LOADING DATA INTO JUPITER NOTEBOOK

https://www.anaconda.com/download

Jupyter Notebook is an open-source web application that allows you to create and share documents containing live code, equations, visualizations, and narrative text.

1. Importing .CSV Files / Flat files

a. Download the relevant .CSV file from a trusted source.
b. Upload the file using Jupiter file upload option available in the home page.
c. Create a new notebook.
d. Import the Pandas library into your Jupyter Notebook. You can do this by running the following command: **import pandas as pd**. This command imports the Pandas library and assigns it the alias "pd", which is a common convention in the Python community.
e. To load a CSV file into Pandas, you can use the read_csv() function. This function takes the path to the CSV file as a parameter and returns a DataFrame object, which is a two-dimensional table-like data structure that can hold data of different types. Assuming that your CSV file is stored in the same directory as your Jupyter Notebook, you can load it by running the following command: **df = pd.read_csv('data.csv')**. This command reads the CSV file named "mydata.csv" and stores its contents in a DataFrame object named "df". You can replace "data.csv" with the name of your CSV file.
f. You can now explore the data as it the file is now loaded successfully into the environment.

2. Loading the dataset directly from Kaggle

a. Navigate to the dataset required in kaggle.
```
pip install opendatasets
import opendatasets as od
dataset = '<i>Dataset's URL</i>'
od.download(dataset)
```
b. Asks for username and key (make sure to download the API key from the kaggle account)
```
{"username":"","key":""}
```
c. The other way of doing this is uploading the jason file directly int Jupiter notebook to avoid providing username and key.
```
import os
```
while downloading get the path of the file and move it into directory variable.

d. **data_dir = '.\Path'**
os.listdir(data_dir)
import pandas as pd

This information will provide you with basic information on working with datasets and files in Python environment.