

File Edit View Run Kernel Settings Help

Code Trusted

JupyterLab Python [conda env:base] \* 1

```
[15]: import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
[16]: df=pd.read_csv("diabetes (1).xls")
```

```
[17]: df.head()
```

```
[17]:
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	
0	6	148	72	35	0	33.6		0.627	50	1
1	1	85	66	29	0	26.6		0.351	31	0
2	8	183	64	0	0	23.3		0.672	32	1
3	1	89	66	23	94	28.1		0.167	21	0
4	0	137	40	35	168	43.1		2.288	33	1

```
[21]: df.describe()
```

```
[21]:
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000		768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.799479	31.992578		0.471876	33.240885	0.348958
std	3.369578	31.972618	19.355807	15.952218	115.244002	7.884160		0.331329	11.760232	0.476951
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000		0.078000	21.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000		0.243750	24.000000	0.000000



Search



EN IN



## jupyter Untitled74 Last Checkpoint: 9 minutes ago

File Edit View Run Kernel Settings Help

	25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000	0.243750	24.000000	0.000000
	50%	3.000000	117.000000	72.000000	23.000000	30.500000	32.000000	0.372500	29.000000	0.000000
	75%	6.000000	140.250000	80.000000	32.000000	127.250000	36.600000	0.626250	41.000000	1.000000
	max	17.000000	199.000000	122.000000	99.000000	846.000000	67.100000	2.420000	81.000000	1.000000

[23]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Pregnancies      768 non-null    int64  
 1   Glucose          768 non-null    int64  
 2   BloodPressure    768 non-null    int64  
 3   SkinThickness    768 non-null    int64  
 4   Insulin          768 non-null    int64  
 5   BMI              768 non-null    float64 
 6   DiabetesPedigreeFunction 768 non-null    float64 
 7   Age              768 non-null    int64  
 8   Outcome          768 non-null    int64  
dtypes: float64(2), int64(7)
memory usage: 54.1 KB
```

[27]: h1=df['Glucose']
h2=df['BloodPressure']
legend=['Glucose','BP']
plt.hist([h1,h2],color=['pink','purple'])
plt.xlabel("Level")
plt.ylabel("No. of Patient")
plt.legend(legend)
plt.title('Glucose vs BloodPressure')
plt.show()



Search



# jupyter Untitled74 Last Checkpoint: 9 minutes ago

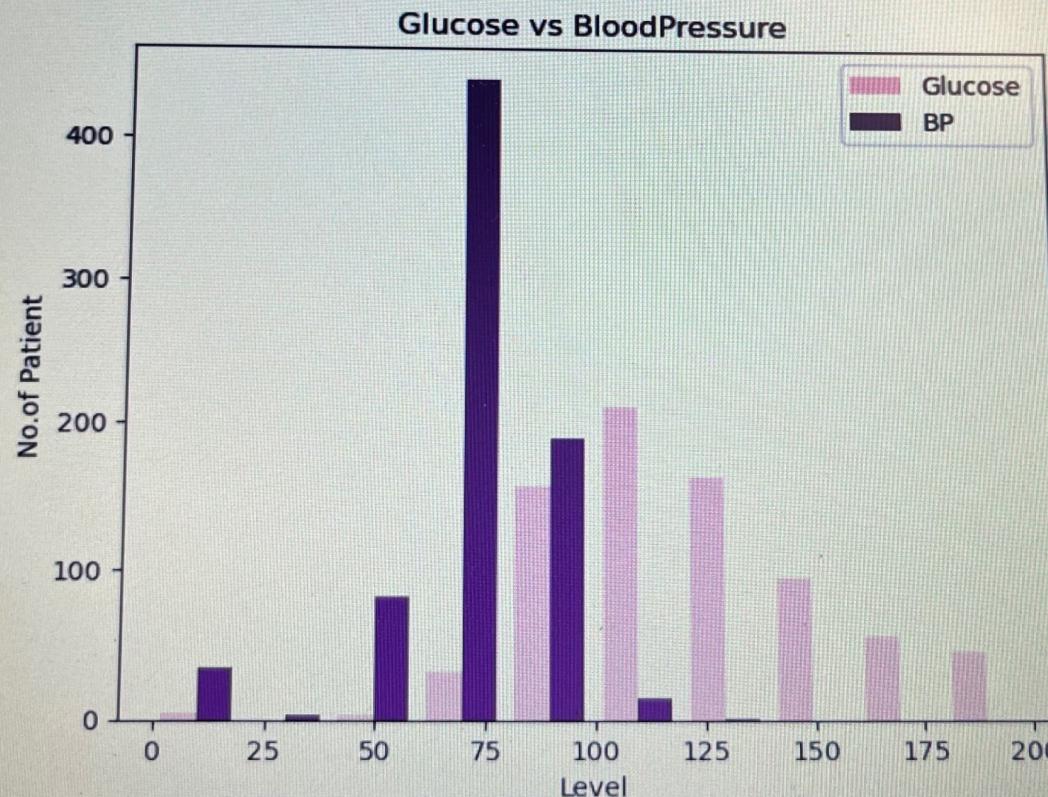
File Edit View Run Kernel Settings Help

Code

JupyterLab

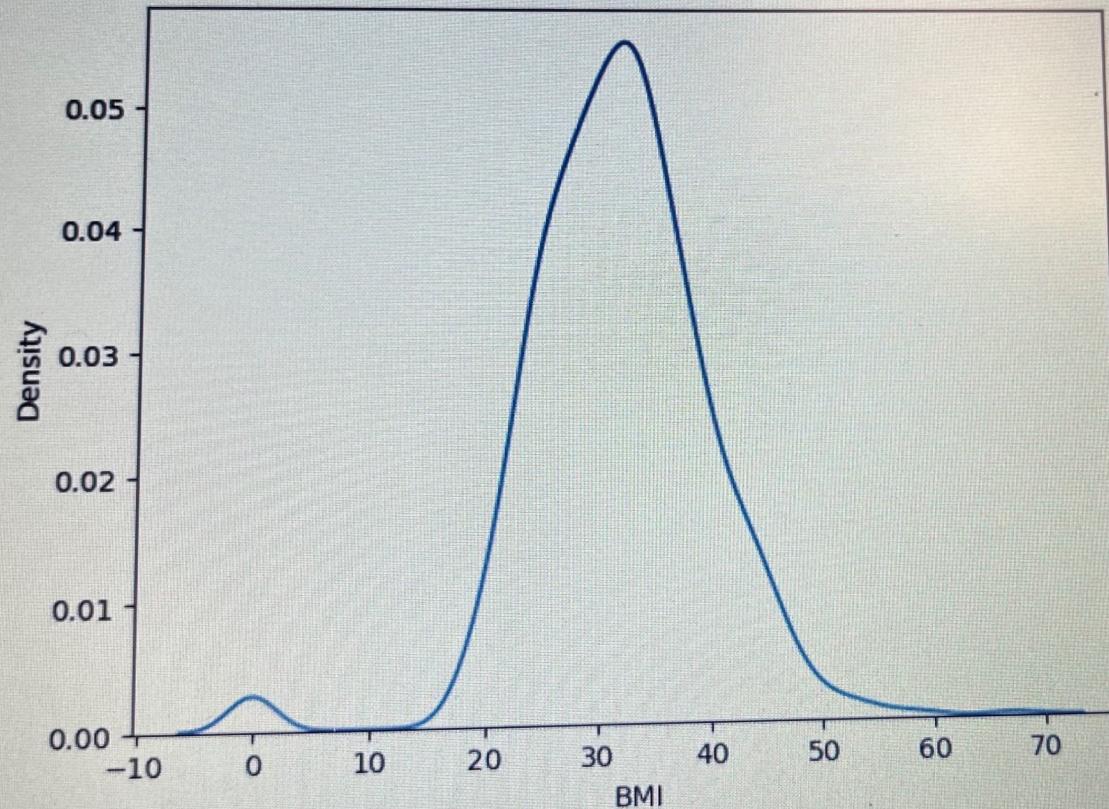
Python [conda env:base]

```
[27]: h1=df['Glucose']
       h2=df['BloodPressure']
       legend=['Glucose', 'BP']
       plt.hist([h1,h2],color=['pink','purple'])
       plt.xlabel("Level")
       plt.ylabel("No.of Patient")
       plt.legend(legend)
       plt.title('Glucose vs BloodPressure')
       plt.show()
```



Level

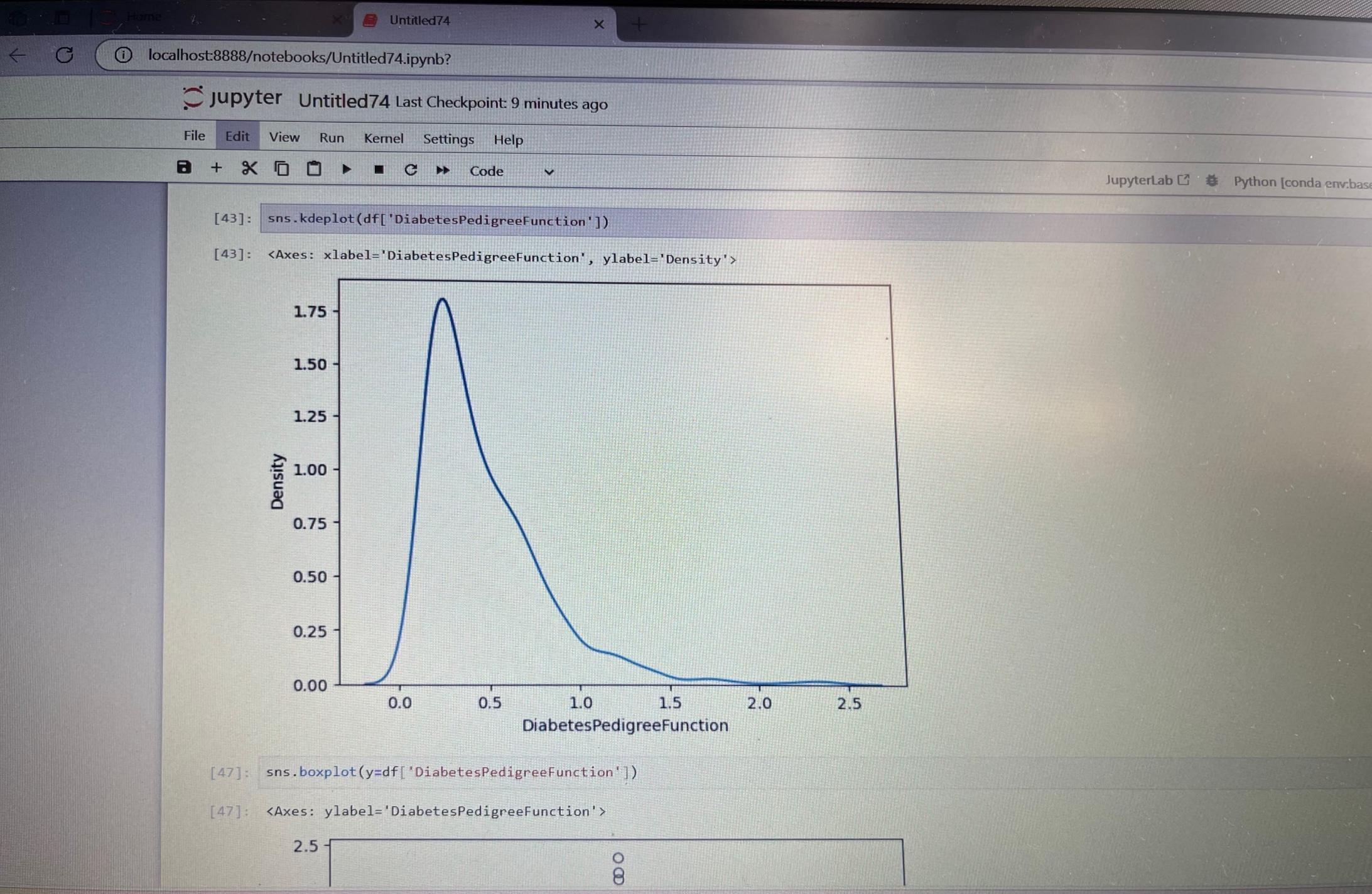
```
[39]: sns.kdeplot(df['BMI'])
plt.show()
```

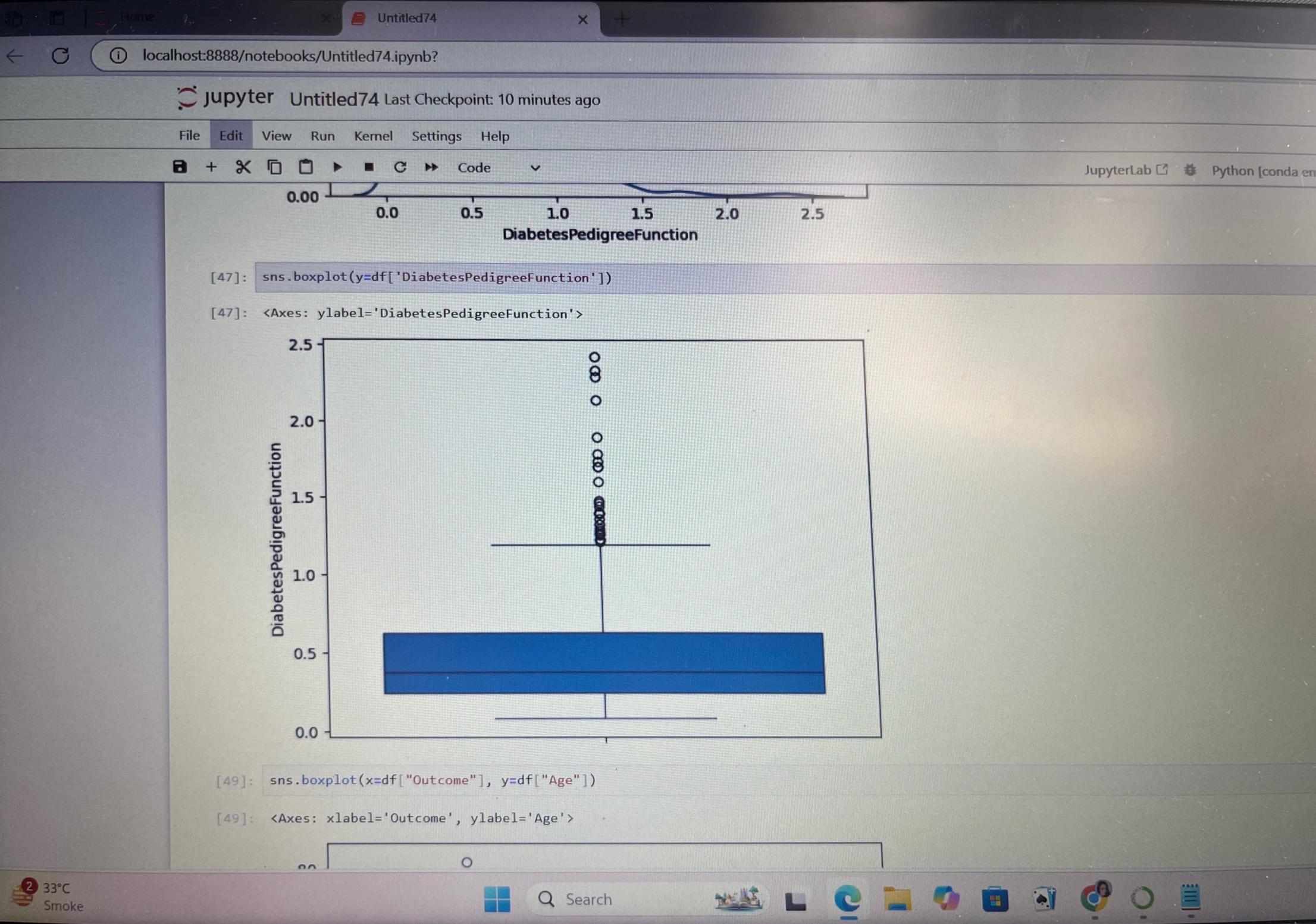


```
[43]: sns.kdeplot(df['DiabetesPedigreeFunction']).  
[43]: <Axes: xlabel='DiabetesPedigreeFunction', ylabel='Density'>
```

1.75







 jupyter Untitled74 Last Checkpoint: 10 minutes ago

File Edit View Run Kernel Settings Help

File + X □ ▶ □ C ▶ Code ▾

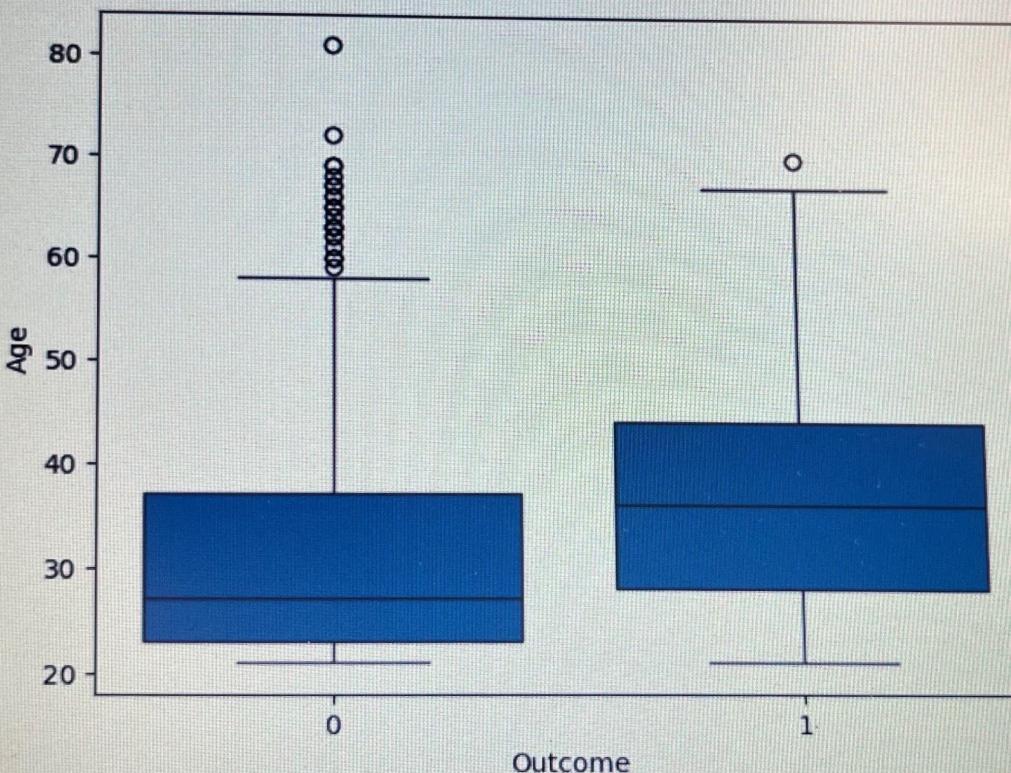
JupyterLab

Python [conda env]

0.0

[49]: sns.boxplot(x=df["Outcome"], y=df["Age"])

[49]: &lt;Axes: xlabel='Outcome', ylabel='Age'&gt;

[55]: corr = df.corr()  
ax = sns.heatmap(  
 corr,  
 vmin=-1, vmax=1, center=0,

Home Untitled74

localhost:8888/notebooks/Untitled74.ipynb?

## Jupyter Untitled74 Last Checkpoint: 10 minutes ago

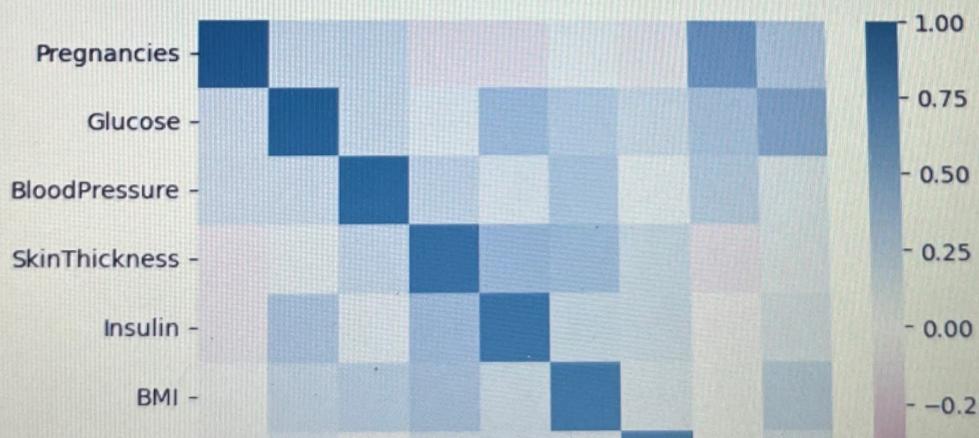
File Edit View Run Kernel Settings Help

Trusted

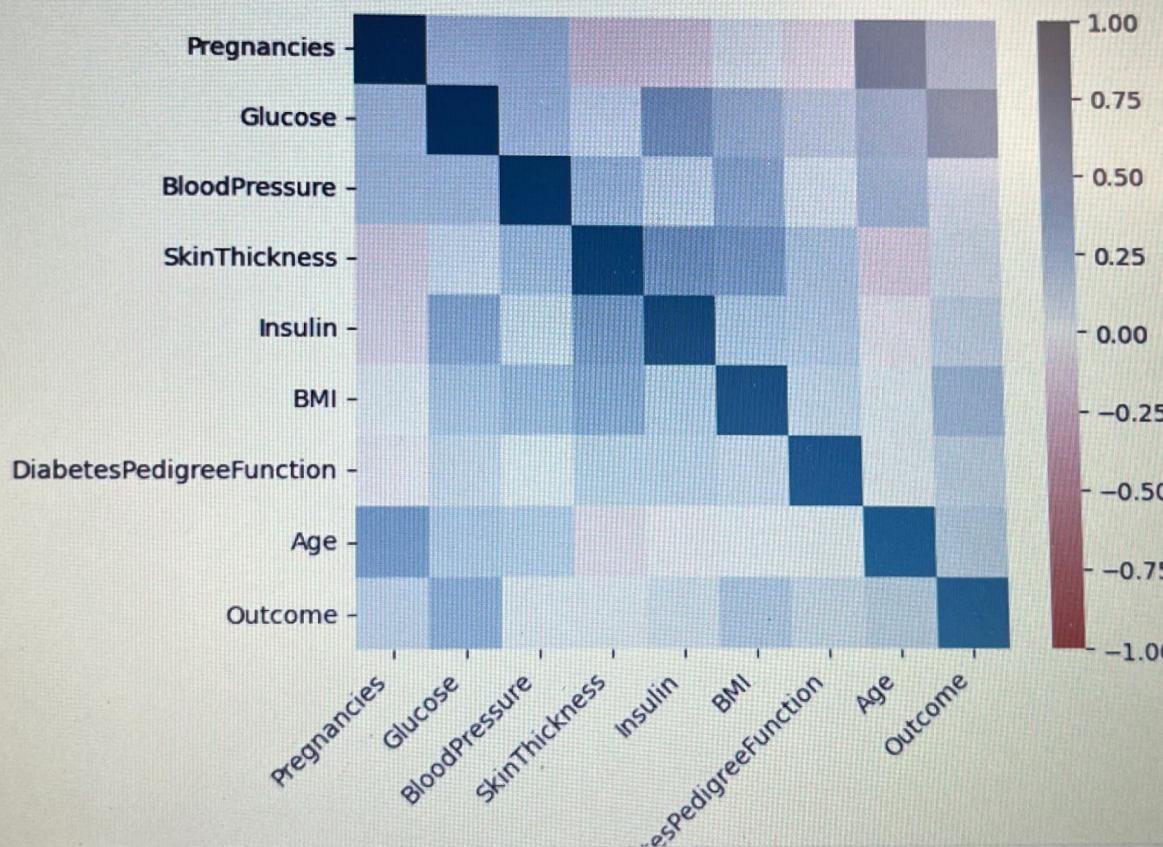
[55]:

```
corr = df.corr()
ax = sns.heatmap(
    corr,
    vmin=-1, vmax=1, center=0,
    cmap=sns.diverging_palette(20,220, n=200),
    square=True
)
ax.set_xticklabels(
    ax.get_xticklabels(),
    rotation=45,
    horizontalalignment='right')
```

```
[55]: [Text(0.5, 0, 'Pregnancies'),
       Text(1.5, 0, 'Glucose'),
       Text(2.5, 0, 'BloodPressure'),
       Text(3.5, 0, 'SkinThickness'),
       Text(4.5, 0, 'Insulin'),
       Text(5.5, 0, 'BMI'),
       Text(6.5, 0, 'DiabetesPedigreeFunction'),
       Text(7.5, 0, 'Age'),
       Text(8.5, 0, 'Outcome')]
```



```
Text(1.5, 0, 'Glucose'),  
Text(2.5, 0, 'BloodPressure'),  
Text(3.5, 0, 'SkinThickness'),  
Text(4.5, 0, 'Insulin'),  
Text(5.5, 0, 'BMI'),  
Text(6.5, 0, 'DiabetesPedigreeFunction'),  
Text(7.5, 0, 'Age'),  
Text(8.5, 0, 'Outcome'))]
```



```
[57]: df=df.sample(frac=1).reset_index(drop=True)

[59]: from sklearn.model_selection import train_test_split

[63]: x=df.iloc[:,df.columns =='Outcome']
y=df.iloc[:,df.columns == 'Outcome']

[65]: x_train, x_test, y_train, y_test=train_test_split(x,y,test_size=0.3)

[67]: x_train.shape

[67]: (537, 1)

[69]: y_test.shape

[69]: (231, 1)

[71]: from sklearn.preprocessing import minmax_scale

[73]: X_train=minmax_scale(x_train)

[75]: X_test=minmax_scale(x_test)

[77]: from sklearn import svm
from sklearn.metrics import accuracy_score,confusion_matrix

[79]: model1=svm.SVC()

[81]: model1.fit(X_train,y_train)
```

C:\Users\HP\anaconda3\Lib\site-packages\sklearn\utils\validation.py:1339: DataConversionWarning: A column-vector y was passed when  
x was expected. Please change the shape of y to (n\_samples, ) for example using ravel().

# Jupyter Untitled74 Last Checkpoint: 10 minutes ago

File Edit View Run Kernel Settings Help



JupyterLab Python [conda env:]

```
[81]: model1.fit(X_train,y_train)
```

```
C:\Users\HP\anaconda3\Lib\site-packages\sklearn\utils\validation.py:1339: DataConversionWarning: A column-vector y was passed when a 1d array was expected.
d. Please change the shape of y to (n_samples, ), for example using ravel().
y = column_or_1d(y, warn=True)
```

```
[81]: ▾ SVC ⓘ ?
```

```
SVC()
```

```
[83]: predict1=model1.predict(X_test)
```

```
[85]: svm=accuracy_score(y_test,predict1)
print(svm)
```

```
1.0
```

```
[87]: from sklearn.linear_model import LogisticRegression
```

```
[89]: lr=LogisticRegression()
lr.fit(X_train, y_train)
predict2=lr.predict(X_test)
```

```
C:\Users\HP\anaconda3\Lib\site-packages\sklearn\utils\validation.py:1339: DataConversionWarning: A column-vector y was passed when a 1d array was expected.
d. Please change the shape of y to (n_samples, ), for example using ravel().
y = column_or_1d(y, warn=True)
```

```
[91]: lr=accuracy_score(y_test,predict2)
print(lr)
```

```
1.0
```

```
[ ]:
```



Search

