



# Weka – Using CSV Files

---

Kyuseok Shim

Seoul National University

<http://kdd.snu.ac.kr/~shim>



# CSV Files

---

- CSV files could be easily exported from Microsoft Excel or Google Spreadsheet

	A	B	C	D	E
1	outlook	temperature	humidity	windy	play
2	sunny	85	85	FALSE	no
3	sunny	80	90	TRUE	no
4	overcast	83	86	FALSE	yes
5	rainy	70	96	FALSE	yes
6	rainy	68	80	FALSE	yes
7	rainy	65	70	TRUE	no

```
outlook,temperature,humidity,windy,play
sunny,85,85,FALSE,no
sunny,80,90,TRUE,no
overcast,83,86,FALSE,yes
rainy,70,96,FALSE,yes
rainy,68,80,FALSE,yes
rainy,65,70,TRUE,no
```

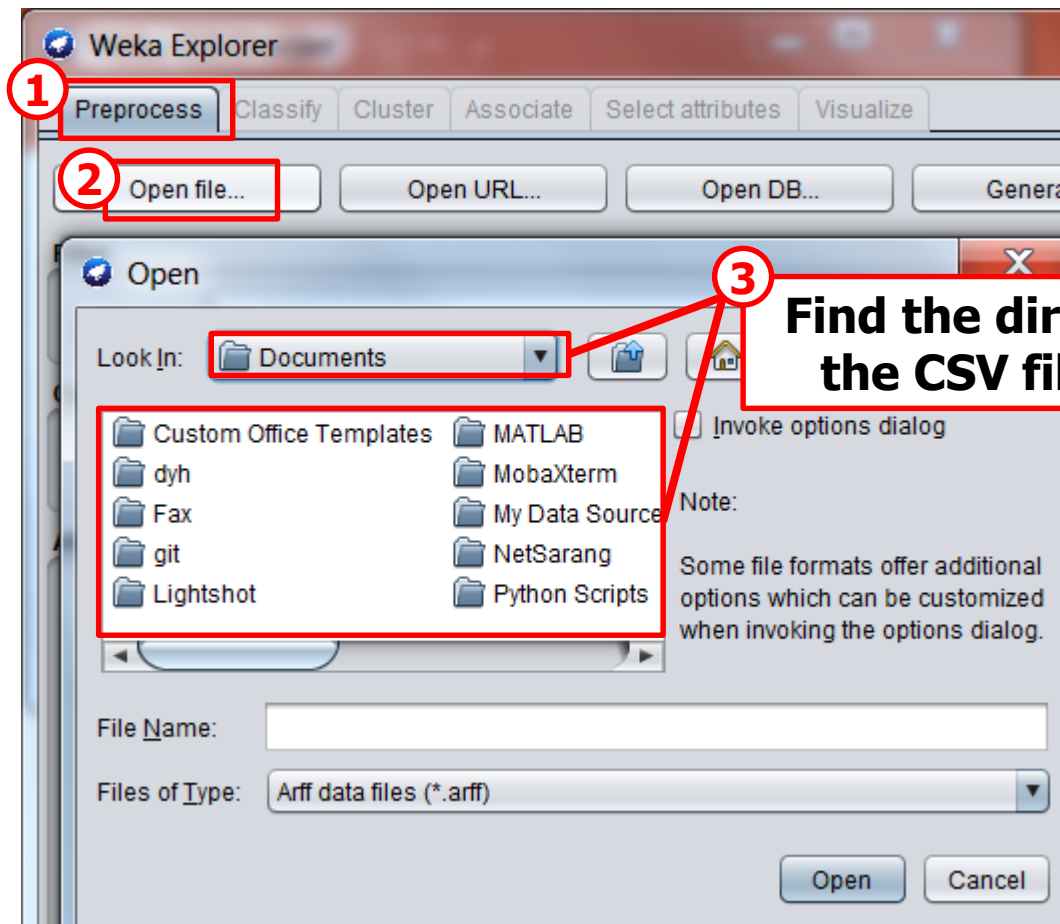


# Download a Sample CSV File

---

- <http://kdd.snu.ac.kr/weka/>
  - Download the **test.csv**

# Using CSV Files for Weka



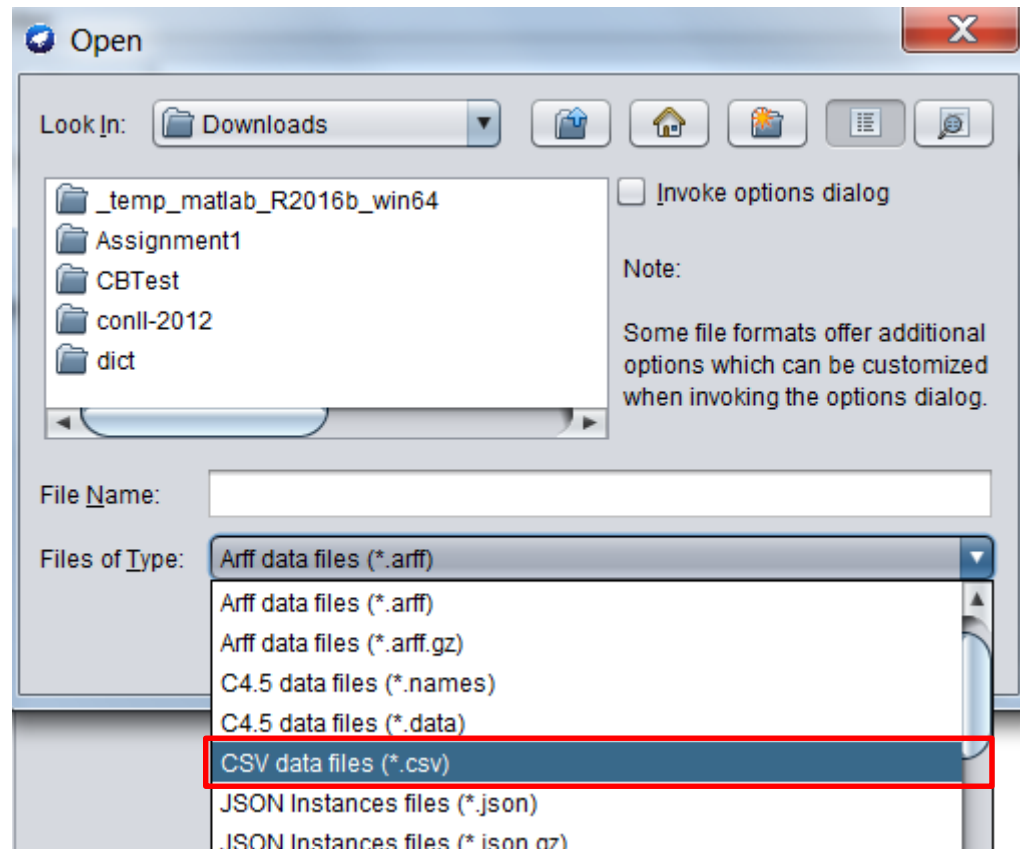
# Using CSV Files for Weka

- If the test.csv is downloaded in C:\Users\dyhong\Downloads



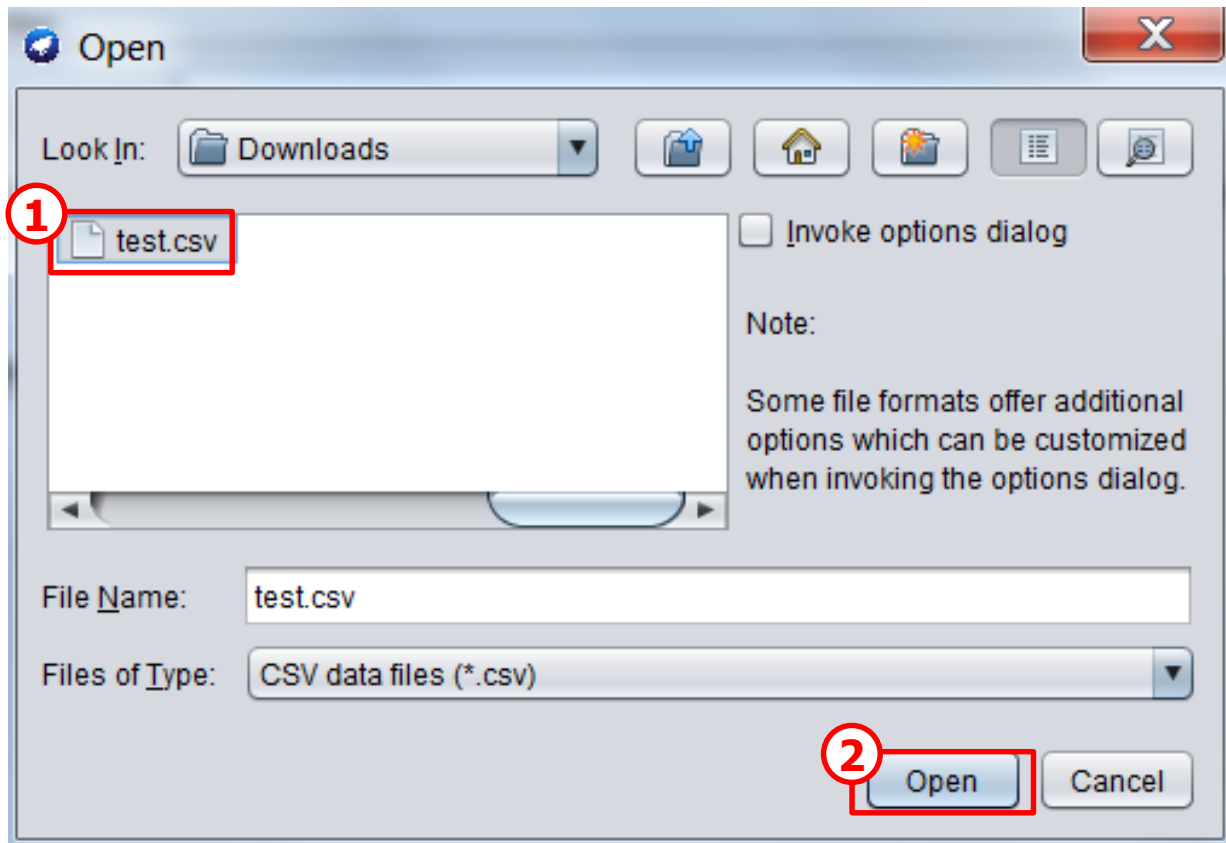
# Using CSV Files for Weka

- If the test.csv is downloaded in C:\Users\dyhong\Downloads



# Using CSV Files for Weka

- If the test.csv is downloaded in C:\Users\dyhong\Downloads



Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file...

Open UR...

Open DB...

Generate...

Undo

Edit...

Save...

Filter

Choose

No

The data is loaded successfully

Apply

Stop

Current relation

Relation: test  
Instances: 14

Attributes: 5  
Sum of weights: 14

Selected attribute

Name: outlook  
Missing: 0 (0%)  
Distinct: 3  
Type: Nominal  
Unique: 0 (0%)

Attributes

All

None

Invert

Pattern

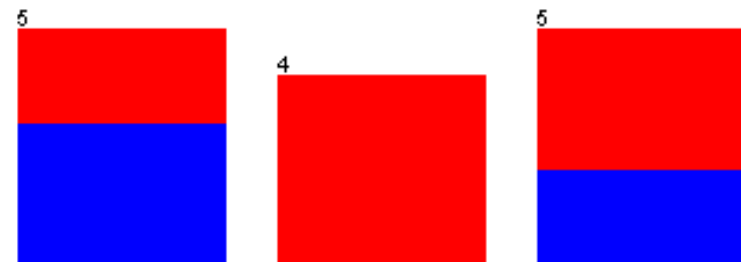
No.		Name
1	<input checked="" type="checkbox"/>	outlook
2	<input type="checkbox"/>	temperature
3	<input type="checkbox"/>	humidity
4	<input type="checkbox"/>	windy
5	<input type="checkbox"/>	play

Remove

No.	Label	Count	Weight
1	sunny	5	5.0
2	overcast	4	4.0
3	rainy	5	5.0

Class: play (Nom)

Visualize All



Status



# Saving the CSV File as ARFF File

