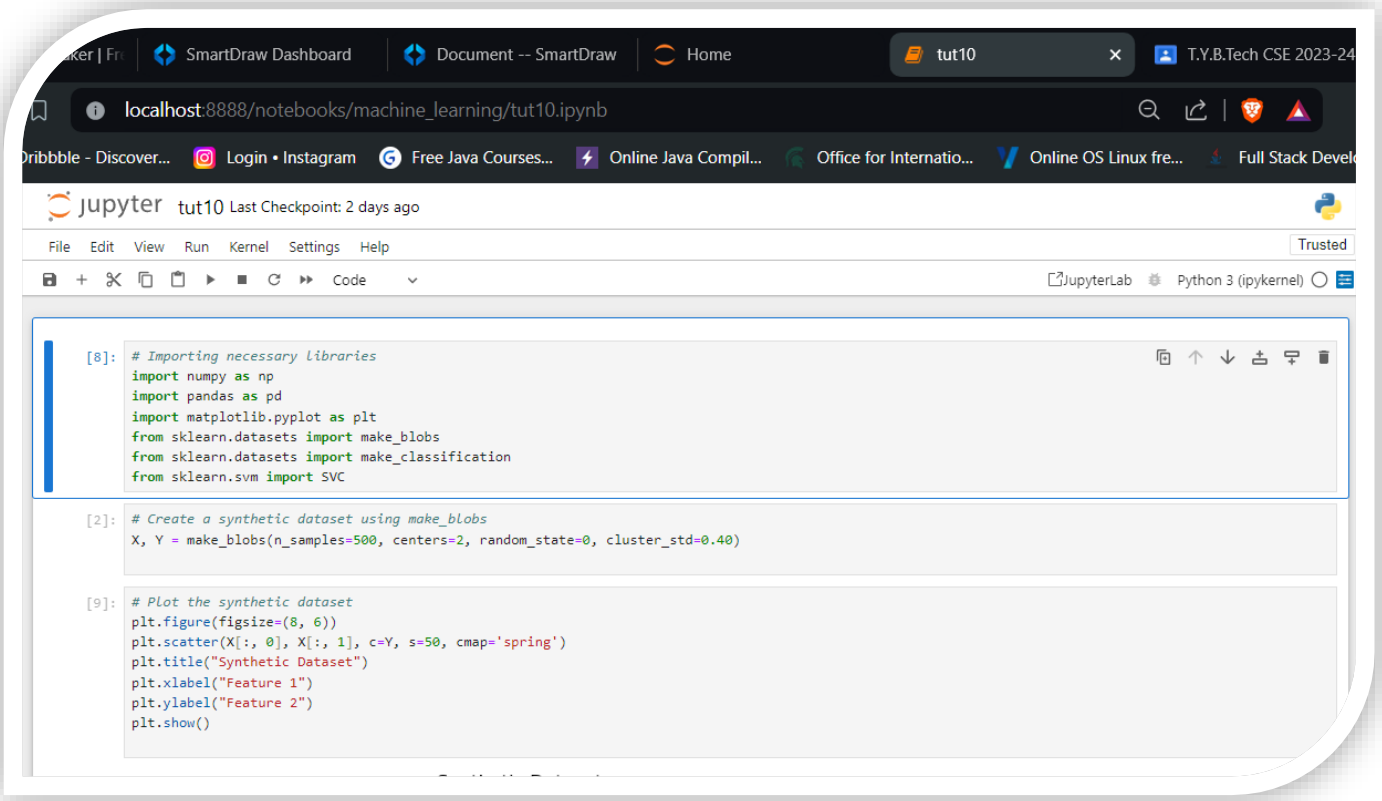
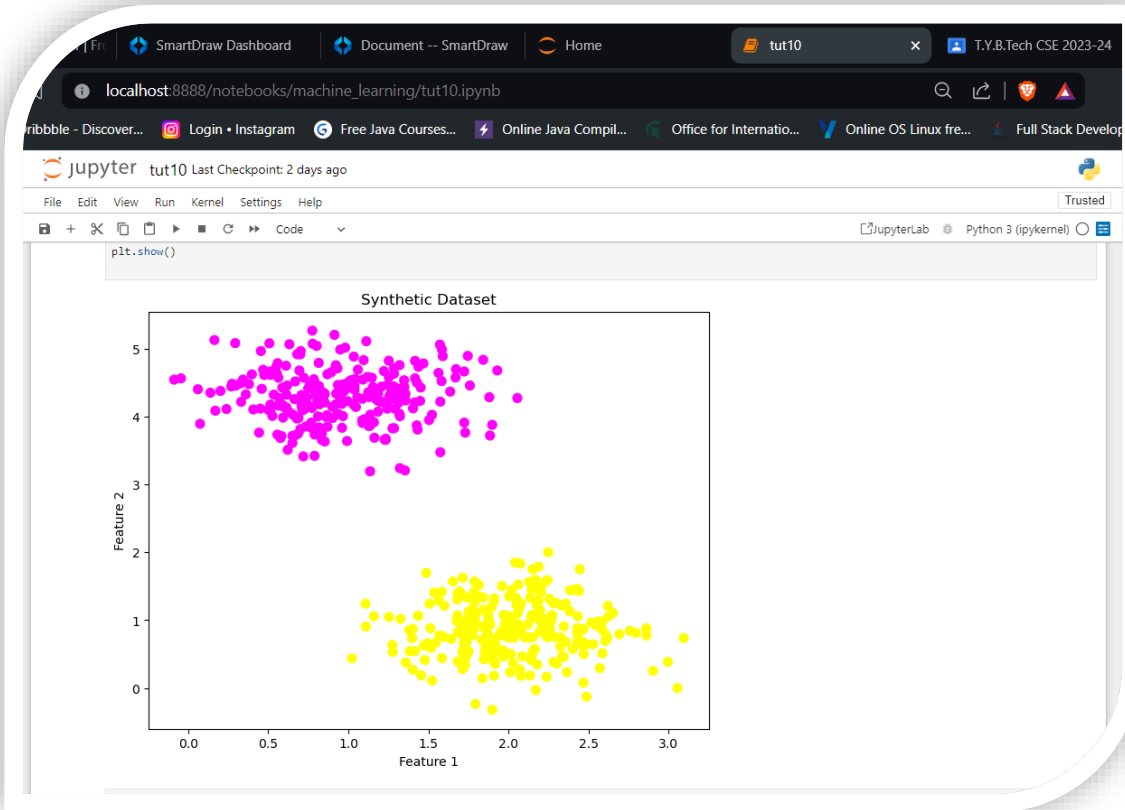
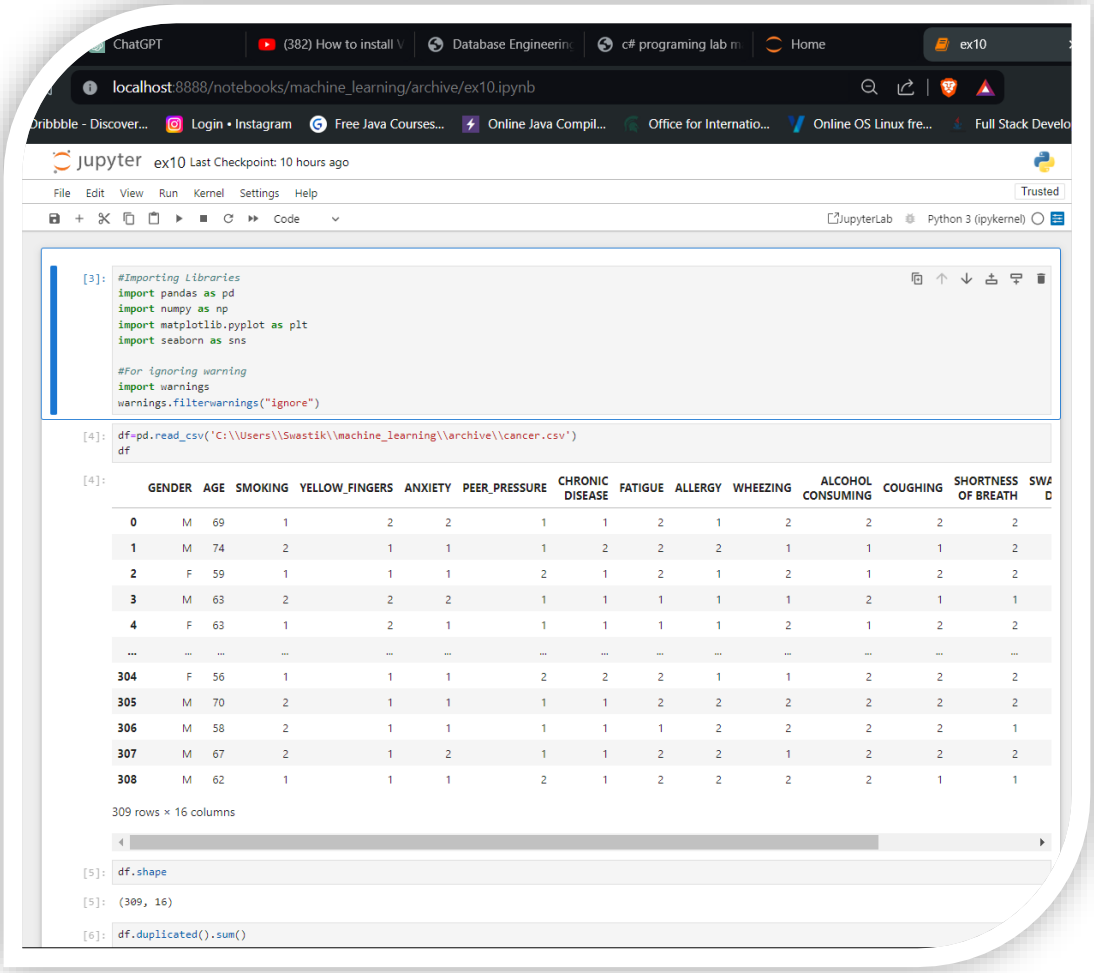
Support vector machine algorithm.





Tut 10



```
[3]: #Importing Libraries
     import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import seaborn as sns

     #For ignoring warning
     import warnings
     warnings.filterwarnings("ignore")
```

```
[4]: df=pd.read_csv('C:\\Users\\Swastik\\machine_learning\\archive\\cancer.csv')
     df
```

[4]:

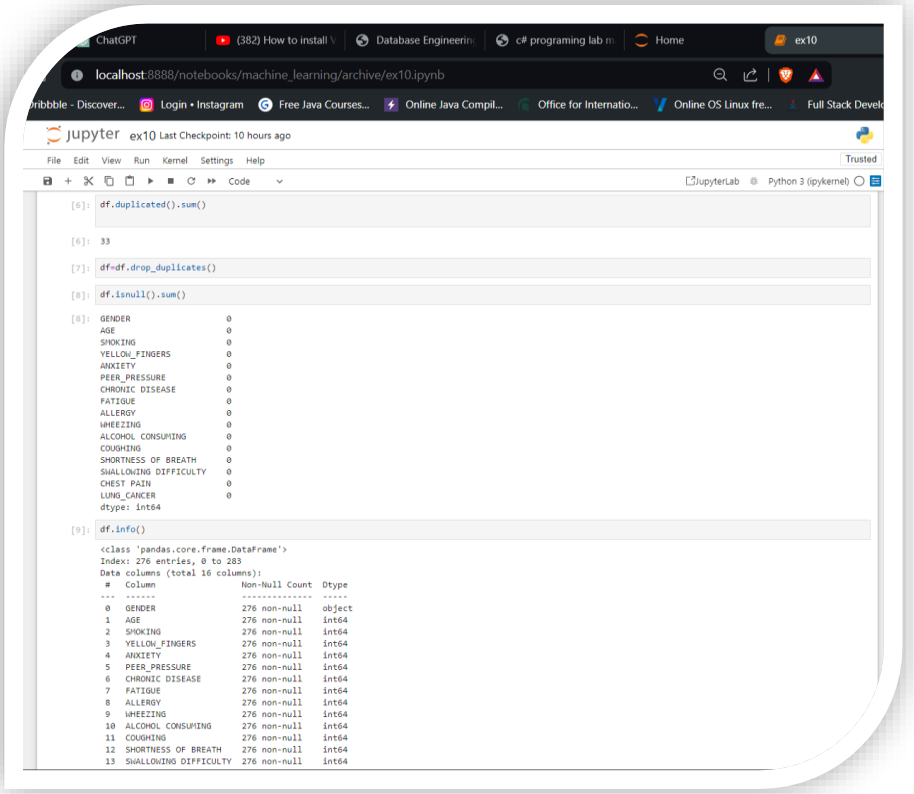
	GENDER	AGE	SMOKING	YELLOW_FINGERS	ANXIETY	PEER_PRESSURE	CHRONIC DISEASE	FATIGUE	ALLERGY	WHEEZING	ALCOHOL CONSUMING	COUGHING	SHORTNESS OF BREATH	SWA D
0	M	69	1	2	2	1	1	2	1	2	2	2	2	
1	M	74	2	1	1	1	2	2	2	1	1	1	2	
2	F	59	1	1	1	2	1	2	1	2	1	2	2	
3	M	63	2	2	2	1	1	1	1	1	2	1	1	
4	F	63	1	2	1	1	1	1	1	2	1	2	2	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
304	F	56	1	1	1	2	2	2	1	1	2	2	2	
305	M	70	2	1	1	1	1	2	2	2	2	2	2	
306	M	58	2	1	1	1	1	1	2	2	2	2	1	
307	M	67	2	1	2	1	1	2	2	1	2	2	2	
308	M	62	1	1	1	2	1	2	2	2	2	1	1	

309 rows × 16 columns

```
[5]: df.shape
```

[5]: (309, 16)

```
[6]: df.duplicated().sum()
```



```
[6]: df.duplicated().sum()
```

[6]: 33

```
[7]: df=df.drop_duplicates()
```

```
[8]: df.isnull().sum()
```

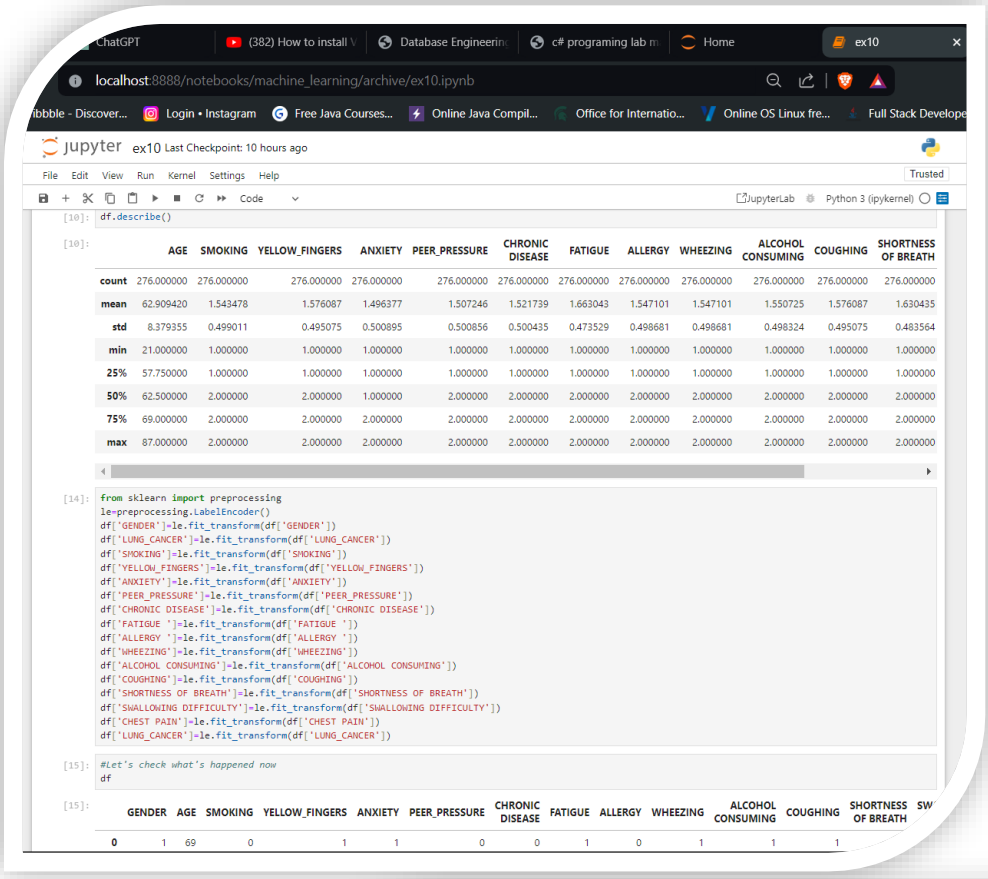
[8]:

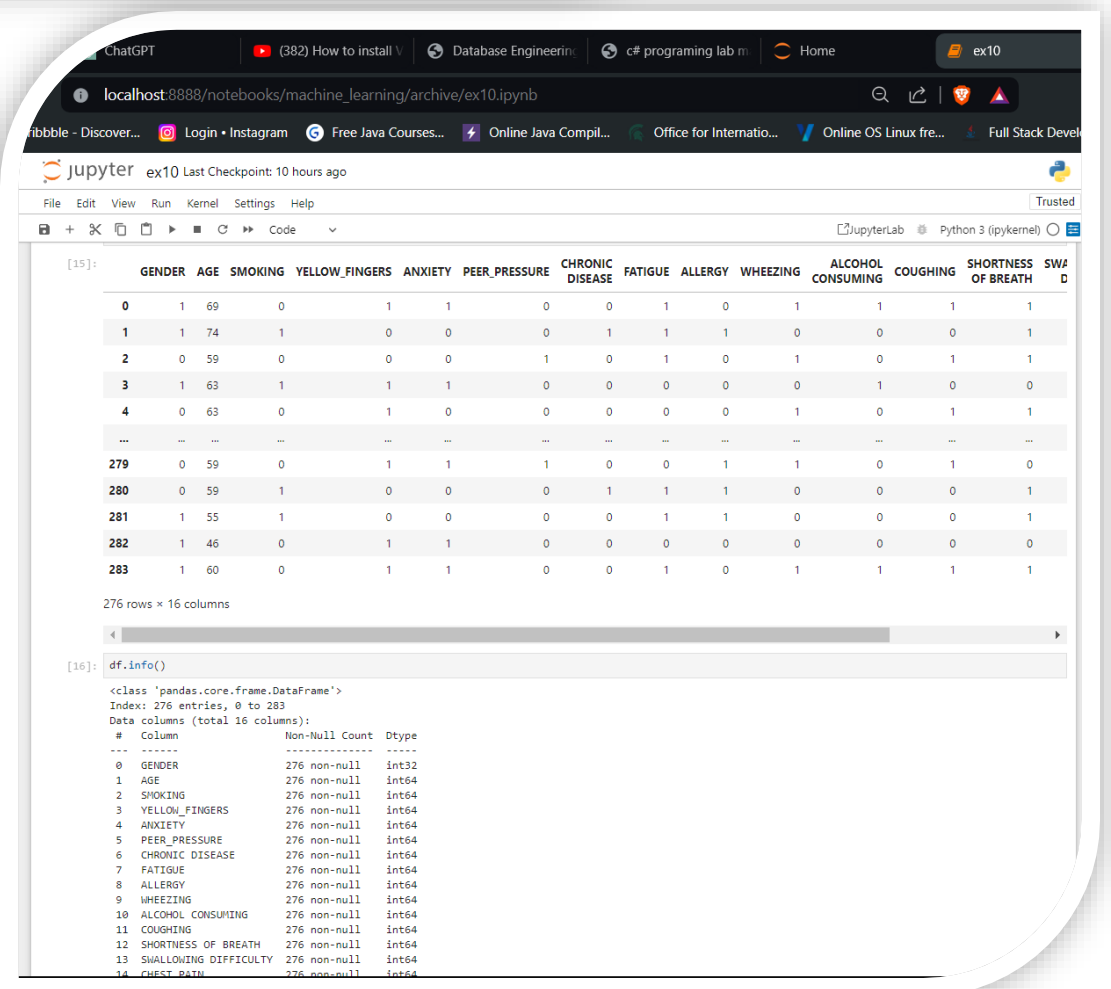
```
GENDER                  0
AGE                     0
SMOKING                 0
YELLOW_FINGERS          0
ANXIETY                 0
PEER_PRESSURE           0
CHRONIC DISEASE         0
FATIGUE                 0
ALLERGY                 0
WHEEZING                0
ALCOHOL CONSUMING       0
COUGHING                0
SHORTNESS OF BREATH     0
SWALLOWING DIFFICULTY   0
CHEST PAIN              0
LUNG_CANCER             0
dtype: int64
```

```
[9]: df.info()
```

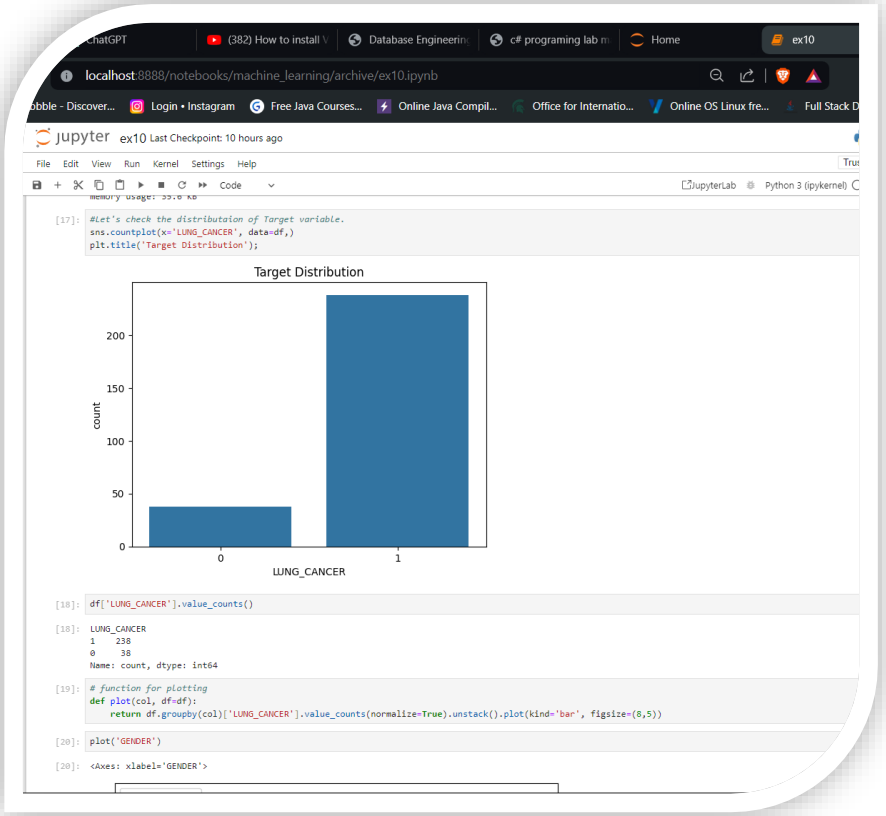
```
<class 'pandas.core.frame.DataFrame'>
Index: 276 entries, 0 to 283
Data columns (total 16 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   GENDER                 276 non-null    object
 1   AGE                    276 non-null    int64
 2   SMOKING                276 non-null    int64
 3   YELLOW_FINGERS         276 non-null    int64
 4   ANXIETY                276 non-null    int64
 5   PEER_PRESSURE          276 non-null    int64
 6   CHRONIC DISEASE        276 non-null    int64
 7   FATIGUE                276 non-null    int64
 8   ALLERGY                276 non-null    int64
 9   WHEEZING               276 non-null    int64
 10  ALCOHOL CONSUMING      276 non-null    int64
 11  COUGHING               276 non-null    int64
 12  SHORTNESS OF BREATH    276 non-null    int64
 13  SWALLOWING DIFFICULTY  276 non-null    int64
```

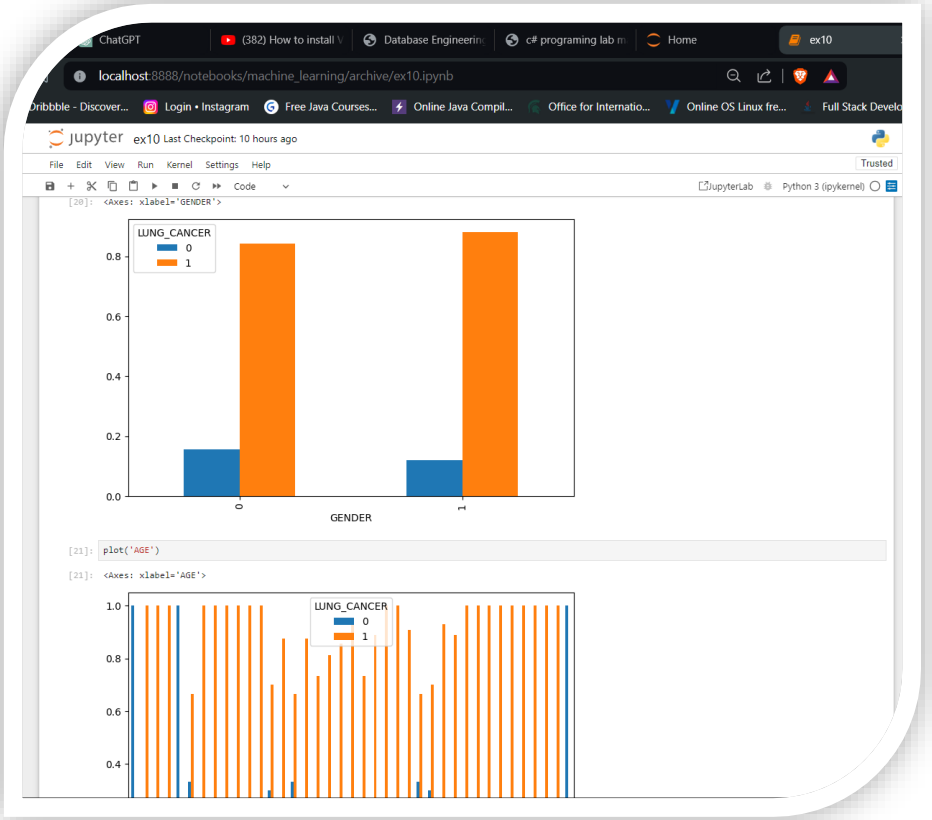
Tut 10


ChatGPT
(382) How to install V
Database Engineerin
c# programing lab m
Home
ex10
localhost:8888/notebooks/machine_learning/archive/ex10.ipynb
Dribbble - Discover...
Login • Instagram
Free Java Courses...
Online Java Compil...
Office for Internatio...
Online OS Linux fre...
Full Stack Develope
jupyter ex10 Last Checkpoint: 10 hours ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
[10]: df.describe()
[10]:
AGE SMOKING YELLOW_FINGERS ANXIETY PEER_PRESSURE CHRONIC DISEASE FATIGUE ALLERGY WHEEZING ALCOHOL CONSUMING COUGHING SHORTNESS OF BREATH
count 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000 276.000000
mean 62.909420 1.543478 1.576087 1.496377 1.507246 1.521739 1.663043 1.547101 1.547101 1.550725 1.576087 1.630435
std 8.379355 0.499011 0.495075 0.500895 0.500856 0.500435 0.473529 0.498681 0.498681 0.498324 0.495075 0.483564
min 21.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000
25% 57.750000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000
50% 62.500000 2.000000 2.000000 1.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000
75% 69.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000
max 87.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000 2.000000
[14]:
from sklearn import preprocessing
le=preprocessing.LabelEncoder()
df['GENDER']=le.fit_transform(df['GENDER'])
df['LUNG_CANCER']=le.fit_transform(df['LUNG_CANCER'])
df['SMOKING']=le.fit_transform(df['SMOKING'])
df['YELLOW_FINGERS']=le.fit_transform(df['YELLOW_FINGERS'])
df['ANXIETY']=le.fit_transform(df['ANXIETY'])
df['PEER_PRESSURE']=le.fit_transform(df['PEER_PRESSURE'])
df['CHRONIC DISEASE']=le.fit_transform(df['CHRONIC DISEASE'])
df['FATIGUE ']=le.fit_transform(df['FATIGUE '])
df['ALLERGY ']=le.fit_transform(df['ALLERGY '])
df['WHEEZING']=le.fit_transform(df['WHEEZING'])
df['ALCOHOL CONSUMING']=le.fit_transform(df['ALCOHOL CONSUMING'])
df['COUGHING']=le.fit_transform(df['COUGHING'])
df['SHORTNESS OF BREATH']=le.fit_transform(df['SHORTNESS OF BREATH'])
df['SWALLOWING DIFFICULTY']=le.fit_transform(df['SWALLOWING DIFFICULTY'])
df['CHEST PAIN']=le.fit_transform(df['CHEST PAIN'])
df['LUNG_CANCER']=le.fit_transform(df['LUNG_CANCER'])
[15]:
#Let's check what's happened now
df
[15]:
GENDER AGE SMOKING YELLOW_FINGERS ANXIETY PEER_PRESSURE CHRONIC DISEASE FATIGUE ALLERGY WHEEZING ALCOHOL CONSUMING COUGHING SHORTNESS OF BREATH SW
0 1 69 0 1 1 0 0 1 0 1 1 1

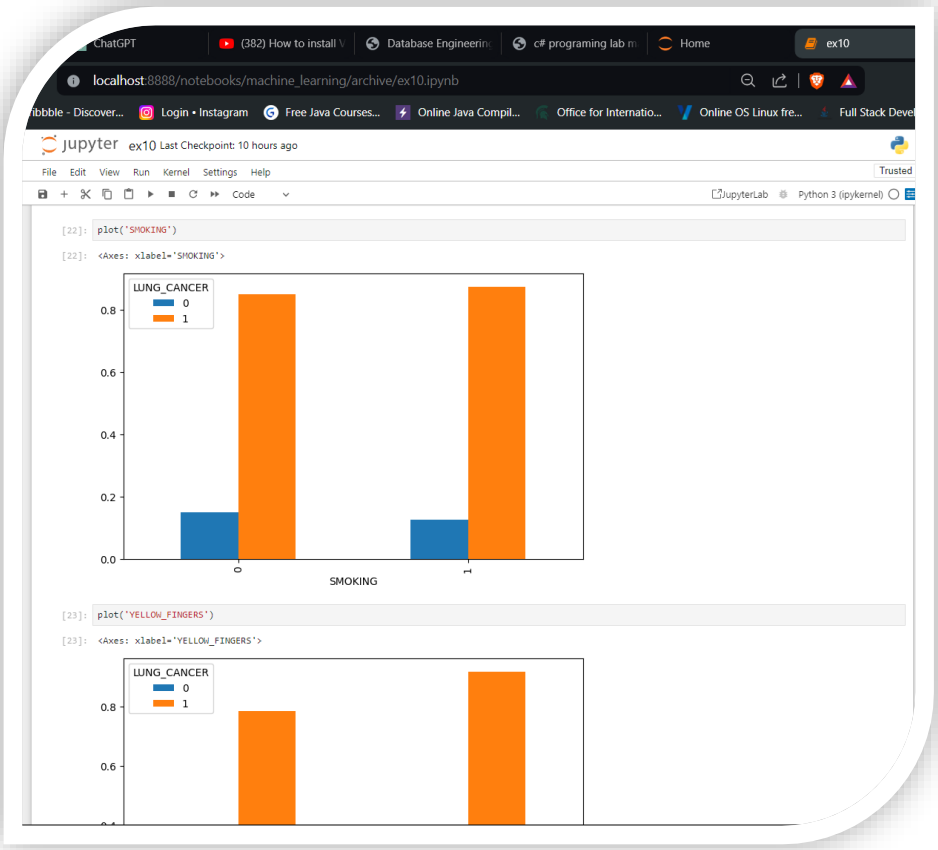

ChatGPT
(382) How to install V
Database Engineerin
c# programing lab m
Home
ex10
localhost:8888/notebooks/machine_learning/archive/ex10.ipynb
Dribbble - Discover...
Login • Instagram
Free Java Courses...
Online Java Compil...
Office for Internatio...
Online OS Linux fre...
jupyter ex10 Last Checkpoint: 10 hours ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
[15]:
GENDER AGE SMOKING YELLOW_FINGERS ANXIETY PEER_PRESSURE CHRONIC DISEASE FATIGUE ALLERGY WHEEZING ALCOHOL CONSUMING COUGHING SHORTNESS OF BREATH SWA D
0 1 69 0 1 1 0 0 1 0 1 1 1 1
1 1 74 1 0 0 0 1 1 1 0 0 0 1
2 0 59 0 0 0 1 0 1 0 1 0 1 1
3 1 63 1 1 1 0 0 0 0 0 1 0 0
4 0 63 0 1 0 0 0 0 0 1 0 1 1
... ...
279 0 59 0 1 1 1 0 0 1 1 0 1 0
280 0 59 1 0 0 0 1 1 1 0 0 0 1
281 1 55 1 0 0 0 0 1 1 0 0 0 1
282 1 46 0 1 1 0 0 0 0 0 0 0 0
283 1 60 0 1 1 0 0 1 0 1 1 1 1
276 rows × 16 columns
[16]:
df.info()
<class 'pandas.core.frame.DataFrame'>
Index: 276 entries, 0 to 283
Data columns (total 16 columns):
Column Non-Null Count Dtype
--- ------ -------------- -----
0 GENDER 276 non-null int32
1 AGE 276 non-null int64
2 SMOKING 276 non-null int64
3 YELLOW_FINGERS 276 non-null int64
4 ANXIETY 276 non-null int64
5 PEER_PRESSURE 276 non-null int64
6 CHRONIC DISEASE 276 non-null int64
7 FATIGUE 276 non-null int64
8 ALLERGY 276 non-null int64
9 WHEEZING 276 non-null int64
10 ALCOHOL CONSUMING 276 non-null int64
11 COUGHING 276 non-null int64
12 SHORTNESS OF BREATH 276 non-null int64
13 SWALLOWING DIFFICULTY 276 non-null int64

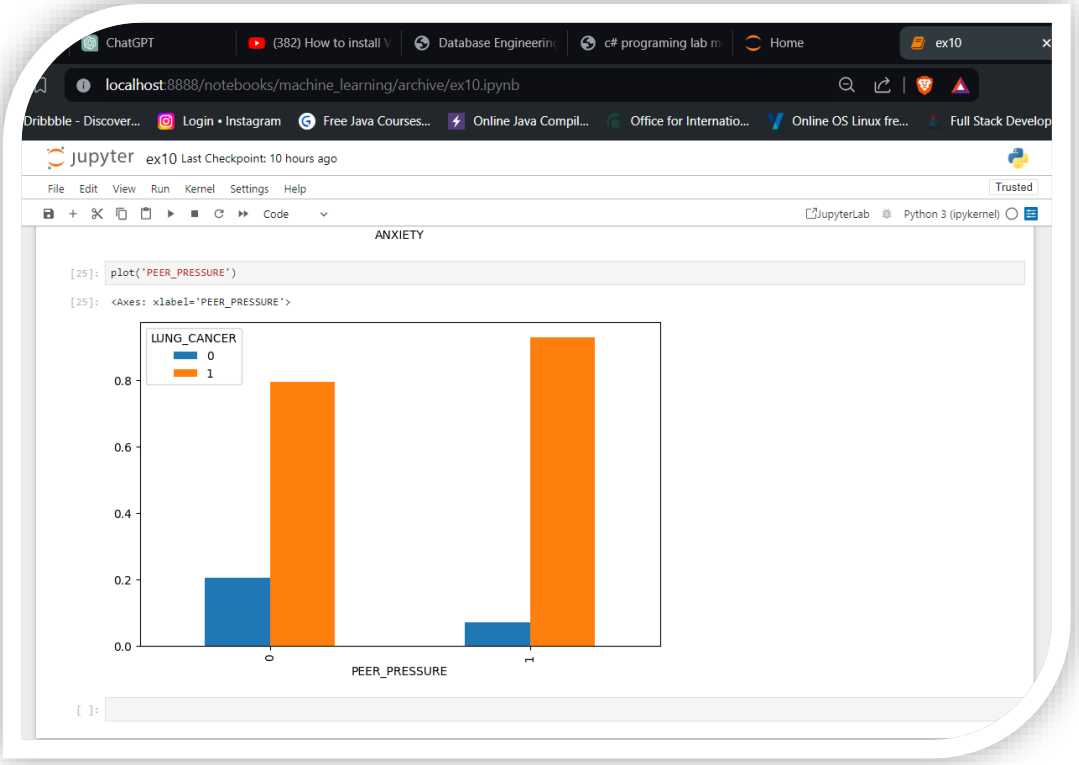
Tut 10


ChatGPT
(382) How to install
Database Engineerin
c# programing lab m
Home
ex10
localhost:8888/notebooks/machine_learning/archive/ex10.ipynb
Login • Instagram
Free Java Courses...
Online Java Compil...
Office for Internatio...
Online OS Linux fre...
jupyter ex10 Last Checkpoint: 10 hours ago
File Edit View Run Kernel Settings Help
Code
JupyterLab
Python 3 (ipykernel)
[17]: #Let's check the distribution of Target variable.
sns.countplot(x='LUNG_CANCER', data=df,)
plt.title('Target Distribution');
Target Distribution
count
LUNG_CANCER
[18]: df['LUNG_CANCER'].value_counts()
[18]: LUNG_CANCER
1 238
0 38
Name: count, dtype: int64
[19]: # function for plotting
def plot(col, df=df):
return df.groupby(col)['LUNG_CANCER'].value_counts(normalize=True).unstack().plot(kind='bar', figsize=(8,5))
[20]: plot('GENDER')
[20]: <Axes: xlabel='GENDER'>


ChatGPT
(382) How to install
Database Engineerin
c# programing lab m
Home
ex10
localhost:8888/notebooks/machine_learning/archive/ex10.ipynb
Login • Instagram
Free Java Courses...
Online Java Compil...
Office for Internatio...
Online OS Linux fre...
jupyter ex10 Last Checkpoint: 10 hours ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
[20]: <Axes: xlabel='GENDER'>
LUNG_CANCER
0
1
GENDER
[21]: plot('AGE')
[21]: <Axes: xlabel='AGE'>

Tut 10


ChatGPT
(382) How to install
Database Engineering
c# programing lab
Home
ex10
localhost:8888/notebooks/machine_learning/archive/ex10.ipynb
Login • Instagram
Free Java Courses...
Online Java Compil...
Office for Internatio...
Online OS Linux fre...
jupyter ex10 Last Checkpoint: 10 hours ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
[22]: plot('SMOKING')
[22]: <Axes: xlabel='SMOKING'>
LUNG_CANCER
0
1
0.0
0.2
0.4
0.6
0.8
0
1
SMOKING
[23]: plot('YELLOW_FINGERS')
[23]: <Axes: xlabel='YELLOW_FINGERS'>
LUNG_CANCER
0
1
0.6
0.8


ChatGPT
(382) How to install
Database Engineering
c# programing lab
Home
ex10
localhost:8888/notebooks/machine_learning/archive/ex10.ipynb
Dribbble - Discover...
Login • Instagram
Free Java Courses...
Online Java Compil...
Office for Internatio...
Online OS Linux fre...
Full Stack Develop
jupyter ex10 Last Checkpoint: 10 hours ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
ANXIETY
[25]: plot('PEER_PRESSURE')
[25]: <Axes: xlabel='PEER_PRESSURE'>
LUNG_CANCER
0
1
0.0
0.2
0.4
0.6
0.8
0
1
PEER_PRESSURE
[ ]: