



Unit 13

csv module

Spread sheet and corresponding CSV file

Name	Exam1	Exam2	Final Exam	Overall Grade
Bill	75.00	100.00	50.00	75.00
Fred	50.00	50.00	50.00	50.00
Irving	0.00	0.00	0.00	0.00
Monty	100.00	100.00	100.00	100.00
Average				56.25

FIGURE 14.2 A simple spreadsheet from Microsoft Excel 2008.

```
Name,Exam1,Exam2,Final Exam,Overall Grade
Bill,75.00,100.00,50.00,75.00
Fred,50.00,50.00,50.00,50.00
Irving,0.00,0.00,0.00,0.00
Monty,100.00,100.00,100.00,100.00

Average,,,,56.25
```

csv file

- <https://docs.python.org/3/library/csv.html>
- Là file text mà mỗi dòng các phần tử phân tách nhau bởi dấu phẩy (,), chấm phẩy (;), hay tab (\t).
- Python hỗ trợ module csv, gồm:
 - csv.field_size_limit – return maximum field size
 - csv.get_dialect – get the dialect which is associated with the name
 - csv.list_dialects – show all registered dialects
 - csv.register_dialect - associate dialect with name
 - csv.unregister_dialect - delete the dialect associated with the name the dialect registry

csv module

- `csv.reader` – read data from a csv file
- `csv.writer` – write data to a csv file
- **`csv.QUOTE_ALL`** - Quote everything, regardless of type.
- **`csv.QUOTE_MINIMAL`** - Quote fields with special characters
- **`csv.QUOTE_NONNUMERIC`** - Quote all fields that aren't numbers value
- **`csv.QUOTE_NONE`** – Don't quote anything in output

Read csv

```
import csv
workbook_file = open('Workbook1.csv', 'r')
workbook_reader = csv.reader(workbook_file)

for row in workbook_reader:
    print(row)

workbook_file.close()
```

```
>>>
['Name', 'Exam1', 'Exam2', 'Final Exam', 'Overall Grade']
['Bill', '75.00', '100.00', '50.00', '75.00']
['Fred', '50.00', '50.00', '50.00', '50.00']
['Irving', '0.00', '0.00', '0.00', '0.00']
['Monty', '100.00', '100.00', '100.00', '100.00']
[]
['Average', '', '', '', '56.25']
```

Read file csv into Dictionary

```
import csv
with ('employee_birthday.txt', mode='r') as csv_file:
    csv_reader = csv.DictReader(csv_file)
    for row in csv_reader:
        print(f'\t{row["name"]} - {row["age"]}.')
```

Write csv

```
import csv
with ('employee_file.csv', mode='w') as employee_file:
    employee_writer = csv.writer(employee_file, delimiter=',',
                                  quotechar='"', quoting=csv.QUOTE_MINIMAL)

    employee_writer.writerow(['John Smith', 'Accounting', 'November'])
    employee_writer.writerow(['Erica Meyers', 'IT', 'March'])
```

Write csv from Dict

```
import csv
with ('employee_file2.csv', mode='w') as csv_file:
    fieldnames = ['emp_name', 'dept']
    writer = csv.DictWriter(csv_file, fieldnames=fieldnames)
    writer.writeheader()
    writer.writerow({'emp_name': 'John Smith', 'dept': 'Accounting'})
    writer.writerow({'emp_name': 'Erica Meyers', 'dept': 'IT'})
```


Xử lý csv trong thư viện Pandas

```
import pandas  
df = pandas.read_csv('hrdata.csv', index_col='Name')  
print(df)
```

```
import pandas  
df = pandas.read_csv('hrdata.csv', index_col='Name',  
parse_dates=['Birth Date'])  
print(df)
```

Xử lý csv trong thư viện Pandas

```
df = pandas.read_csv('hrdata.csv',  
    index_col='Employee',  
    parse_dates=['Hired'],  
    header=0,  
    names=['Employee', 'Hired', 'Salary', 'Sick Days'])  
  
# Process data  
  
# Write new file  
df.to_csv('hrdata_modified.csv')
```