH	25.355	DrugY	23
1	47	M	LOW	HIGH	13.093	drugC	47
2	47	M	LOW	HIGH	10.114	drugC	47
3	28	F	NORMAL	HIGH	7.798	drugX	28
4	61	F	LOW	HIGH	18.043	DrugY	61

In [26]:

```
df['Drug'].unique #for unique
```

Out[26]:

```
<bound method Series.unique of 0      DrugY
1      drugC
2      drugC
3      drugX
4      DrugY
...
195    drugC
196    drugC
197    drugX
198    drugX
```



199 drugX

Name: Drug, Length: 200, dtype: object>

In [27]:

```
df=df.drop(columns=['Age2'])
```

In [28]:

```
df.head()
```

Out[28]:

	Age	Sex	BP	Cholesterol	Sodium to Potassium Ratio	Drug
0	23	F	HIGH	HIGH	25.355	DrugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	DrugY

In [29]:

```
newdf=df[['Sex','Drug']] #new dataframe  
newdf
```

Out[29]:

	Sex	Drug
0	F	DrugY
1	M	drugC
2	M	drugC
3	F	drugX
4	F	DrugY
...	...	...
195	F	drugC
196	M	drugC

197	M	drugX
198	M	drugX
199	F	drugX

200 rows × 2 columns

In [30]:

```
df.iloc[5] #row extraction
```

Out[30]:

```
Age                22
Sex                F
BP                NORMAL
Cholesterol        HIGH
Sodium to Potassium Ratio  8.607
Drug              drugX
```

Name: 5, dtype: object

In [31]:

```
type(df) #type check
```

Out[31]:

```
pandas.core.frame.DataFrame
```

In [32]:

```
type(df.iloc[5])
```

Out[32]:

```
pandas.core.series.Series
```

In [33]:

```
df.iloc[0:20] #slicing
```

Out[33]:

	Age	Sex	BP	Cholesterol	Sodium to Potassium Ratio	Drug
0	23	F	HIGH	HIGH	25.355	DrugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	DrugY
5	22	F	NORMAL	HIGH	8.607	drugX

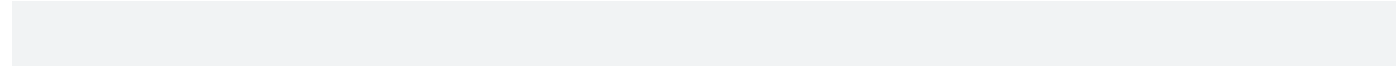
6	49	F	NORMAL	HIGH	16.275	DrugY
7	41	M	LOW	HIGH	11.037	drugC
8	60	M	NORMAL	HIGH	15.171	DrugY
9	43	M	LOW	NORMAL	19.368	DrugY
10	47	F	LOW	HIGH	11.767	drugC
11	34	F	HIGH	NORMAL	19.199	DrugY
12	43	M	LOW	HIGH	15.376	DrugY
13	74	F	LOW	HIGH	20.942	DrugY
14	50	F	NORMAL	HIGH	12.703	drugX
15	16	F	HIGH	NORMAL	15.516	DrugY
16	69	M	LOW	NORMAL	11.455	drugX

17	43	M	HIGH	HIGH	13.972	drugA
18	23	M	LOW	HIGH	7.298	drugC
19	32	F	HIGH	NORMAL	25.974	DrugY

In [34]:

# linkcode

```
for i in df.iterrows():
    print(i)
```



(0, Age 23  
Sex F  
BP HIGH  
Cholesterol HIGH  
Sodium to Potassium Ratio 25.355  
Drug DrugY  
Name: 0, dtype: object)

(1, Age 47  
Sex M  
BP LOW  
Cholesterol HIGH  
Sodium to Potassium Ratio 13.093  
Drug drugC

Name: 1, dtype: object)

(2, Age 47

Sex M

BP LOW

Cholesterol HIGH

Sodium to Potassium Ratio 10.114

Drug drugC

Name: 2, dtype: object)

(3, Age 28

Sex F

BP NORMAL

Cholesterol HIGH

Sodium to Potassium Ratio 7.798

Drug drugX

Name: 3, dtype: object)

(4, Age 61

Sex F

BP LOW

Cholesterol HIGH

Sodium to Potassium Ratio 18.043

Drug DrugY

Name: 4, dtype: object)

(5, Age 22

Sex F

BP NORMAL

Cholesterol HIGH

Sodium to Potassium Ratio	8.607
Drug	drugX
Name: 5, dtype: object)	
(6, Age	49
Sex	F
BP	NORMAL
Cholesterol	HIGH
Sodium to Potassium Ratio	16.275
Drug	DrugY
Name: 6, dtype: object)	
(7, Age	41
Sex	M
BP	LOW
Cholesterol	HIGH
Sodium to Potassium Ratio	11.037
Drug	drugC
Name: 7, dtype: object)	
(8, Age	60
Sex	M
BP	NORMAL
Cholesterol	HIGH
Sodium to Potassium Ratio	15.171
Drug	DrugY
Name: 8, dtype: object)	
(9, Age	43



Sex	M
BP	LOW
Cholesterol	NORMAL
Sodium to Potassium Ratio	19.368
Drug	DrugY
Name: 9, dtype: object)	
(10, Age	47
Sex	F
BP	LOW
Cholesterol	HIGH
Sodium to Potassium Ratio	11.767
Drug	drugC
Name: 10, dtype: object)	
(11, Age	34
Sex	F
BP	HIGH
Cholesterol	NORMAL
Sodium to Potassium Ratio	19.199
Drug	DrugY
Name: 11, dtype: object)	
(12, Age	43
Sex	M
BP	LOW
Cholesterol	HIGH
Sodium to Potassium Ratio	15.376
Drug	DrugY

Name: 12, dtype: object)

(13, Age 74

Sex F

BP LOW

Cholesterol HIGH

Sodium to Potassium Ratio 20.942

Drug DrugY

Name: 13, dtype: object)

(14, Age 50

Sex F

BP NORMAL

Cholesterol HIGH

Sodium to Potassium Ratio 12.703

Drug drugX

Name: 14, dtype: object)

(15, Age 16

Sex F

BP HIGH

Cholesterol NORMAL

Sodium to Potassium Ratio 15.516

Drug DrugY

Name: 15, dtype: object)

(16, Age 69

Sex M

BP LOW

Cholesterol	NORMAL
Sodium to Potassium Ratio	11.455
Drug	drugX
Name: 16, dtype: object)	
(17, Age	43
Sex	M
BP	HIGH
Cholesterol	HIGH
Sodium to Potassium Ratio	13.972
Drug	drugA
Name: 17, dtype: object)	
(18, Age	23
Sex	M
BP	LOW
Cholesterol	HIGH
Sodium to Potassium Ratio	7.298
Drug	drugC
Name: 18, dtype: object)	
(19, Age	32
Sex	F
BP	HIGH
Cholesterol	NORMAL
Sodium to Potassium Ratio	25.974
Drug	DrugY
Name: 19, dtype: object)	
(20, Age	57

Sex M

BP LOW

Cholesterol NORMAL

Sodium to Potassium Ratio 19.128

Drug DrugY

Name: 20, dtype: object)

(21, Age 63

Sex M

BP NORMAL

Cholesterol HIGH

Sodium to Potassium Ratio 25.917

Drug DrugY

Name: 21, dtype: object)

(22, Age 47

Sex M

BP LOW

Cholesterol NORMAL

Sodium to Potassium Ratio 30.568

Drug DrugY

Name: 22, dtype: object)

(23, Age 48

Sex F

BP LOW

Cholesterol HIGH

Sodium to Potassium Ratio 15.036

Drug	DrugY
Name: 23, dtype: object)	
(24, Age	33
Sex	F
BP	LOW
Cholesterol	HIGH
Sodium to Potassium Ratio	33.486
Drug	DrugY
Name: 24, dtype: object)	
(25, Age	28
Sex	F
BP	HIGH
Cholesterol	NORMAL
Sodium to Potassium Ratio	18.809
Drug	DrugY
Name: 25, dtype: object)	
(26, Age	31
Sex	M
BP	HIGH
Cholesterol	HIGH
Sodium to Potassium Ratio	30.366
Drug	DrugY
Name: 26, dtype: object)	
(27, Age	49
Sex	F
BP	NORMAL

Cholesterol                      NORMAL

Sodium to Potassium Ratio      9.381

Drug                              drugX

Name: 27, dtype: object)

(28, Age                              39

Sex                                F

BP                                 LOW

Cholesterol                      NORMAL

Sodium to Potassium Ratio      22.697

Drug                              DrugY

Name: 28, dtype: object)

Cholesterol                      HIGH

Sodium to Potassium Ratio      19.221

Drug                              DrugY

Name: 40, dtype: object)

(41, Age                              58

Sex                                F

BP                                 HIGH

Cholesterol                      NORMAL

Sodium to Potassium Ratio      14.239

Drug                              drugB

Name: 41, dtype: object)

(42, Age                              50

Sex                                M

BP                                 NORMAL

Cholesterol                      NORMAL

Sodium to Potassium Ratio      15.79

Drug      DrugY

Name: 42, dtype: object)

(43, Age      23

Sex      M

BP      NORMAL

Cholesterol      HIGH

Sodium to Potassium Ratio      12.26

Drug      drugX

Name: 43, dtype: object)

(44, Age      50

Sex      F

BP      NORMAL

Cholesterol      NORMAL

Sodium to Potassium Ratio      12.295

Drug      drugX

Name: 44, dtype: object)

(45, Age      66

Sex      F

BP      NORMAL

Cholesterol      NORMAL

Sodium to Potassium Ratio      8.107

Drug      drugX

Name: 45, dtype: object)

(46, Age      37

Sex	F
BP	HIGH
Cholesterol	HIGH
Sodium to Potassium Ratio	13.091
Drug	drugA
Name: 46, dtype: object)	
(47, Age	68
Sex	M

1 [Reference link](#)

2 [Reference link](#) for ML project