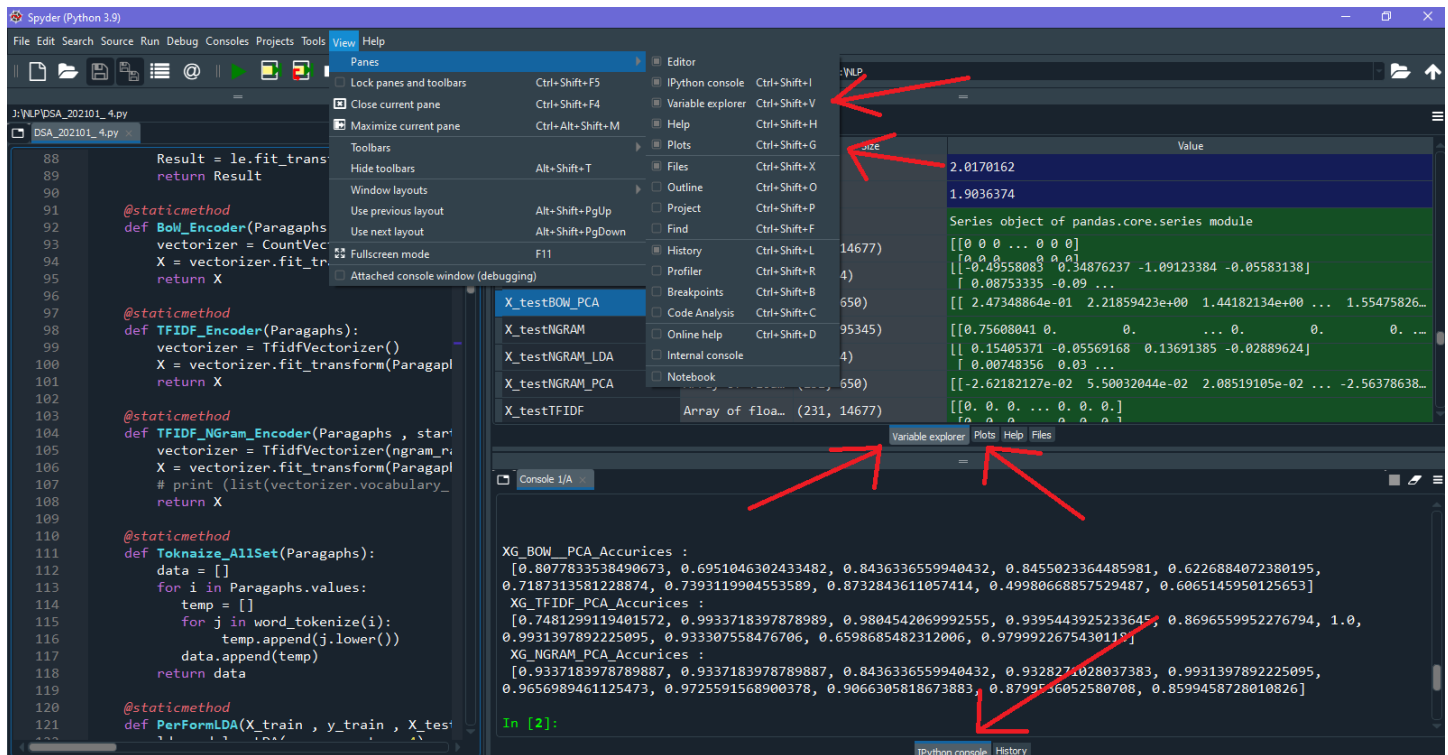


For best experience please consider to run **DSA_202101_4.py** on Spyder to see all created variables on variables explorer tab and see all plots on plots tab.



Please make sure that you have installed these libraries before importing it ->

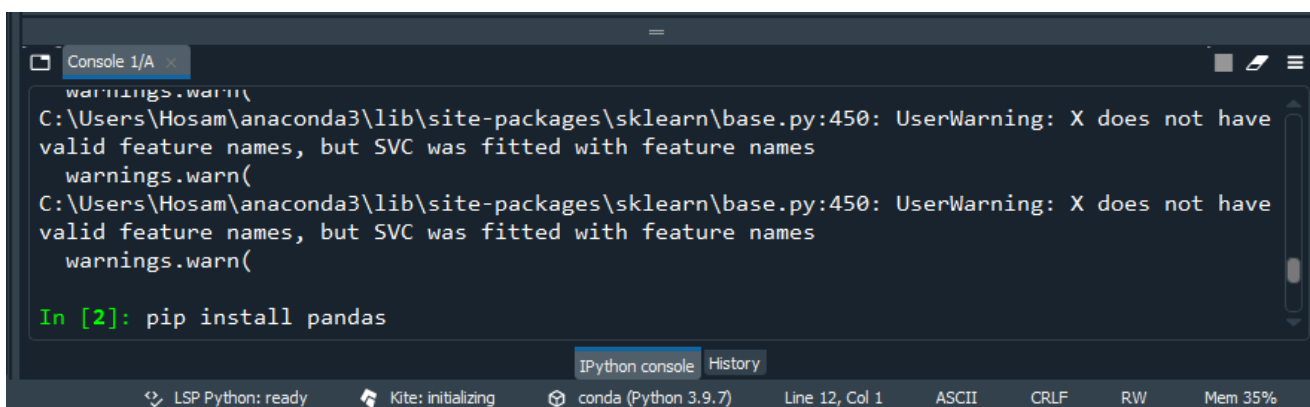
```
import nltk,re,pandas as pd,random,numpy as np  
from nltk.corpus import stopwords  
from sklearn import preprocessing  
from sklearn.feature_extraction.text import CountVectorizer  
from nltk.tokenize import word_tokenize  
from sklearn.metrics import accuracy_score  
from sklearn.metrics import classification_report  
from nltk import ngrams  
import matplotlib.pyplot as plt  
from sklearn.feature_extraction.text import TfidfVectorizer
```

```
from sklearn.model_selection import train_test_split  
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis as LDA  
import gensim  
from sklearn.svm import SVC  
from sklearn.tree import DecisionTreeClassifier  
from sklearn.metrics import confusion_matrix  
import seaborn as sns  
from sklearn.model_selection import StratifiedKFold  
from sklearn.neighbors import KNeighborsClassifier  
from sklearn.ensemble import ExtraTreesClassifier  
import xgboost as xgb  
from sklearn.decomposition import PCA  
from sklearn.ensemble import RandomForestClassifier
```

To install the missing libraries, you have to write something like this ->

```
>>> pip install pandas
```

In Spyder console.



```
warnings.warn(  
C:\Users\Hosam\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have  
valid feature names, but SVC was fitted with feature names  
warnings.warn(  
C:\Users\Hosam\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have  
valid feature names, but SVC was fitted with feature names  
warnings.warn(  
  
In [2]: pip install pandas
```

The screenshot shows the Spyder Python console interface. At the top, there's a tab labeled 'Console 1/A'. The main area displays two identical warnings from sklearn's base.py, indicating that a model (SVC) was fitted with feature names but the data (X) does not have valid feature names. Below the warnings, the command 'In [2]: pip install pandas' is entered. At the bottom, there's a status bar with several indicators: 'LSP Python: ready', 'Kite: initializing', 'conda (Python 3.9.7)', 'Line 12, Col 1', 'ASCII', 'CRLF', 'RW', and 'Mem 35%'.

or

```
>>> conda install pandas
```

In Anaconda prompt.

```
Select Anaconda Prompt (anaconda3)

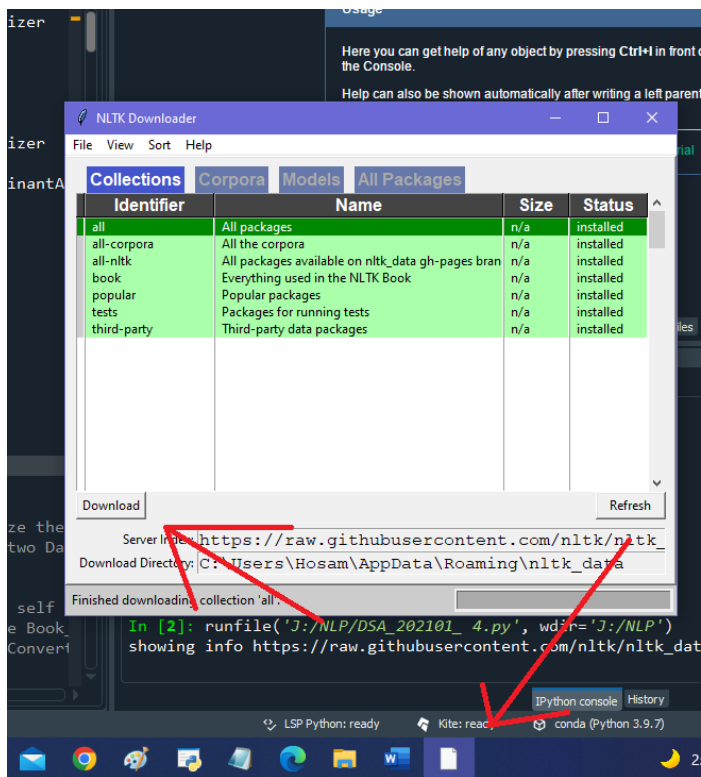
(base) C:\Users\Hosam>conda install pandas
```

List of required libraries:

```
pip install regex
```

```
>>> import nltk
```

```
>>> nltk.download()
```



```
pip install pandas
```

```
pip install numpy
```

```
pip install seaborn
```

```
pip install -U scikit-learn
```

```
pip install -U matplotlib
```

```
pip install -U genism
```

```
pip install xgboost
```

Or

```
Conda install xgboost
```

And so on...