Assignment 01

Name : Kiran K. Patil ID:211070904

Sem : 06 Course : Machine Learning

Kaggle :

Kaggle is a platform where data scientists and machine learning developers can compete against each other to solve complex data science problems. Companies and organizations also use Kaggle to host competitions in order to crowdsource solutions to problems they are facing.

The platform provides a variety of tools for participants, including a cloud-based workbench for developing and running code, as well as a large dataset repository. Kaggle also hosts a community forum where participants can ask questions, share ideas and collaborate on projects. Additionally, Kaggle offers a number of educational resources, including tutorials, courses, and a blog to help users learn data science and machine learning. It is also one of the most popular platform for data scientist to showcase their skills and find job opportunities

Pandas :

Pandas is a open-source library for the Python programming language for data manipulation and analysis. It provides data structures such as dataframes and series that allow for easy manipulation and analysis of large datasets. Pandas is built on top of the NumPy library and is widely used in the data science community for its powerful data manipulation capabilities. Some of the main features of Pandas include.

Pandas is a powerful library with a wide range of functions for data manipulation and analysis. Here is a list of some of the most commonly used functions in Pandas, along with a brief description of what they do:

Pandas Functions :

- read_csv(): reads a CSV file and returns a dataframe
- to_csv(): writes a dataframe to a CSV file
- head(): returns the first n rows of a dataframe (default is 5)
- tail(): returns the last n rows of a dataframe (default is 5)
- shape: returns the dimensions of a dataframe (rows, columns)
- info(): returns information about a dataframe, including the data types of each column and the number of non-null values
- describe(): returns basic statistics for each numeric column in a dataframe
- columns: returns the column labels of a dataframe
- index: returns the index (row labels) of a dataframe
- value_counts(): returns the frequency counts for each unique value in a column
- sort_values(): sorts a dataframe by one or more columns
- groupby(): groups a dataframe by one or more columns and applies a function to each group
- merge(): merges two dataframes on one or more columns
- pivot_table(): creates a pivot table from a dataframe
- melt(): "melts" a dataframe and returns a new, reshaped dataframe
- stack(): "stacks" the columns of a dataframe and returns a new, reshaped dataframe

- unstack(): "unstacks" the rows of a dataframe and returns a new, reshaped dataframe
- crosstab(): creates a cross-tabulation (frequency table) of two or more factors

Seaborn :

Seaborn is a Python data visualization library based on Matplotlib. It is built on top of Matplotlib and allows for easy creation of beautiful, informative, and highly-customizable statistical graphics. Some of the main features of Seaborn include:

Here is a list of some commonly used functions in Seaborn:

- sns.lineplot(): creates a line plot
- sns.barplot(): creates a bar plot
- sns.scatterplot(): creates a scatter plot
- sns.histplot(): creates a histogram
- sns.boxplot(): creates a box plot
- sns.violinplot(): creates a violin plot
- sns.catplot(): creates a categorical plot
- sns.pairplot(): creates a pair plot
- sns.jointplot(): creates a joint plot
- sns.heatmap(): creates a heat map
- sns.regplot(): creates a regression plot
- sns.kdeplot(): creates a kernel density estimate plot
- sns.lmplot(): creates a scatter plot with linear regression line
- sns.countplot(): creates a bar plot of counts
- sns.despine(): removes the top and right spines from the plot
- sns.set_style(): sets the background theme of the plot
- sns.set_context(): sets the context of the plot (paper, notebook, talk, poster)

Matplotlib :

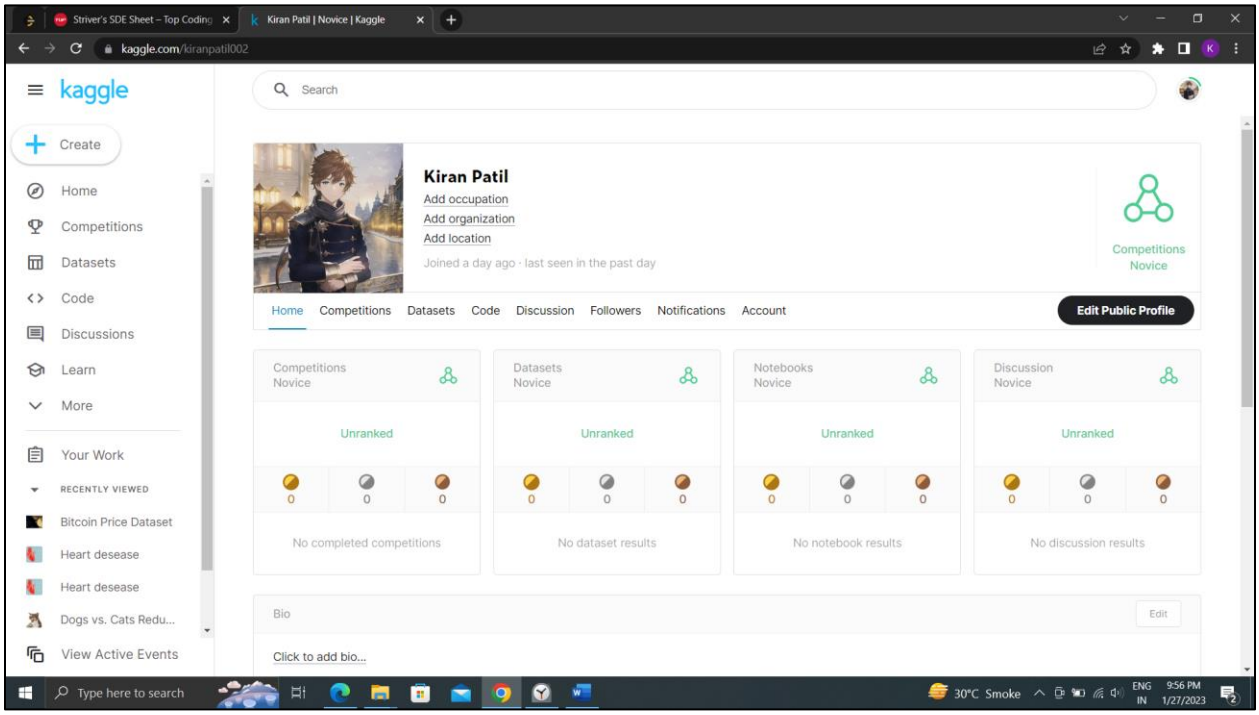
Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy

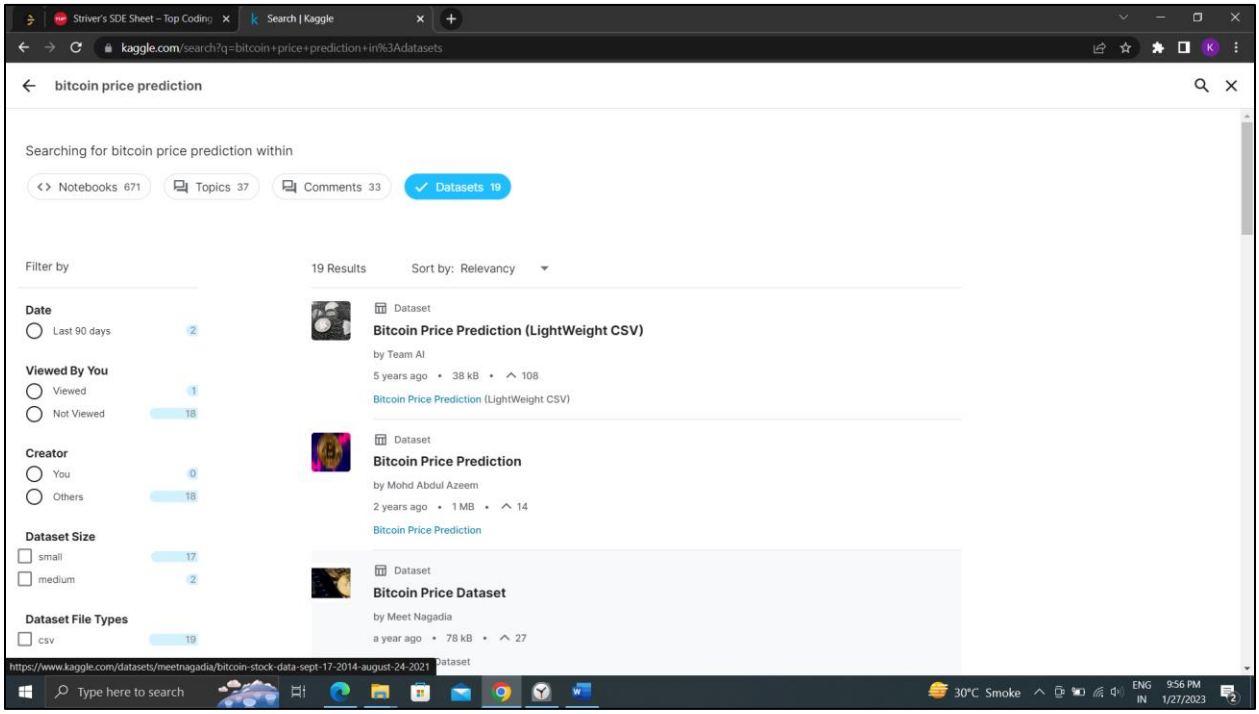
Matplotlib is a powerful data visualization library that provides a wide variety of customizable plots, but it can be complex to use for creating more advanced visualizations. Seaborn is built on top of Matplotlib and provides a simpler, more convenient interface for creating many common types of plots, especially those used in statistical data visualization.

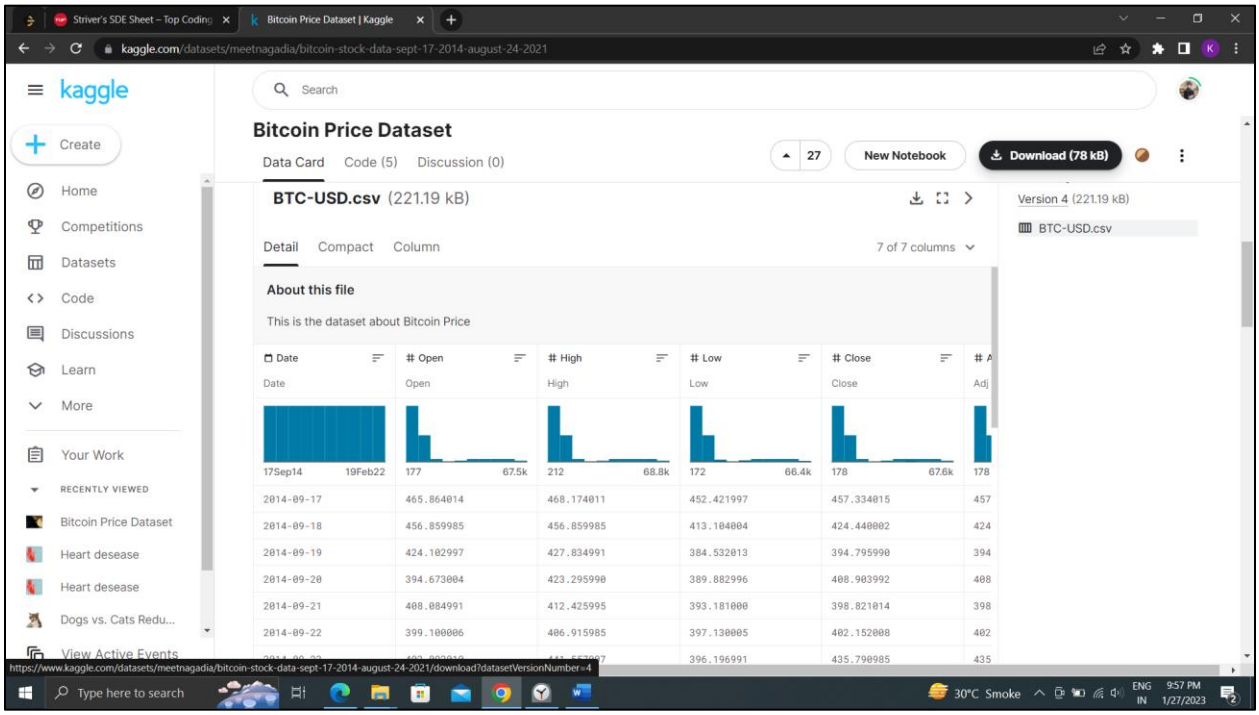
Here is a list of some commonly used functions in Matplotlib:

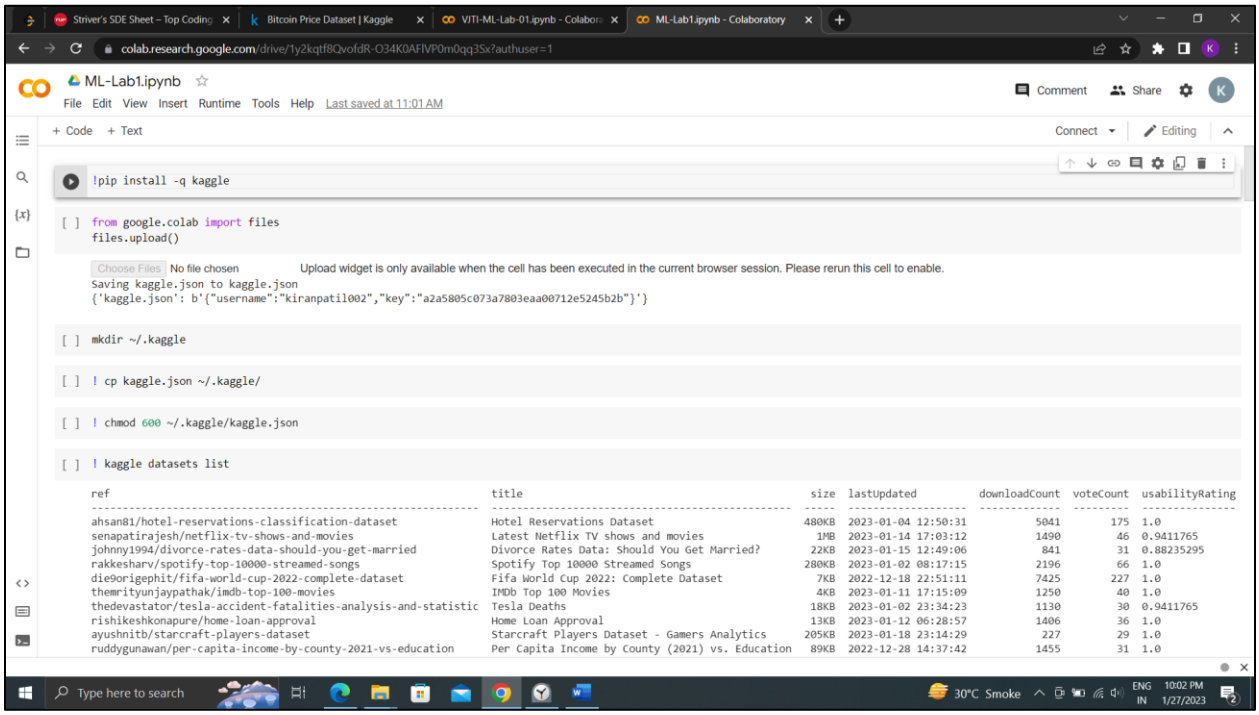
- plt.plot(): creates a line plot
- plt.scatter(): creates a scatter plot
- plt.bar(): creates a bar plot
- plt.hist(): creates a histogram
- plt.pie(): creates a pie chart
- plt.boxplot(): creates a box plot
- plt.stem(): creates a stem plot
- plt.quiver(): creates a vector field plot
- plt.semilogx(): creates a semilogarithmic plot of x-axis
- plt.semilogy(): creates a semilogarithmic plot of y-axis
- plt.loglog(): creates a log-log plot
- plt.polar(): creates a polar plot
- plt.contour(): creates a contour plot
- plt.imshow(): displays an image
- plt.colorbar(): adds a colorbar to a plot
- plt.xlabel(): adds a label to the x-axis
- plt.ylabel(): adds a label to the y-axis
- plt.title(): adds a title to the plot
- plt.legend(): adds a legend to the plot
- plt.show(): displays the plot

Setting up an account and download the dataset from Kaggle :






Striver's SDE Sheet – Top Coding
Bitcoin Price Dataset | Kaggle
kaggle.com/datasets/meetnagadia/bitcoin-stock-data-sept-17-2014-august-24-2021
kaggle
Search
Create
Home
Competitions
Datasets
Code
Discussions
Learn
More
Your Work
RECENTLY VIEWED
Bitcoin Price Dataset
Heart desease
Heart desease
Dogs vs. Cats Redu...
View Active Events
Bitcoin Price Dataset
Data Card
Code (5)
Discussion (0)
27
New Notebook
Download (78 kB)
BTC-USD.csv (221.19 kB)
Version 4 (221.19 kB)
BTC-USD.csv
Detail
Compact
Column
7 of 7 columns
About this file
This is the dataset about Bitcoin Price
Date
Open
High
Low
Close
17Sep14
19Feb22
177
67.5k
212
68.8k
172
66.4k
178
67.6k
2014-09-17 465.864014 468.174011 452.421997 457.334015
2014-09-18 456.859985 456.859985 413.104004 424.440002
2014-09-19 424.102997 427.834991 384.532013 394.795990
2014-09-20 394.673004 423.295990 389.882996 408.903992
2014-09-21 408.084991 412.425995 393.181000 398.821014
2014-09-22 399.100006 406.915985 397.130005 402.152008
396.196991 435.790985
https://www.kaggle.com/datasets/meetnagadia/bitcoin-stock-data-sept-17-2014-august-24-2021/download?datasetVersionNumber=4
Type here to search
30°C Smoke
ENG IN
9:57 PM
1/27/2023


Striver's SDE Sheet – Top Coding
Bitcoin Price Dataset | Kaggle
VITI-ML-Lab-01.ipynb - Colabora
ML-Lab1.ipynb - Colaboratory
colab.research.google.com/drive/1y2kqtf8QvofdR-O34K0AFlVP0m0qq3Sx?authuser=1
ML-Lab1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 11:01 AM
Comment
Share
+ Code + Text
Connect
Editing
!pip install -q kaggle
from google.colab import files
files.upload()
Choose Files No file chosen
Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving kaggle.json to kaggle.json
{'kaggle.json': b'{"username":"kiranpatil002","key":"a2a5805c073a7803eaa00712e5245b2b"}'}
mkdir ~/.kaggle
! cp kaggle.json ~/.kaggle/
! chmod 600 ~/.kaggle/kaggle.json
! kaggle datasets list
ref title size lastUpdated downloadCount voteCount usabilityRating
ahsan81/hotel-reservations-classification-dataset Hotel Reservations Dataset 480KB 2023-01-04 12:50:31 5041 175 1.0
senapatirajesh/netflix-tv-shows-and-movies Latest Netflix TV shows and movies 1MB 2023-01-14 17:03:12 1490 46 0.9411765
johnny1994/divorce-rates-data-should-you-get-married Divorce Rates Data: Should You Get Married? 22KB 2023-01-15 12:49:06 841 31 0.88235295
rakkesharv/spotify-top-10000-streamed-songs Spotify Top 10000 Streamed Songs 280KB 2023-01-02 08:17:15 2196 66 1.0
die9origephit/fifa-world-cup-2022-complete-dataset Fifa World Cup 2022: Complete Dataset 7KB 2022-12-18 22:51:11 7425 227 1.0
themrityunjaypathak/imdb-top-100-movies IMDb Top 100 Movies 4KB 2023-01-11 17:15:09 1250 40 1.0
thedevastator/tesla-accident-fatalities-analysis-and-statistic Tesla Deaths 18KB 2023-01-02 23:34:23 1130 30 0.9411765
rishikeshkonapure/home-loan-approval Home Loan Approval 13KB 2023-01-12 06:28:57 1406 36 1.0
ayushnitb/starcraft-players-dataset Starcraft Players Dataset - Gamers Analytics 205KB 2023-01-18 23:14:29 227 29 1.0
ruddygunawan/per-capita-income-by-county-2021-vs-education Per Capita Income by County (2021) vs. Education 89KB 2022-12-28 14:37:42 1455 31 1.0
Type here to search
30°C Smoke
ENG IN
10:02 PM
1/27/2023

Striver's SDE Sheet – Top Coding | Bitcoin Price Dataset | Kaggle | VITI-ML-Lab-01.ipynb - Colaborat | ML-Lab1.ipynb - Colaboratory

colab.research.google.com/drive/1y2kqtf8QvofdR-O34K0AFlVP0m0qq3Sx?authuser=1#scrollTo=qQJ-NSGhxJD1

ML-Lab1.ipynb

File Edit View Insert Runtime Tools Help Last saved at 11:01 AM

Comment Share

+ Code + Text

Connect Editing

```
[ ] !kaggle competitions download -c heart-desease

    403 - Forbidden
```

```
from google.colab import files
files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving heart-desease-publicleaderboard.csv to heart-desease-publicleaderboard.csv
{'heart-desease-publicleaderboard.csv': b'\xef\xbb\xbfTeamId,TeamName,SubmissionDate,Score\n9535242,"Ivan Novoselov","2022-12-01 10:29:58",1.00000\n9528960,timelessowl,"2022-11-30
09:39:52",1.00000\n9527112,AndreyS,"2022-12-01 10:43:28",1.00000\n'}
```

```
[ ] !ls

    heart-desease-publicleaderboard.csv  sample_data
```

```
[ ] from google.colab import files
    files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving BTC-USD.csv to BTC-USD.csv
{'BTC-USD.csv': b'Date,Open,High,Low,Close,Adj Close,Volume\n2014-09-17,465.864014,468.174011,452.421997,457.334015,457.334015,21056800\n2014-09-
18,456.859985,456.859985,413.104004,424.440002,424.440002,34483200\n2014-09-19,424.102997,427.834991,384.532013,394.795990,394.795990,37919700\n2014-09-
20,394.673004,423.295990,389.882996,408.903992,408.903992,36863600\n2014-09-21,408.084991,412.425995,393.181000,398.821014,398.821014,26580100\n2014-09-
22,399.100006,406.915985,397.130005,402.152008,402.152008,24127600\n2014-09-23,402.092010,441.557007,396.196991,435.790985,435.790985,45099500\n2014-09-
24,435.751007,436.112000,421.131989,423.204987,423.204987,30627700\n2014-09-25,423.156006,423.519989,409.467987,411.574005,411.574005,26814400\n2014-09-
26,411.428986,414.937988,400.009003,404.424988,404.424988,21460800\n2014-09-27,403.556000,406.622986,397.372009,399.519989,399.519989,15029300\n2014-09-
28,399.471008,401.016998,374.332001,377.181000,377.181000,23613300\n2014-09-29,376.928009,385.210999,372.239990,375.467010,375.467010,32497700\n2014-09-
30,376.088013,390.976990,373.442993,386.944000,386.944000,34707300\n2014-10-01,387.427002,391.378998,380.779999,383.614990,383.614990,26229400\n2014-10-
02,383.988007,385.497009,372.946014,375.071991,375.071991,21777700\n2014-10-03,375.181000,377.695007,357.859009,359.511993,359.511993,30901200\n2014-10-
04,359.891998,364.487000,325.885986,328.865997,328.865997,47236500\n2014-10-05,328.915985,341.800995,289.295990,320.510010,320.510010,83308096\n2014-10-
```

Striver's SDE Sheet – Top Coding | Bitcoin Price Dataset | Kaggle | VITI-ML-Lab-01.ipynb - Colaborat | ML-Lab1.ipynb - Colaboratory

colab.research.google.com/drive/1y2kqtf8QvofdR-O34K0AFlVP0m0qq3Sx?authuser=1#scrollTo=qQJ-NSGhxJD1

ML-Lab1.ipynb

File Edit View Insert Runtime Tools Help Last saved at 11:01 AM

Comment Share

+ Code + Text

Connect Editing

```
18,40552.132813,40929.152344,39637.617188,40030.976563,40030.976563,23310007704\n2022-02-19,40022.132813,40246.027344,40010.867188,40126.429688,40126.429688,22263900160'}
```

```
[ ] !ls

    BTC-USD.csv  heart-desease-publicleaderboard.csv  sample_data
```

```
import pandas as pd

data_frame = pd.read_csv('BTC-USD.csv', names = ['Date', 'Open', 'High', 'Low', 'Close','Adj Close','Volume'])
data_frame.head()
```

	Date	Open	High	Low	Close	Adj Close	Volume
0	Date	Open	High	Low	Close	Adj Close	Volume
1	2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
2	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
3	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700
4	2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600

```
[ ] data_frame.rename(columns=data_frame.iloc[0]).drop(data_frame.index[0])
```

	Date	Open	High	Low	Close	Adj Close	Volume
1	2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
2	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
3	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700

Striver's SDE Sheet – Top Coding | Bitcoin Price Dataset | Kaggle | VITI-ML-Lab-01.ipynb - Colabora | ML-Lab1.ipynb - Colaboratory

colab.research.google.com/drive/1y2kqtf8QvofdR-O34K0AFlVP0m0qq3Sx?authuser=1#scrollTo=qQJ-NSGhxJD1

ML-Lab1.ipynb

File Edit View Insert Runtime Tools Help Last saved at 11:01 AM

Comment Share

+ Code + Text

Connect Editing

```
[ ] data_frame.describe()
```

	Date	Open	High	Low	Close	Adj Close	Volume
count	2714	2714	2714	2714	2714	2714	2714
unique	2714	2710	2711	2713	2711	2711	2714
top	Date	17760.300781	6536.919922	429.076996	6741.750000	6741.750000	Volume
freq	1	2	2	2	2	2	1

```
print(data_frame)
```

```
           Date          Open          High           Low         Close  \
0          Date          Open          High           Low         Close
1    2014-09-17    465.864014    468.174011    452.421997    457.334015
2    2014-09-18    456.859985    456.859985    413.104004    424.440002
3    2014-09-19    424.102997    427.834991    384.532013    394.795990
4    2014-09-20    394.673004    423.295990    389.882996    408.903992
...         ...           ...           ...           ...           ...
2709 2022-02-15  42586.464844  44667.218750  42491.035156  44575.203125
2710 2022-02-16  44578.277344  44578.277344  43456.691406  43961.859375
2711 2022-02-17  43937.070313  44132.972656  40249.371094  40538.011719
2712 2022-02-18  40552.132813  40929.152344  39637.617188  40030.976563
2713 2022-02-19  40022.132813  40246.027344  40010.867188  40126.429688

         Adj Close       Volume
0        Adj Close       Volume
1       457.334015     21056800
2       424.440002     34483200
3       394.795990     37919700
4       408.903992     36863600
```

Striver's SDE Sheet – Top Coding | Bitcoin Price Dataset | Kaggle | VITI-ML-Lab-01.ipynb - Colabora | ML-Lab1.ipynb - Colaboratory

colab.research.google.com/drive/1y2kqtf8QvofdR-O34K0AFlVP0m0qq3Sx?authuser=1#scrollTo=qQJ-NSGhxJD1

ML-Lab1.ipynb

File Edit View Insert Runtime Tools Help Last saved at 11:01 AM

Comment Share

+ Code + Text

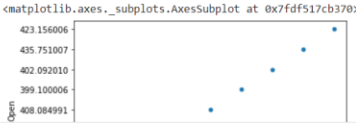
Connect Editing

```
[ ] print(df)
```

```
         Date        Open        High         Low       Close   Adj Close  \
1  2014-09-17  465.864014  468.174011  452.421997  457.334015  457.334015
2  2014-09-18  456.859985  456.859985  413.104004  424.440002  424.440002
3  2014-09-19  424.102997  427.834991  384.532013  394.795990  394.795990
4  2014-09-20  394.673004  423.295990  389.882996  408.903992  408.903992
5  2014-09-21  408.084991  412.425995  393.181000  398.821014  398.821014
6  2014-09-22  399.100006  406.915985  397.130005  402.152008  402.152008
7  2014-09-23  402.092010  441.557007  396.196991  435.790985  435.790985
8  2014-09-24  435.751007  436.112000  421.131989  423.204987  423.204987
9  2014-09-25  423.156006  423.519989  409.467987  411.574005  411.574005

     Volume
1  21056800
2  34483200
3  37919700
4  36863600
5  26580100
6  24127600
7  45099500
8  30627700
9  26814400
```

```
df.plot(kind="scatter", x='Date', y='Open')
```

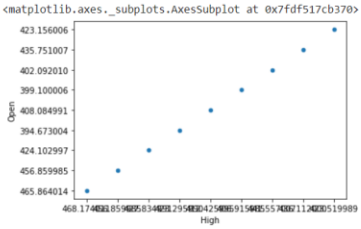
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdf517cb370>
```



```
2  34483200
3  37919700
4  36863600
5  26580100
6  24127600
7  45099500
8  30627700
9  26814400
```

```
df.plot(kind="scatter", x='Date', y='Open')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdf517cb370>
```




```
import seaborn as sns
```

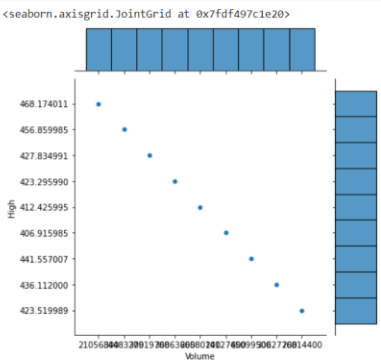
```
sns.jointplot(x='Volume', y='High', data=df, height=6)
```

```
<seaborn.axisgrid.JointGrid at 0x7fdf497c1e20>
```

```
import seaborn as sns
```

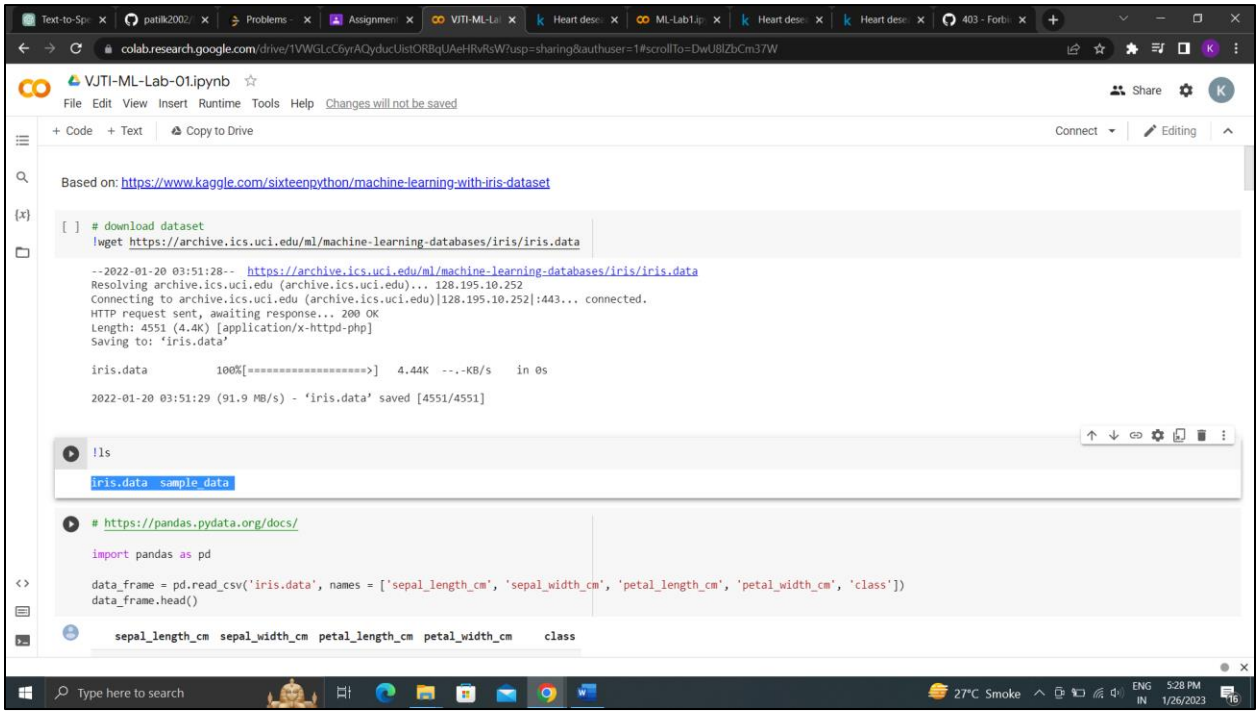
```
sns.jointplot(x='Volume', y='High', data=df, height=6)
```

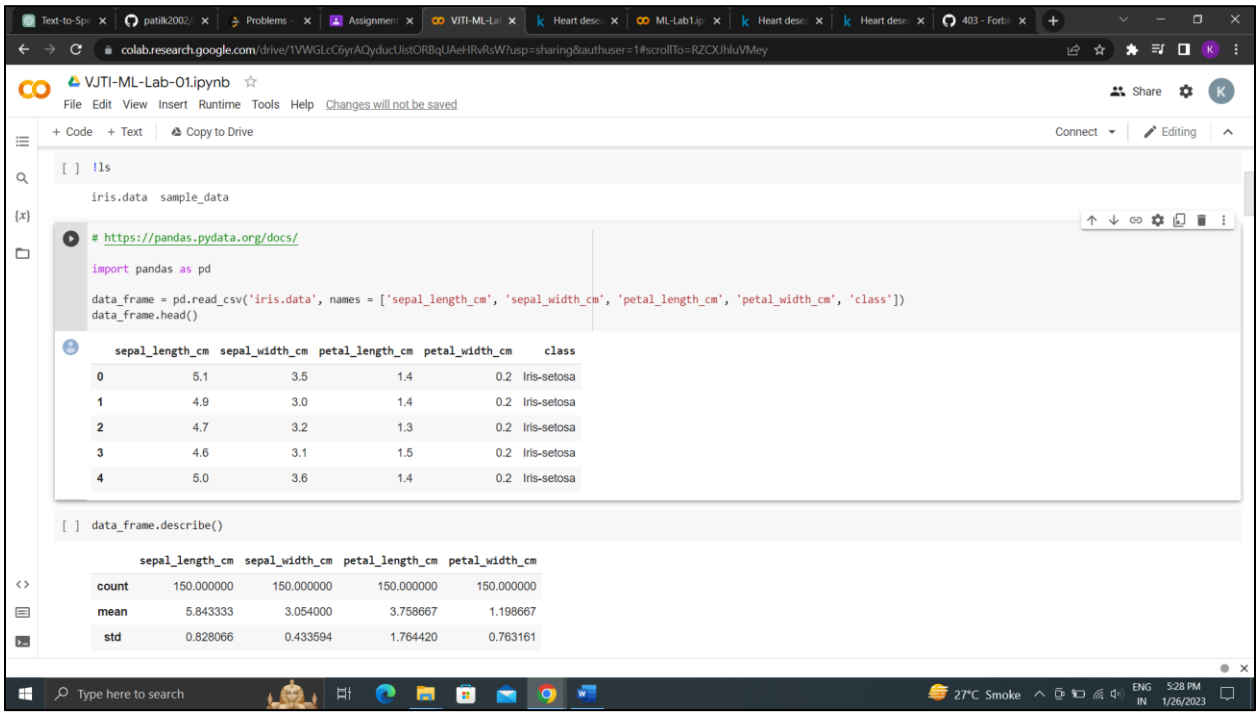
```
<seaborn.axisgrid.JointGrid at 0x7fdf497c1e20>
```

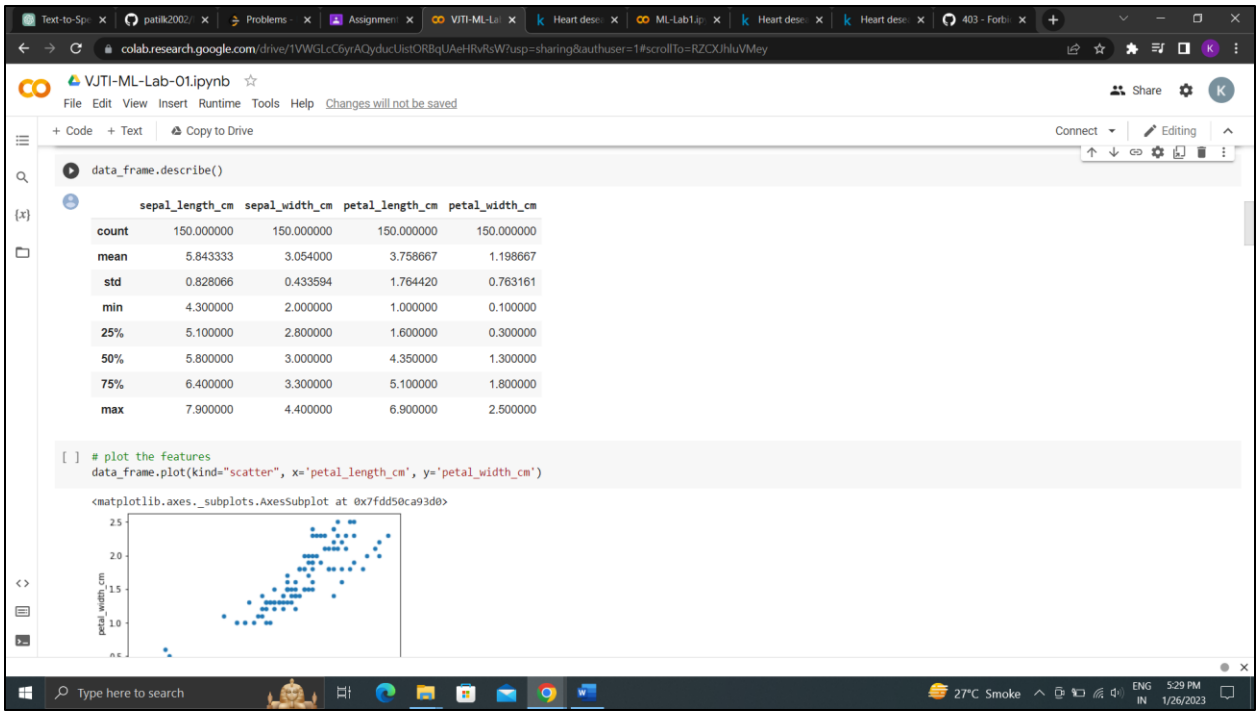



```
import matplotlib.pyplot as plt
plt.scatter()
```

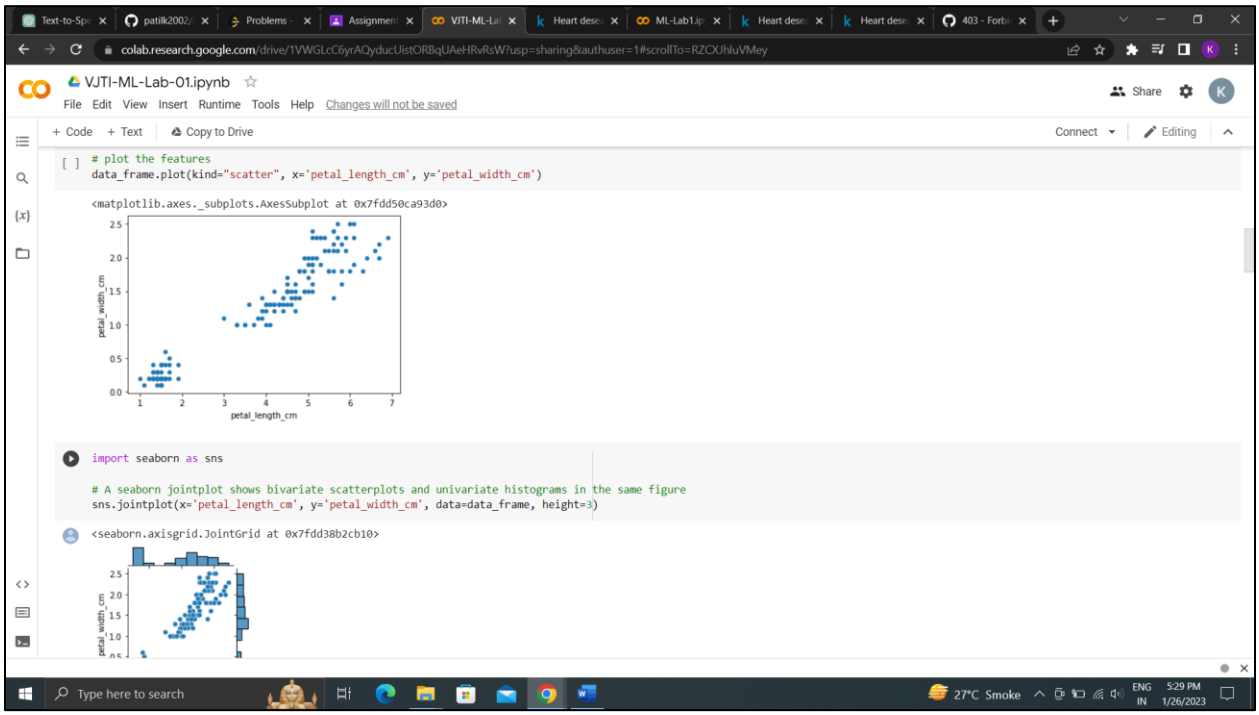
Functions performed using Pandas :



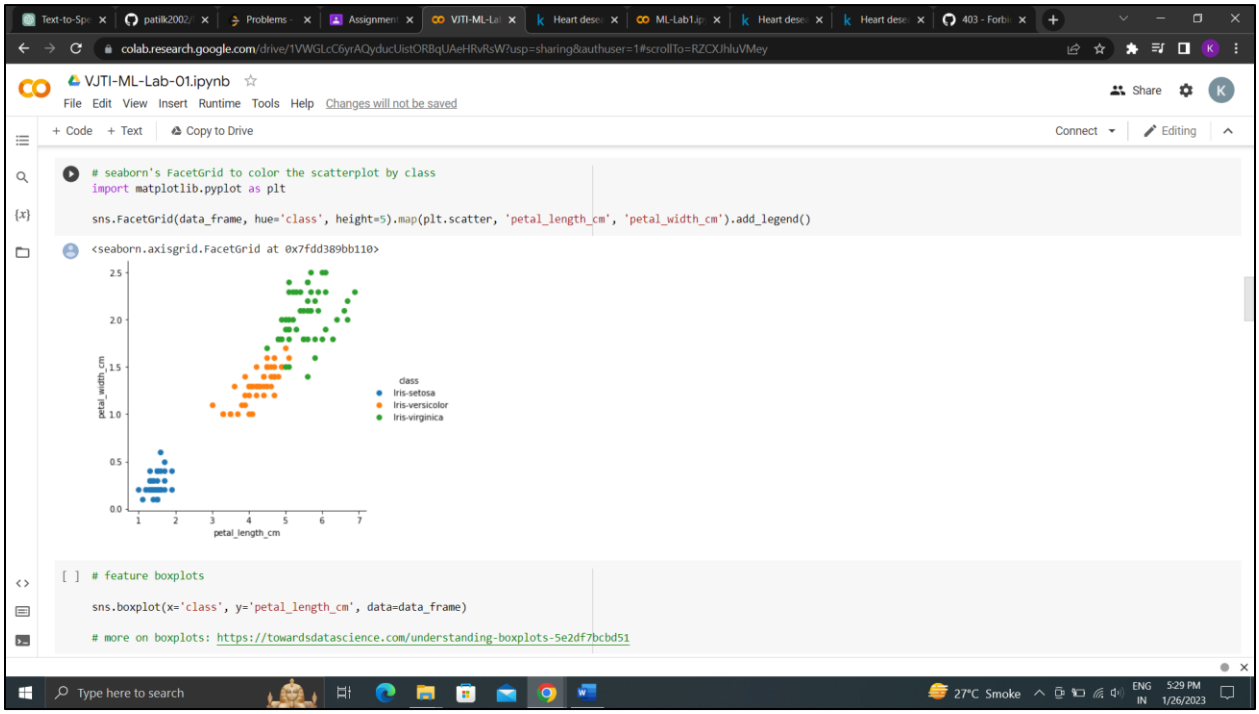


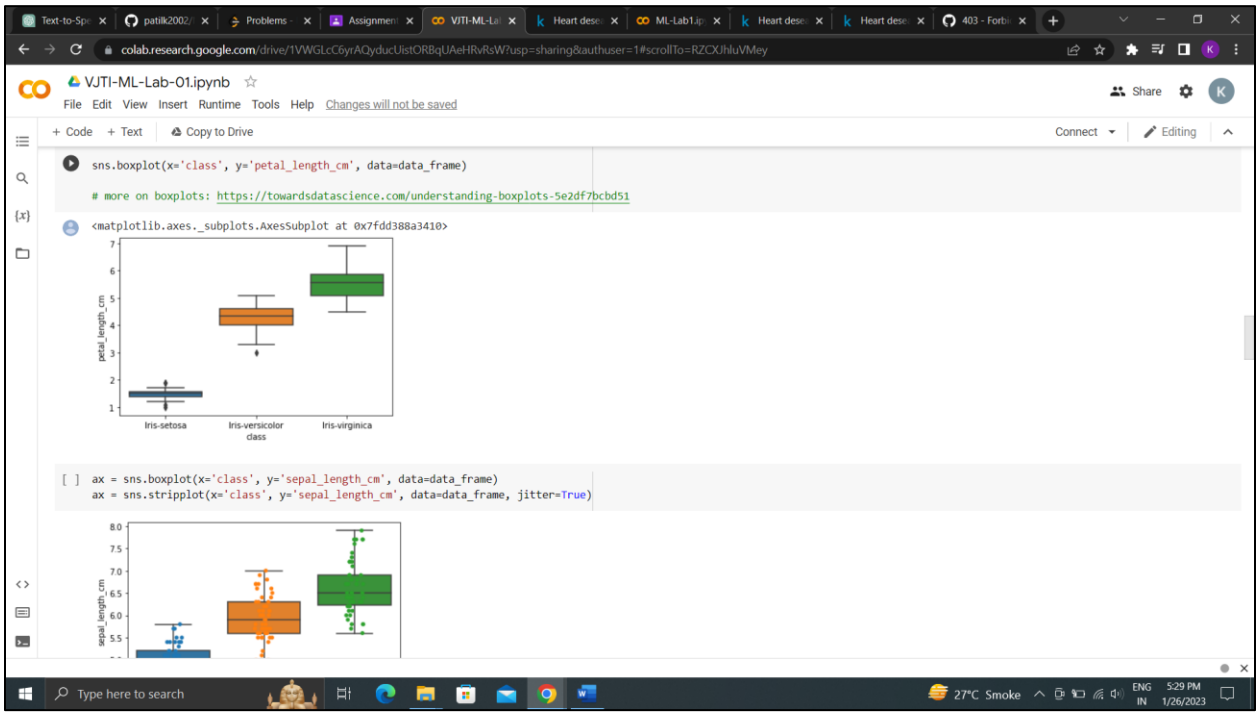


Functions Performed Seaborn :



Functions performed using matplotlib





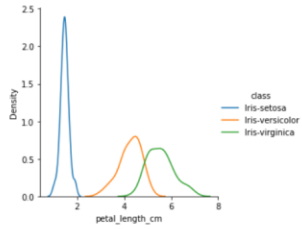
colab.research.google.com/drive/1VWGLcC6yrAQyducUistORBqUAeHRvRsW?usp=sharing&authuser=1#scrollTo=RZCXJhluVMey

VJTI-ML-Lab-01.ipynb

File Edit View Insert Runtime Tools Help Changes will not be saved

```
sns.FacetGrid(data_frame, hue='class', height=4) \
    .map(sns.kdeplot, 'petal_length_cm') \
    .add_legend()
```

```
<seaborn.axisgrid.FacetGrid at 0x7fdd389400d0>
```




```
# violoin plot

sns.violinplot(x='class', y='petal_length_cm', data=data_frame, height=6)

# read more: https://chartio.com/learn/charts/violin-plot-complete-guide/
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdd36660cd0>
```

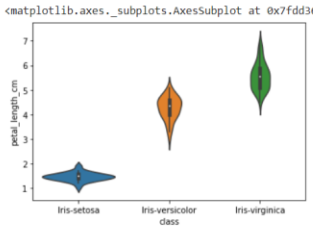
colab.research.google.com/drive/1VWGLcC6yrAQyducUistORBqUAeHRvRsW?usp=sharing&authuser=1#scrollTo=RZCXJhluVMey

VJTI-ML-Lab-01.ipynb

File Edit View Insert Runtime Tools Help Changes will not be saved

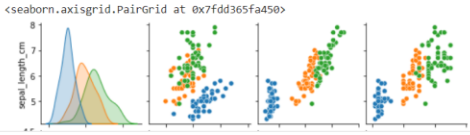
```
sns.violinplot(x='class', y='petal_length_cm', data=data_frame, height=6)

# read more: https://chartio.com/learn/charts/violin-plot-complete-guide/
```

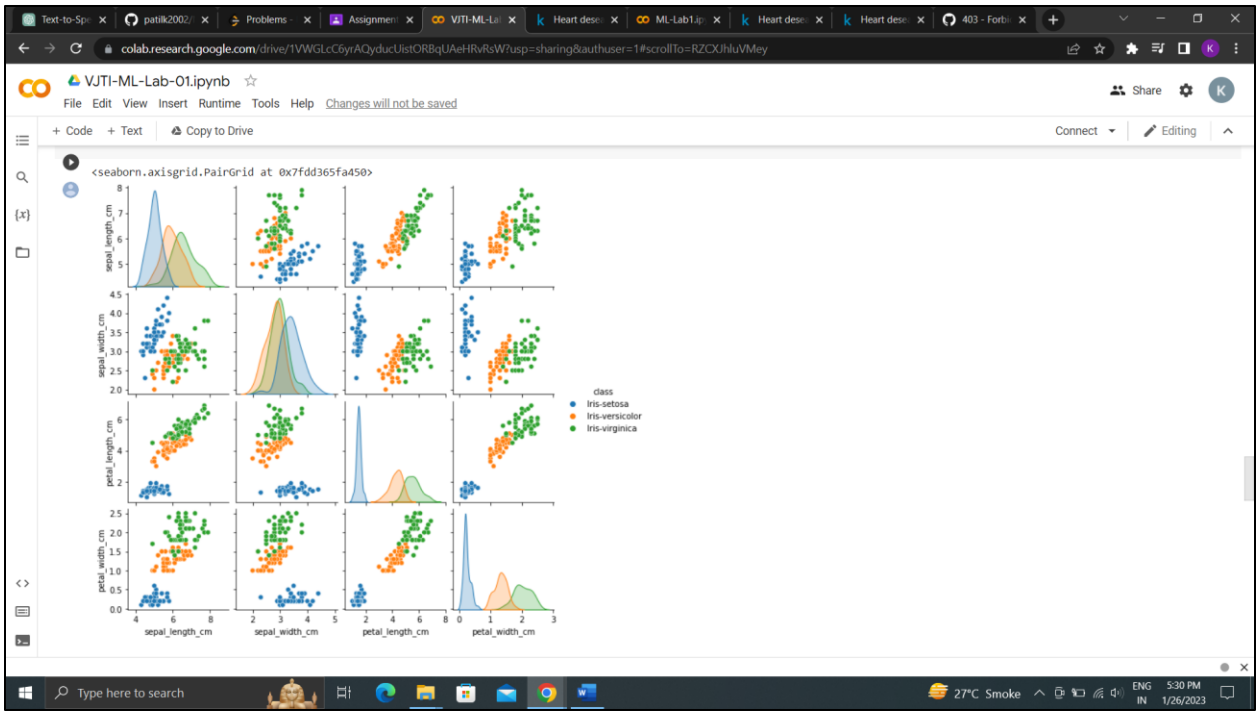
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdd36660cd0>
```

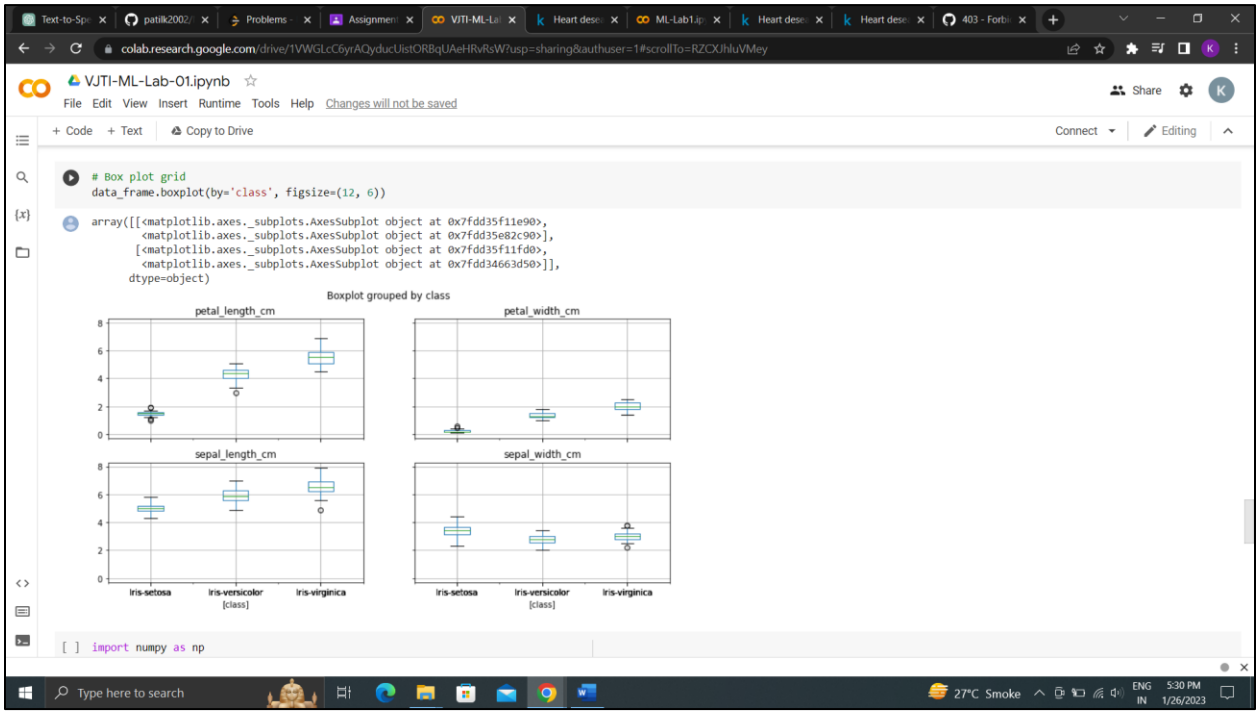



```
# pair-wise feature combination plots
sns.pairplot(data_frame, hue='class', height=2)
```

```
<seaborn.axisgrid.PairGrid at 0x7fdd365fa450>
```




VJTI-ML-Lab-01.ipynb
File Edit View Insert Runtime Tools Help Changes will not be saved
+ Code + Text Copy to Drive
Connect Editing
Share
<seaborn.axisgrid.PairGrid at 0x7fdd365fa450>
sepal_length_cm
sepal_width_cm
petal_length_cm
petal_width_cm
class
Iris-setosa
Iris-versicolor
Iris-virginica


VJTI-ML-Lab-01.ipynb
File Edit View Insert Runtime Tools Help Changes will not be saved
+ Code + Text Copy to Drive
Connect Editing
Share
# Box plot grid
data_frame.boxplot(by='class', figsize=(12, 6))
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7fdd35f11e90>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7fdd35e82c90>],
[<matplotlib.axes._subplots.AxesSubplot object at 0x7fdd35f11fd0>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7fdd34663d50>]],
dtype=object)
Boxplot grouped by class
petal_length_cm
petal_width_cm
sepal_length_cm
sepal_width_cm
Iris-setosa
Iris-versicolor
Iris-virginica
[class]
import numpy as np

colab.research.google.com/drive/1VWGLcC6yrAQyducUistORBqUAeHRvRsW?usp=sharing&authuser=1#scrollTo=RZCXJhluVMey

VJTI-ML-Lab-01.ipynb

File Edit View Insert Runtime Tools Help Changes will not be saved

```python
import numpy as np
import math

# Seperating the data into dependent and independent variables
X = data_frame.iloc[:, :-1].values
y = data_frame.iloc[:, -1].values

# Create train and test split

def train_test_split(X, y, train=0.7):
    # first shuffle
    permute = np.random.permutation(len(X))
    X = X[permute]
    y = y[permute]

    X_train, y_train = X[0:math.floor(0.7*len(X))], y[0:math.floor(0.7*len(X))]
    X_test, y_test = X[math.floor(0.7*len(X)):], y[math.floor(0.7*len(X)):]

    return X_train, y_train, X_test, y_test
```

```python
X_train, y_train, X_test, y_test = train_test_split(X, y, train=0.7)
```

```python
# Logistic Regression

from sklearn.linear_model import LogisticRegression

# fit logistic regression model on training data
classifier = LogisticRegression()
classifier.fit(X_train, y_train)

# predict model on testing data
```

colab.research.google.com/drive/1VWGLcC6yrAQyducUistORBqUAeHRvRsW?usp=sharing&authuser=1#scrollTo=RZCXJhluVMey

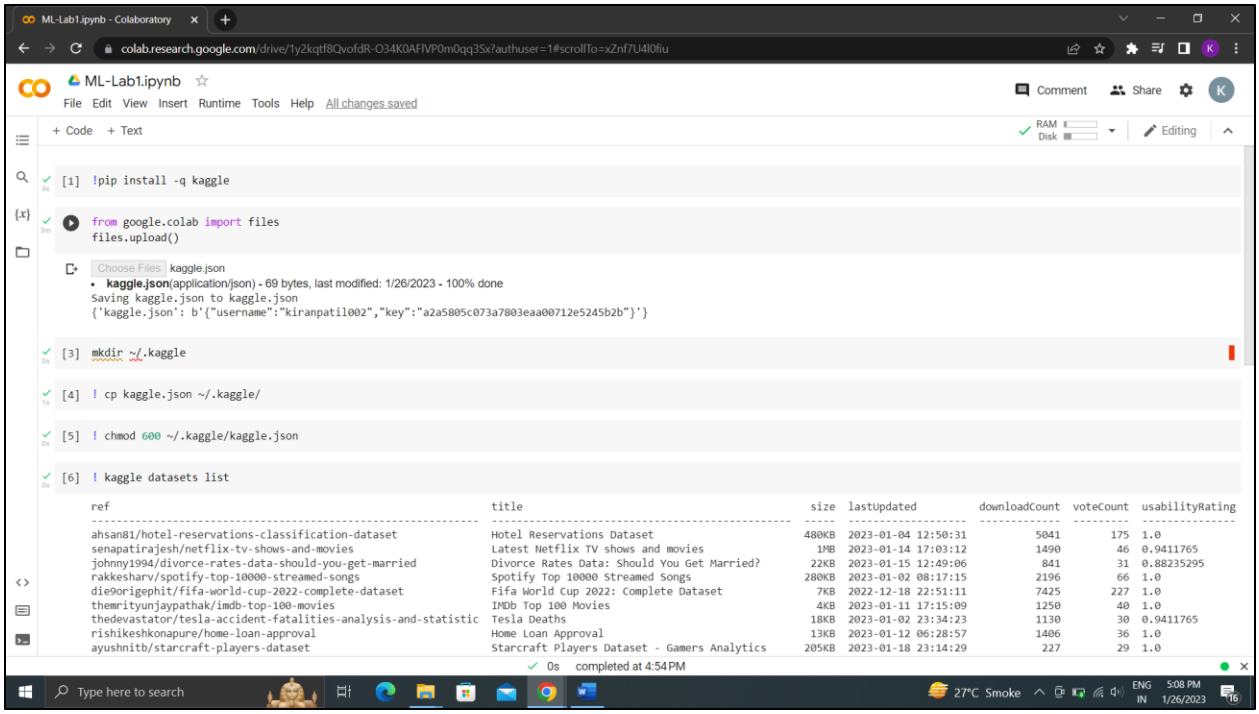
VJTI-ML-Lab-01.ipynb

File Edit View Insert Runtime Tools Help Changes will not be saved

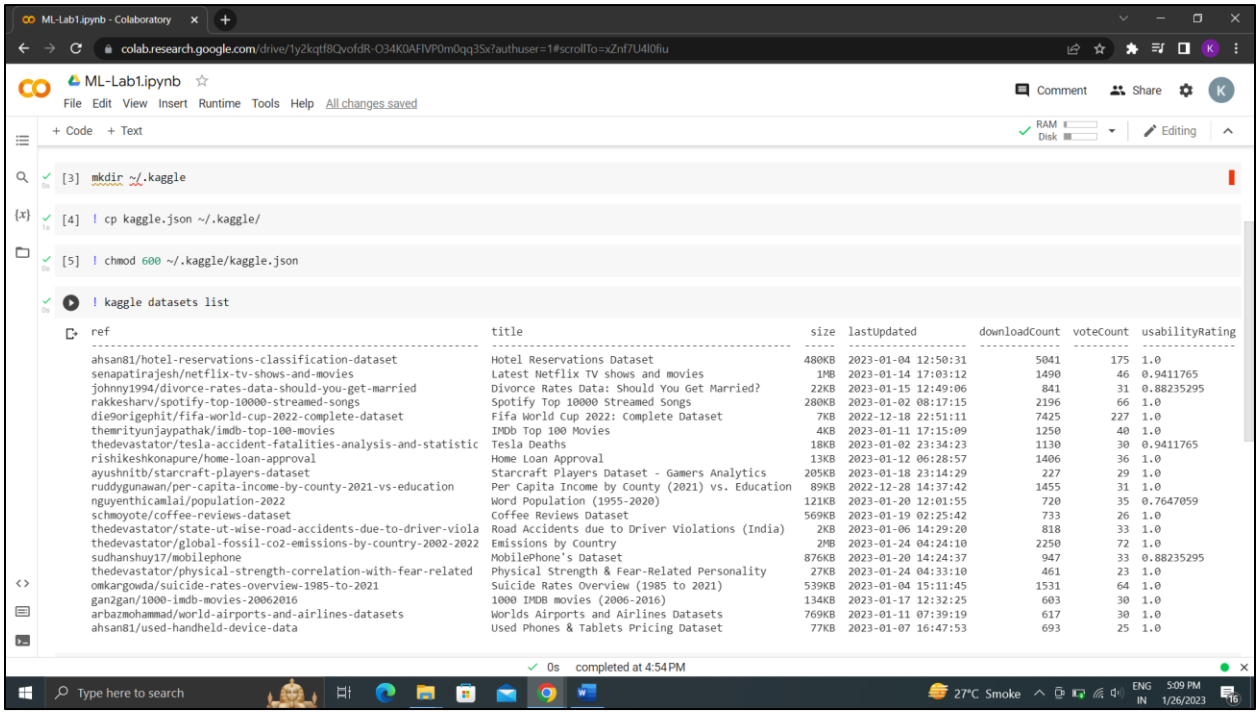
```python
from sklearn.linear_model import LogisticRegression

# fit logistic regression model on training data
classifier = LogisticRegression()
classifier.fit(X_train, y_train)

# predict model on testing data
y_pred = classifier.predict(X_test)
```

```python
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report

# Summary of the predictions made by the classifier
print(classification_report(y_test, y_pred, digits=3))
print(confusion_matrix(y_test, y_pred))
# Accuracy score
from sklearn.metrics import accuracy_score
print('accuracy is',accuracy_score(y_pred,y_test))
```

```
                 precision    recall  f1-score   support

    Iris-setosa      1.000     1.000     1.000        13
Iris-versicolor      0.941     1.000     0.970        16
 Iris-virginica      1.000     0.938     0.968        16

       accuracy                          0.978        45
      macro avg      0.980     0.979     0.979        45
   weighted avg      0.979     0.978     0.978        45

[[13  0  0]
 [ 0 16  0]
 [ 0  1 15]]
accuracy is 0.9777777777777777
```



```
[1] !pip install -q kaggle

from google.colab import files
files.upload()

Choose Files kaggle.json
• kaggle.json(application/json) - 69 bytes, last modified: 1/26/2023 - 100% done
Saving kaggle.json to kaggle.json
{'kaggle.json': b'{"username":"kiranpatil002","key":"a2a5805c073a7803eaa00712e5245b2b"}'}

[3] mkdir ~/.kaggle

[4] ! cp kaggle.json ~/.kaggle/

[5] ! chmod 600 ~/.kaggle/kaggle.json

[6] ! kaggle datasets list

ref                                                         title                                          size  lastUpdated          downloadCount  voteCount  usabilityRating
----------------------------------------------------------  ---------------------------------------------  ----  -------------------  -------------  ---------  ---------------
ahsan81/hotel-reservations-classification-dataset           Hotel Reservations Dataset                     480KB 2023-01-04 12:50:31           5041        175  1.0
senapatirajesh/netflix-tv-shows-and-movies                  Latest Netflix TV shows and movies               1MB 2023-01-14 17:03:12           1490         46  0.9411765
johnny1994/divorce-rates-data-should-you-get-married        Divorce Rates Data: Should You Get Married?     22KB 2023-01-15 12:49:06            841         31  0.88235295
rakkesharv/spotify-top-10000-streamed-songs                 Spotify Top 10000 Streamed Songs               280KB 2023-01-02 08:17:15           2196         66  1.0
die9origephit/fifa-world-cup-2022-complete-dataset          Fifa World Cup 2022: Complete Dataset            7KB 2022-12-18 22:51:11           7425        227  1.0
themrityunjaypathak/imdb-top-100-movies                     IMDb Top 100 Movies                              4KB 2023-01-11 17:15:09           1250         40  1.0
thedevastator/tesla-accident-fatalities-analysis-and-statistic Tesla Deaths                                 18KB 2023-01-02 23:34:23           1130         30  0.9411765
rishikeshkonapure/home-loan-approval                        Home Loan Approval                              13KB 2023-01-12 06:28:57           1406         36  1.0
ayushnitb/starcraft-players-dataset                         Starcraft Players Dataset - Gamers Analytics   205KB 2023-01-18 23:14:29            227         29  1.0
```



```
[3] mkdir ~/.kaggle

[4] ! cp kaggle.json ~/.kaggle/

[5] ! chmod 600 ~/.kaggle/kaggle.json

[6] ! kaggle datasets list

ref                                                         title                                          size  lastUpdated          downloadCount  voteCount  usabilityRating
----------------------------------------------------------  ---------------------------------------------  ----  -------------------  -------------  ---------  ---------------
ahsan81/hotel-reservations-classification-dataset           Hotel Reservations Dataset                     480KB 2023-01-04 12:50:31           5041        175  1.0
senapatirajesh/netflix-tv-shows-and-movies                  Latest Netflix TV shows and movies               1MB 2023-01-14 17:03:12           1490         46  0.9411765
johnny1994/divorce-rates-data-should-you-get-married        Divorce Rates Data: Should You Get Married?     22KB 2023-01-15 12:49:06            841         31  0.88235295
rakkesharv/spotify-top-10000-streamed-songs                 Spotify Top 10000 Streamed Songs               280KB 2023-01-02 08:17:15           2196         66  1.0
die9origephit/fifa-world-cup-2022-complete-dataset          Fifa World Cup 2022: Complete Dataset            7KB 2022-12-18 22:51:11           7425        227  1.0
themrityunjaypathak/imdb-top-100-movies                     IMDb Top 100 Movies                              4KB 2023-01-11 17:15:09           1250         40  1.0
thedevastator/tesla-accident-fatalities-analysis-and-statistic Tesla Deaths                                 18KB 2023-01-02 23:34:23           1130         30  0.9411765
rishikeshkonapure/home-loan-approval                        Home Loan Approval                              13KB 2023-01-12 06:28:57           1406         36  1.0
ayushnitb/starcraft-players-dataset                         Starcraft Players Dataset - Gamers Analytics   205KB 2023-01-18 23:14:29            227         29  1.0
ruddygunawan/per-capita-income-by-county-2021-vs-education  Per Capita Income by County (2021) vs. Education  89KB 2022-12-28 14:37:42           1455         31  1.0
nguyenthicamlai/population-2022                             Word Population (1955-2020)                    121KB 2023-01-20 12:01:55            720         35  0.7647059
schmoyote/coffee-reviews-dataset                            Coffee Reviews Dataset                         569KB 2023-01-19 02:25:42            733         26  1.0
thedevastator/state-ut-wise-road-accidents-due-to-driver-viola Road Accidents due to Driver Violations (India)  2KB 2023-01-06 14:29:20            818         33  1.0
thedevastator/global-fossil-co2-emissions-by-country-2002-2022 Emissions by Country                          2MB 2023-01-24 04:24:10           2250         72  1.0
sudhanshuy17/mobilephone                                    MobilePhone's Dataset                          876KB 2023-01-20 14:24:37            947         33  0.88235295
thedevastator/physical-strength-correlation-with-fear-related Physical Strength & Fear-Related Personality  27KB 2023-01-24 04:33:10            461         23  1.0
omkargowda/suicide-rates-overview-1985-to-2021              Suicide Rates Overview (1985 to 2021)          539KB 2023-01-04 15:11:45           1531         64  1.0
gan2gan/1000-imdb-movies-20062016                           1000 IMDB movies (2006-2016)                   134KB 2023-01-17 12:32:25            603         30  1.0
arbazmohammad/world-airports-and-airlines-datasets          Worlds Airports and Airlines Datasets          769KB 2023-01-11 07:39:19            617         30  1.0
ahsan81/used-handheld-device-data                           Used Phones & Tablets Pricing Dataset           77KB 2023-01-07 16:47:53            693         25  1.0
```

Conclusion : From this experiment we learned to Download a dataset from Kaggle and Performed data exploration using Pandas, Seaborn, Matplotlib python libraries