

## Data Loading, Storage

```
In [ ]: ▶ import pandas as pd
```

```
In [ ]: ▶ mydata = {  
    'name' : ['Ali', 'Sara', 'Taha', 'Omid'],  
    'age' : [27, 24, 25, 26],  
    'Score': [19, 18, 20, 13]  
}
```

```
In [ ]: ▶ mydf = pd.DataFrame(mydata, columns = ['name', 'age', 'Score'])  
mydf
```

## CSV : Comma-separated values

```
In [ ]: ▶ mydf.to_csv('Files/score.csv', index=False)
```

```
In [ ]: ▶ pd.read_csv('Files/score.csv')
```

## File without header

```
In [ ]: ▶ pd.read_csv('Files/student.csv', header=None)
```

```
In [ ]: ▶ pd.read_csv('Files/student.csv', names=['name', 'Python', 'C++', 'Java'])
```

## skiprows

```
In [ ]: ▶ pd.read_csv('Files/age.csv')
```

```
In [ ]: ▶ pd.read_csv('Files/age.csv', skiprows=[0])
```

```
In [ ]: ▶ pd.read_csv('Files/age.csv', header=[1])
```

## csv.reader

```
In [ ]: ▶ import csv
```

```
In [ ]: ▶ with open('Files/score.csv') as f:  
        x = list(csv.reader(f))  
        x
```

```
In [ ]: ▶ x[0]
```

```
In [ ]: ▶ x[1:]
```

## read\_table

```
In [ ]: ▶ pd.read_table('Files/score.csv')
```

```
In [ ]: ▶ pd.read_table('Files/score.csv', sep=',')
```

```
In [ ]: ▶ pd.read_table('Files/mytext.txt', sep='\s+')
```

## nrows , chunksize

```
In [ ]: ▶ path = 'Files/Alphabet.csv'
```

```
In [ ]: ▶ pd.read_csv(path)
```

```
In [ ]: ▶ pd.read_csv(path, nrows=8)
```

```
In [ ]: ▶ pd.options.display.max_rows = 8  
pd.read_csv(path)
```

```
In [ ]: ▶ pd.read_csv(path, skipfooter=24, engine='python')
```

```
In [ ]: ▶ c = pd.read_csv(path, chunksize=8)  
  
for i in c:  
    print(i)
```

## sys.stdout

```
In [ ]: ▶ import sys
```

```
In [ ]: ▶ mydata = pd.read_csv('Files/score.csv')
```

```
In [ ]: ▶ mydata.to_csv(sys.stdout)
```

```
In [ ]: ▶ mydata.to_csv(sys.stdout, index=False, header=False, sep='|')
```

## df.to\_excel

```
In [ ]: ▶ # pip install openpyxl
```

```
In [ ]: ▶ mydf
```

```
In [ ]: ▶ mydf.to_excel('Files/score.xlsx', index=False)
```

```
In [ ]: ▶ pd.read_excel('Files/score.xlsx')
```

## JSON : JavaScript Object Notation

```
In [ ]: ▶ import json
```

```
In [ ]: ▶ mystr = """
{
  "FirstName": "Taha",
  "Courses" : [{"Name": "Python", "Score": 18}, {"Name": "C++", "Score": 17}
}
"""
```

```
In [ ]: ▶ mydict = json.loads(mystr)
frame = pd.DataFrame(mydict['Courses'], columns=['Name', 'Score'])
frame
```

```
In [ ]: ▶ frame.to_json()
```

```
In [ ]: ▶ frame.to_json(orient='values')
```

```
In [ ]: ▶ frame.to_json(orient='index')
```

```
In [ ]: ▶ frame.to_json(orient='split')
```

```
In [ ]: ▶ pd.read_json(_, orient='split')
```

## HTML : Hypertext Markup Language

```
In [ ]: ▶ #pip install lxml
```

```
In [ ]: ▶ #pip install html5lib
```

```
In [ ]: ▶ mystr = """  
<table>  
  <thead>  
    <tr>  
      <th>name</th>  
      <th>score</th>  
    </tr>  
  </thead>  
  <tbody>  
    <tr>  
      <td>Ali</td>  
      <td>12</td>  
    </tr>  
    <tr>  
      <td>Sara</td>  
      <td>18</td>  
    </tr>  
  </tbody>  
</table>  
"""
```

```
In [ ]: ▶ lst = pd.read_html(mystr)
```

```
In [ ]: ▶ lst[0]
```

```
In [ ]: ▶ r = pd.read_html('Files/test.html')  
r[0]
```

```
In [ ]: ▶ r[0].info()
```

```
In [ ]: ▶ u = 'https://en.wikipedia.org/wiki/Minnesota'  
x = pd.read_html(u , match='Election results from statewide races')  
x[0].head()
```

## PKL file

```
In [ ]: ▶ frame = pd.read_csv('Files/score.csv')  
frame
```

```
In [ ]: ▶ frame.to_pickle('Files/test.pkl')
```

```
In [ ]: ▶ pd.read_pickle('Files/test.pkl')
```

## Hierarchical Data Format (HDF)

```
In [ ]: ▶ #pip install tables
```

```
In [ ]: ▶ import tables
```

```
In [ ]: ▶ mydf = pd.DataFrame({'A': [18, 22], 'B': [3, 15]})  
mydf
```

```
In [ ]: ▶ path = 'Files\data.h5'
```

```
In [ ]: ▶ mydf.to_hdf(path, key='df', mode='w')
```

```
In [ ]: ▶ pd.read_hdf(path, 'df')
```

## requests.get

```
In [ ]: ▶ import requests
```

```
In [ ]: ▶ r = requests.get('https://api.github.com/repos/pandas-dev/pandas/issues')
```

```
In [ ]: ▶ if r.status_code == 200:  
    print('Success!')  
elif r.status_code == 404:  
    print('Not Found.')
```

```
In [ ]: ▶ df = pd.DataFrame(r.json(), columns=['number', 'title'])  
df.head()
```

```
to_csv  
read_csv  
  
read_table  
  
to_excel  
read_excel  
  
to_json  
read_json  
  
to_pickle  
read_pickle
```

```
to_hdf  
read_hdf  
  
read_html
```

دانشگاه شهید مدنی آذربایجان  
برنامه نویسی پیشرفته با پایتون  
امین گلزاری اسکویی  
۱۴۰۰-۱۴۰۱

[Codes and Projects \(click here\) \(https://github.com/Amin-Golzari-Oskouei/Python-Programming-Course-Advanced-2021\)](https://github.com/Amin-Golzari-Oskouei/Python-Programming-Course-Advanced-2021) [slides and videos \(click here\)](#)  
(<https://drive.google.com/drive/folders/1Dx3v7fD1QBWL-MNP2hd7ilxaRbeALkKA>)