

Input, Subset and Output External Data Files using Pandas All Versions

This section shows basic code for reading, sub-setting and writing external data files using pandas.

Examples

Basic Code to Import, Subset and Write External Data Files Using Pandas

```
# Print the working directory
import os
print os.getcwd()
# C:\Python27\Scripts

# Set the working directory
os.chdir('C:/Users/general1/Documents/simple Python files')
print os.getcwd()
# C:\Users\general1\Documents\simple Python files

# load pandas
import pandas as pd

# read a csv data file named 'small_dataset.csv' containing 4 lines and 3 variable
my_data = pd.read_csv("small_dataset.csv")
my_data
#      x  y  z
# 0    1  2  3
# 1    4  5  6
# 2    7  8  9
# 3   10 11 12

my_data.shape      # number of rows and columns in data set
# (4, 3)

my_data.shape[0]    # number of rows in data set
# 4

my_data.shape[1]    # number of columns in data set
# 3

# Python uses 0-based indexing. The first row or column in a data set is located
# at position 0. In R the first row or column in a data set is located
# at position 1.

# Select the first two rows
```

Syntax

Parameters

Remarks