

Instructions :

1. **Install Python**
2. **Install an IDE** (VS Code/Jupyter/PyCharm/Anaconda)- in this tutorial we use PyCharm
3. **Visit Pandas Tutorial** (<https://www.w3schools.com/python/pandas/default.asp>)

Question 1: Create an index series as follows /2 Marks

a	10
b	15
c	18
d	17
e	20

Reference:

https://www.w3schools.com/python/pandas/trypython.asp?filename=demo_pandas_series_label**Question 2: Create data frame as follows: / 3 Marks**

	Gender	Age
0	M	45
1	M	12
2	F	30
3	F	28
4	F	19

Reference:

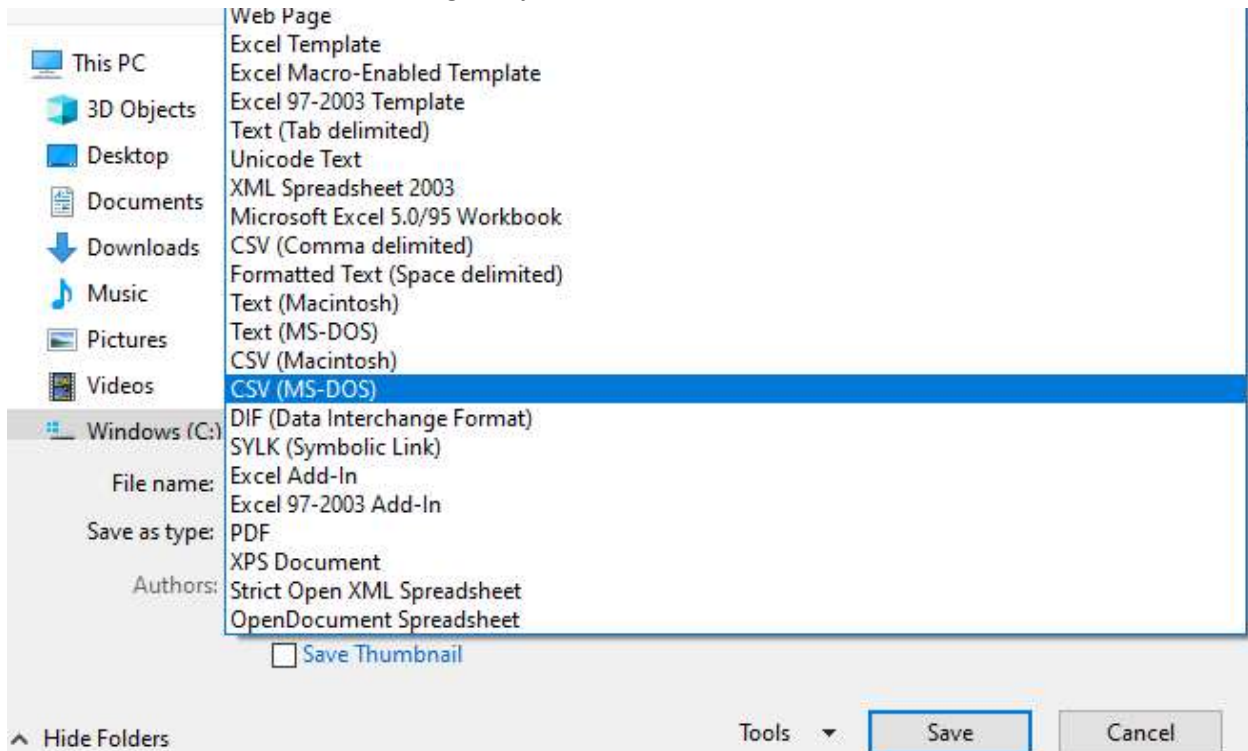
https://www.w3schools.com/python/pandas/trypython.asp?filename=demo_pandas_dataframes1**Question 3:**3.1. Create the following CSV (Comma Separated File) and import it in your project / **7 Marks**

Index	Names	Gender	Age	Height	Weight	Blood Pressure?
0	Peter	Male	45	180	98	Yes
1	John	Male	38	176	71	No
2	Ann	Female	75	169	63	No
3	Leslie	Female	25	178	81	Yes
4	Paul	Male	56	190	88	Yes
5	Alain	Male	35	185	79	No
6	Irene	Female	56	171	110	No
7	Jetheme	Female	27	159	98	Yes

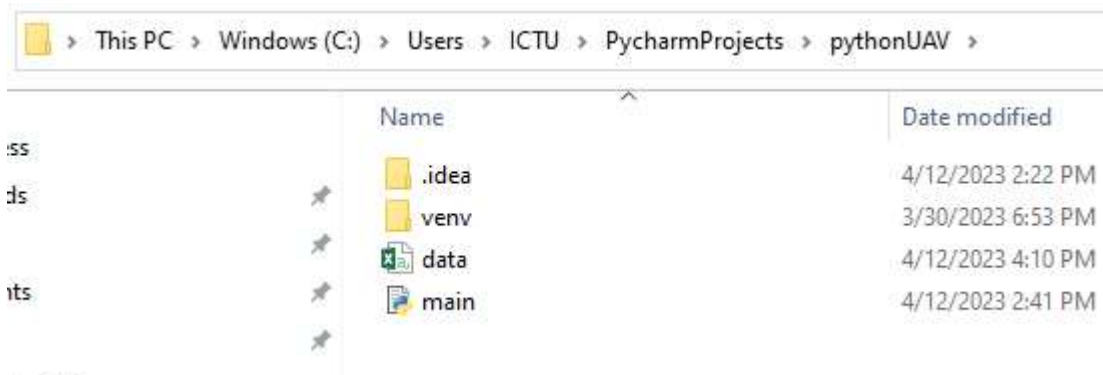
Reference:

- Create the file in Excel
- Save it in any location with the extension .csv (you need to choose the extension while saving. See below for CSV in a Windows Machine. If you have a Mac choose CSV(Macintosh))

Walkthrough Project 1- Introduction to Data Science



- Copy your CSV file in your project's folder. See example below:



- Make sure it appears as an object in your project. If it does, you have imported it and you are ready to manipulate it.

3.2. Read and print the CSV file from your project /2 Marks

Reference:

https://www.w3schools.com/python/pandas/trypandas.asp?filename=demo_pandas_csv

Question 4:

4.1. Download the CSV file from the link below: /1 Mark

https://www.w3schools.com/python/pandas/pandas_csv.asp

Read CSV Files

A simple way to store big data sets is f

CSV files contains plain text and is a w
including Pandas.

In our examples we will be using a CS'

[Download data.csv](#) or [Open data.csv](#)

4.2. Import it in your project and perform data cleaning/cleansing using the references below: /8 Marks

Cleaning empty cells

https://www.w3schools.com/python/pandas/pandas_cleaning_empty_cells.asp

Cleaning wrong format

https://www.w3schools.com/python/pandas/pandas_cleaning_wrong_format.asp

Cleaning wrong data

https://www.w3schools.com/python/pandas/pandas_cleaning_wrong_data.asp

Removing duplicates

https://www.w3schools.com/python/pandas/pandas_cleaning_duplicates.asp

Question 5:

5.1. Create a CSV file of the dataset below: /2 Marks

Index	Age	Height	Weight
0	45	180	98
1	38	176	71
2	75	169	63
3	25	178	81
4	56	190	88
5	35	185	79

Walkthrough Project 1- Introduction to Data Science

6	56	171	110
7	27	159	98
8	24	163	85
9	62	181	115
10	23	167	97
11	42	182	96
12	52	164	69
13	23	173	72
14	45	176	86
15	52	181	120
16	34	186	69
17	48	164	83
18	42	183	68
19	24	158	86

5.2. Plot the dataset using the reference link below: / **5 Marks**

Reference:

https://www.w3schools.com/python/pandas/pandas_plotting.asp