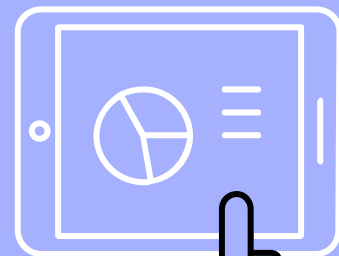
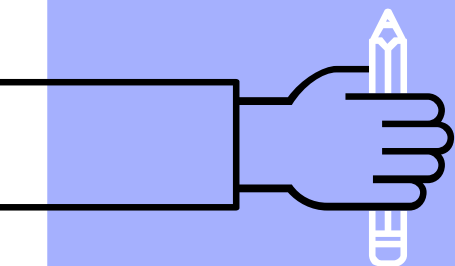
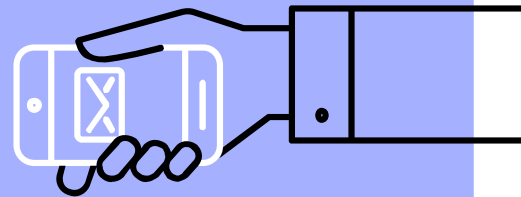


MATLAB Cell Array

Excel Read/Write

盧家鋒 Chia-Feng Lu, Ph.D.
Department of Biomedical Imaging
and Radiological Sciences, NYCU
alvin4016@nycu.edu.tw



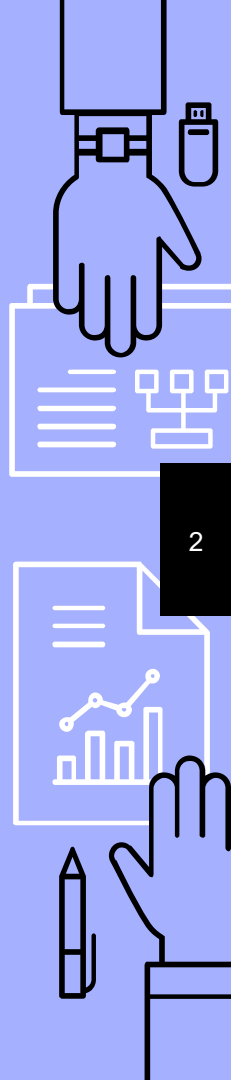


Contents

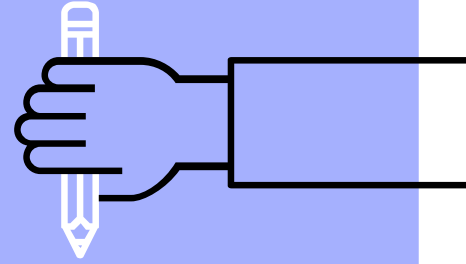
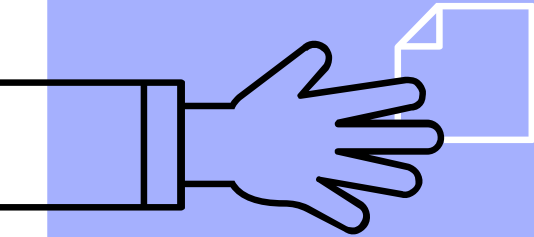
- ▶ Cell array
- ▶ Excel read/write

Please download the handout and materials from

http://cflu.lab.nycu.edu.tw/CFLu_course_matlabimage.html



Cell Array



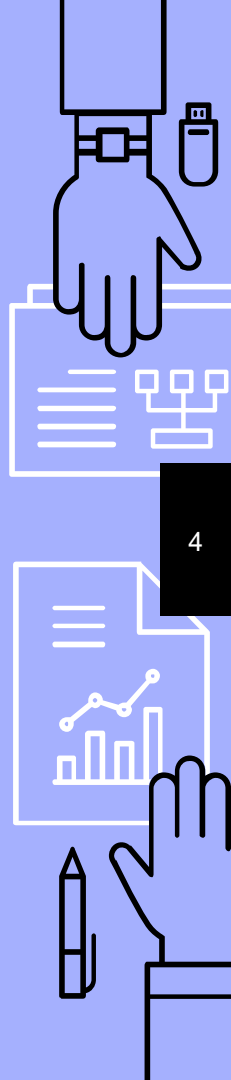


Archive Comprehensive Dataset

For example,

- ▶ 30 normal controls vs. 30 patients with stroke
- ▶ Name, ID → string
- ▶ Age, gender, height, weight → numbers
- ▶ Pre-training data (8 x 450 values) → number array
- ▶ Post-training data (8 x 450 values) → number array

Try to import **ExampleCell.mat** file.
(drag and drop to **Workspace** or use **load**)





Cell Array

- ▶ Store different data formats in a single cell array.
- ▶ Numbers, strings, arrays

NC(2, 6)(1,:) plot bar area pie hist semilc

SELECTION 8x450 double

	1	2	3	4	5	6	7	8
1	810.9025	716.4324	893.5917	597.1968	62.7335	53.5499	208.8670	879.5
2	200.2705	625.3304	684.8781	349.2750	674.7949	259.3526	220.9190	366.7
3	253.9858	615.1605	903.7744	648.5892	865.8245	429.9636	232.5303	519.3
4	959.4531	123.4217	304.1614	894.7045	620.3270	837.7049	203.9535	550.4
5	151.8111	20.7497	715.9496	492.9356	991.8748	644.5207	24.6375	537.9
6	4.1648	971.4546	444.4127	593.4959	121.1350	886.8045	373.2276	922.3
7	702.3257	480.2098	629.6892	747.3605	111.0037	251.1228	335.0475	953.2
8	425.1872	194.5643	912.9675	857.6814	567.9130	364.5518	64.0875	565.2

31x7 cell

	1	2	3	4	5	6	7	8
1	'name'	'age'	'gender'	'height'	'weight'	'predata'	'postda...	
2	'NC01'	62	1	172.1076	69.4049	8x450 d...	8x450 d...	
3	'NC02'	57	0	170.2542	52.2071	8x450 d...	8x450 d...	
4	'NC03'	58	1	174.2668	72.6109	8x450 d...	8x450 d...	
5	'NC04'	63	1	173.4476	69.0174	8x450 d...	8x450 d...	
6	'NC05'	65	1	168.8933	58.8383	8x450 d...	8x450 d...	
7	'NC06'	52	1	168.6337	58.7737	8x450 d...	8x450 d...	
8	'NC07'	65	0	165.8925	79.0742	8x450 d...	8x450 d...	
9	'NC08'	54	0	169.7196	79.6230	8x450 d...	8x450 d...	
10	'NC09'	52	0	174.6123	61.6098	8x450 d...	8x450 d...	
11	'NC10'	68	1	162.9426	64.0754	8x450 d...	8x450 d...	
12	'NC11'	68	1	163.8363	69.6155	8x450 d...	8x450 d...	
13	'NC12'	63	1	172.5047	79.3612	8x450 d...	8x450 d...	
14	'NC13'	55	1	173.8082	77.4354	8x450 d...	8x450 d...	
15	'NC14'	51	0	169.6144	77.6480	8x450 d...	8x450 d...	
16	'NC15'	54	1	169.6135	67.1950	8x450 d...	8x450 d...	

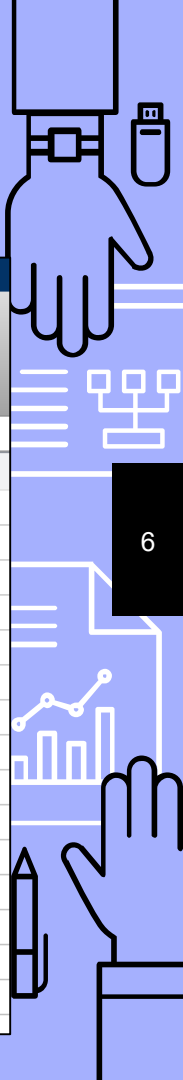
Cell vs. Structure Array

31x7 cell

	1	2	3	4	5	6	7	8
1	'name'	'age'	'gender'	'height'	'weight'	'predata'	'postda...	
2	'NC01'	62	1	172.1076	69.4049	8x450 d...	8x450 d...	
3	'NC02'	57	0	170.2542	52.2071	8x450 d...	8x450 d...	
4	'NC03'	58	1	174.2668	72.6109	8x450 d...	8x450 d...	
5	'NC04'	63	1	173.4476	69.0174	8x450 d...	8x450 d...	
6	'NC05'	65	1	168.8933	58.0383	8x450 d...	8x450 d...	
7	'NC06'	52	1	168.6337	58.7737	8x450 d...	8x450 d...	
8	'NC07'	65	0	165.8925	79.0742	8x450 d...	8x450 d...	
9	'NC08'	54	0	169.7196	79.6230	8x450 d...	8x450 d...	
10	'NC09'	52	0	174.6123	61.6098	8x450 d...	8x450 d...	
11	'NC10'	68	1	162.9426	64.0754	8x450 d...	8x450 d...	
12	'NC11'	68	1	163.8363	69.6155	8x450 d...	8x450 d...	
13	'NC12'	63	1	172.5047	79.3612	8x450 d...	8x450 d...	
14	'NC13'	55	1	173.8082	77.4354	8x450 d...	8x450 d...	
15	'NC14'	51	0	169.6144	77.6480	8x450 d...	8x450 d...	
16	'NC15'	54	1	169.6435	67.4950	8x450 d...	8x450 d...	

1x30 struct with 7 fields

Fields	abc	name	age	gender	height	weight	predata	postdata
1		'NC01'	67	1	168.9951	79.9885	8x450 d...	8x450 d...
2		'NC02'	59	1	170.9523	63.6980	8x450 d...	8x450 d...
3		'NC03'	57	1	161.8978	52.8100	8x450 d...	8x450 d...
4		'NC04'	63	1	168.9596	57.9199	8x450 d...	8x450 d...
5		'NC05'	68	0	167.8267	62.2405	8x450 d...	8x450 d...
6		'NC06'	67	1	169.2874	68.6386	8x450 d...	8x450 d...
7		'NC07'	57	1	173.4562	59.3461	8x450 d...	8x450 d...
8		'NC08'	65	1	161.4962	51.5521	8x450 d...	8x450 d...
9		'NC09'	63	1	163.6110	51.5550	8x450 d...	8x450 d...
10		'NC10'	63	1	167.6989	73.5794	8x450 d...	8x450 d...
11		'NC11'	61	0	160.4540	65.5484	8x450 d...	8x450 d...
12		'NC12'	60	1	162.0846	68.6749	8x450 d...	8x450 d...
13		'NC13'	59	0	165.8496	68.3210	8x450 d...	8x450 d...
14		'NC14'	50	0	164.8730	53.5264	8x450 d...	8x450 d...
15		'NC15'	52	1	172.6919	57.2291	8x450 d...	8x450 d...
16		'NC16'	53	0	163.8809	59.1720	8x450 d...	8x450 d...

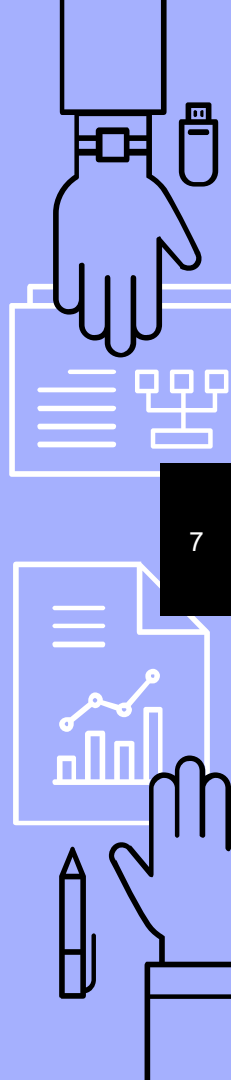




Create a Cell array

When dealing with different length of strings...

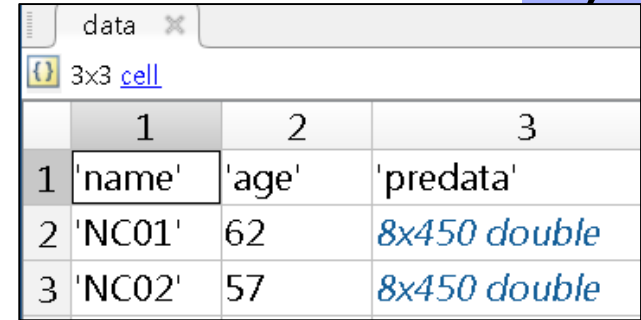
- ▶ `List = ['David '; 'Andy '; 'Jay '; 'Jolin '; 'Selina'];`
- ▶ `List = {'David'; 'Andy'; 'Jay'; 'Jolin'; 'Selina'};`



Create a Cell array

- ▶ `data={};`
- ▶ `data(1,1)={'name'};` % string
- ▶ `data(1,2)={'age'};` % string
- ▶ `data(1,3)={'predata'};` % string

- ▶ `data(2,1)={'NC01'};` % string
- ▶ `data(2,2)={62};` % number
- ▶ `data(2,3)={rand(8,450)};` % number array



data x

3x3 cell

	1	2	3
1	'name'	'age'	'predata'
2	'NC01'	62	8x450 double
3	'NC02'	57	8x450 double

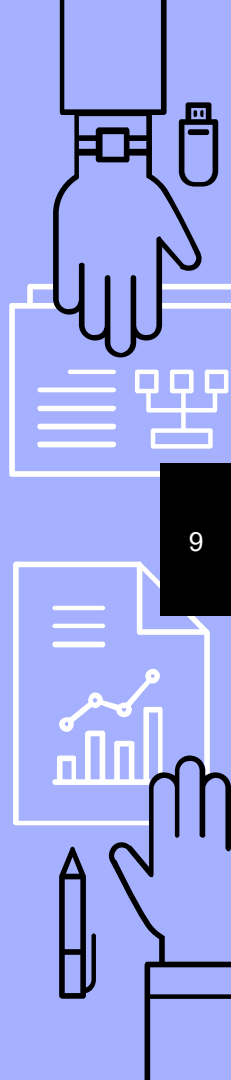
Use variable viewer to check the data structure!



Create a Cell array – Different Ways

- ▶ `data(3,1)={'NC02'};` % string
- ▶ `data(3,2)={57};` % number
- ▶ `data(3,3)={rand(8,450)};` % number array

- ▶ `data{3,1}='NC02';` % string
- ▶ `data{3,2}=57;` % number
- ▶ `data{3,3}=rand(8,450);` % number array

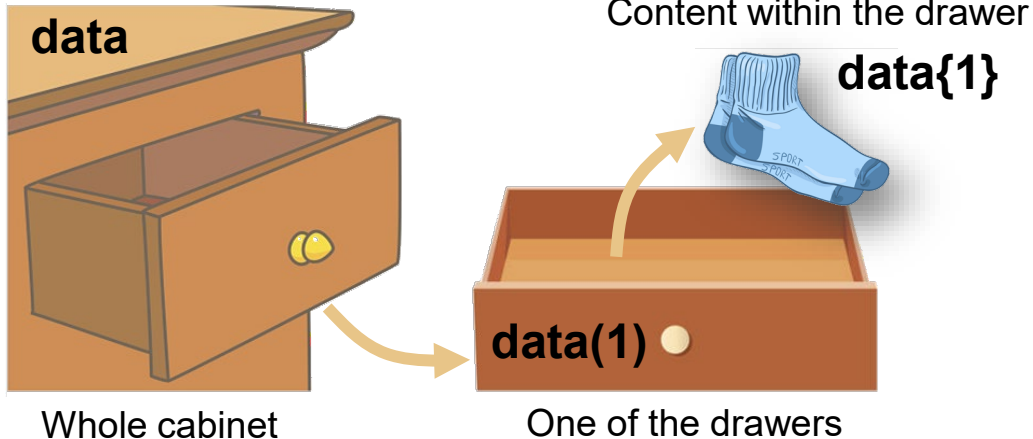


Index of a cell array

- ▶ Is there any difference between following two commands?

`data(1)` → **cell array**

`data{1}` → **string**



```
Command Window
>> data(1)

ans =

1×1 cell array

{'name'}

>> data{1}

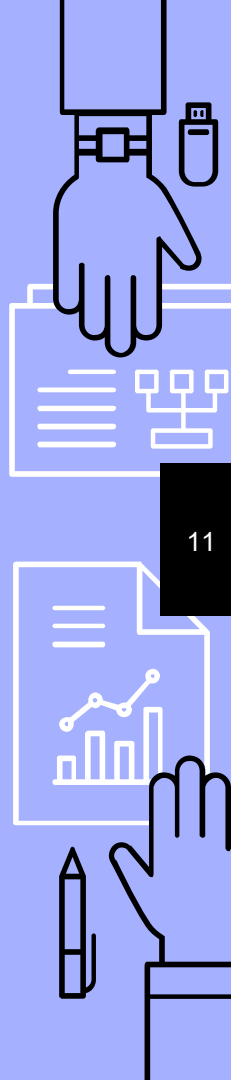
ans =

'name'
```



Useful Functions

- ▶ **cell**
 - Create empty cell array
- ▶ **iscell**
 - True for cell array
- ▶ Try it...
 - `iscell(data(1))`
 - `iscell(data{1})`



Index & Matrix Operations

VARIABLE	SIZE
1x2 cell	
1	2
6x4 double	5x7 double

5x7 double							
	1	2	3	4	5	6	7
1	0.1577	0.1227	0.8109	0.6326	0.3948	0.6181	0.1191
2	0.4218	0.2512	0.0484	0.2468	0.9616	0.0700	0.0431
3	0.5961	0.9376	0.4147	0.8400	0.0082	0.1427	0.1656
4	0.3223	0.6552	0.7255	0.8367	0.3984	0.6542	0.5111
5	0.8307	0.7529	0.1390	0.4531	0.4878	0.2183	0.8673

6x4 double				
	1	2	3	4
1	0.1068	0.3605	0.7252	0.2997
2	0.7349	0.7185	0.9120	0.6019
3	0.3546	0.9953	0.0864	0.1985
4	0.5806	0.0837	0.5200	0.6568
5	0.2989	0.4682	0.5526	0.6984
6	0.7137	0.0938	0.6851	0.4597

$$A\{1,1\}(2:4,2:3) + A\{1,2\}(2:4,5:6)$$

$$A\{1,1\}(2:4,2:3) - A\{1,2\}(2:4,5:6)$$

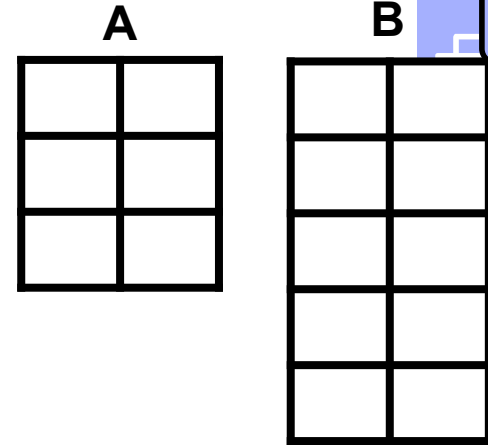
$$A\{1,1\}(2:4,2:3) .* A\{1,2\}(2:4,5:6)$$

$$A\{1,1\}(2:4,2:3) ./ A\{1,2\}(2:4,5:6)$$

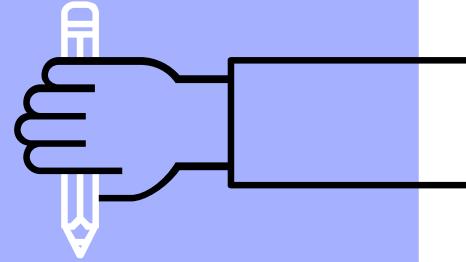
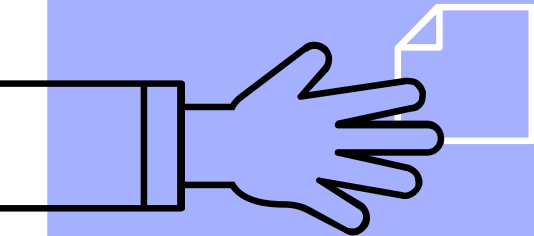
Concatenation of Cell Arrays

- ▶ The same rule as usual...

Command Window	Command Window
<pre>>> A=zeros(3,2); >> B=zeros(5,2); >> C=[A B]; Error using horzcat Dimensions of matrices being concatenated are not consistent. >> C=[A; B]; fx >></pre>	<pre>>> A=cell(3,2); >> B=cell(5,2); >> C=[A B]; Error using horzcat Dimensions of matrices being concatenated are not consistent. >> C=[A; B]; fx >></pre>



Excel Read/Write



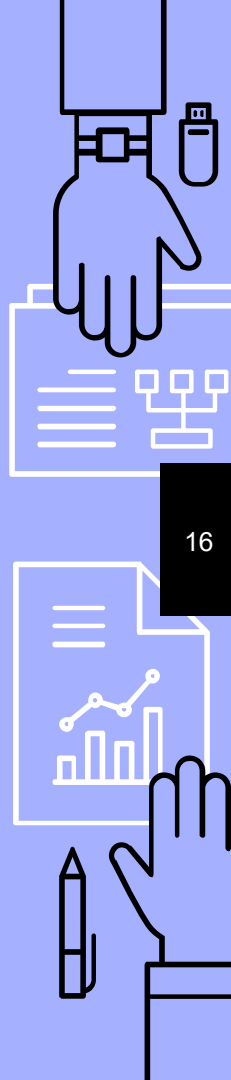


Step 1 – Read Excel File

xlsread Read Microsoft Excel spreadsheet file.

[NUM,TXT,RAW]=xlsread(FILE) reads data from the first worksheet in the Microsoft Excel spreadsheet file named FILE and returns the numeric data in array NUM. Optionally, returns the text fields in cell array TXT, and the unprocessed data (numbers and text) in cell array RAW.

[NUM,TXT,RAW]=xlsread(FILE,SHEET,RANGE) reads from the specified SHEET and RANGE. Specify RANGE using the syntax 'C1:C2', where C1 and C2 are opposing corners of the region.



Step 1 – Import Excel File

- ▶ Read the data from TCGAGBM_data.xlsx
- ▶ `[num1,txt1,raw1]=xlsread('TCGAGBM_data.xlsx');`

	A	B	C	D	E
1	Case	Histology	Grade	Age (year)	Gender
2	TCGA-02-0003	glioblastoma	G4	50	male
3	TCGA-02-0006	glioblastoma	G4	56	female
4	TCGA-02-0009	glioblastoma	G4	61	female
5	TCGA-02-0011	glioblastoma	G4	18	female
6	TCGA-02-0027	glioblastoma	G4	33	female
7	TCGA-02-0033	glioblastoma	G4	54	male
8	TCGA-02-0034	glioblastoma	G4	60	male
9	TCGA-02-0037	glioblastoma	G4	74	female
10	TCGA-02-0046	glioblastoma	G4	61	male

	1	2	3	4	5
1	'Case'	'Histology'	'Grade'	'Age ...'	'Gender'
2	'TCGA-02-0003'	'glioblastoma'	'G4'	50	'male'
3	'TCGA-02-0006'	'glioblastoma'	'G4'	56	'female'
4	'TCGA-02-0009'	'glioblastoma'	'G4'	61	'female'
5	'TCGA-02-0011'	'glioblastoma'	'G4'	18	'female'
6	'TCGA-02-0027'	'glioblastoma'	'G4'	33	'female'
7	'TCGA-02-0033'	'glioblastoma'	'G4'	54	'male'
8	'TCGA-02-0034'	'glioblastoma'	'G4'	60	'male'
9	'TCGA-02-0037'	'glioblastoma'	'G4'	74	'female'
10	'TCGA-02-0046'	'glioblastoma'	'G4'	61	'male'

Step 1 – Import Excel File

- ▶ Read the data from features sheet
- ▶ `[num2,txt2,raw2]=xlsread('TCGAGBM_data.xlsx','features');`

	A	B	C	D	E
1	Subjects	TCGA-0...	TCGA-0...	TCGA-0...	TCGA-0...
2	1. Energy	1.87E+10	9.03E+08	1.6E+09	6.48E+09
3	2. Entropy	8.682872	7.065196	7.262475	7.519484
4	3. Kurtosis	8722566	1365234	197227.5	679031.5
5	4. Maximum	1218	679	589	790
6	5. Mean	655.4755	339.757	395.7246	329.832
7	6. Mean absolute deviation	87.93104	28.52226	31.66313	34.99748
8	7. Median	631	333	401	327
9	8. Minimum	241	207	222	189
10	9. Range	977	472	367	601
11	10. Root mean square	666.6547	342.8137	397.653	333.2275

	1	2	3	4	5
1	'Subjects'	'TCGA-...	'TCGA-...	'TCGA-...	'TCGA-...
2	'1. Energy'	1.8651...	90326...	1.6028...	6.4784...
3	'2. Entropy'	8.6829	7.0652	7.2625	7.5195
4	'3. Kurtosis'	8.7226...	1.3652...	1.9723...	6.7903...
5	'4. Maximum'	1218	679	589	790
6	'5. Mean'	655.47...	339.75...	395.72...	329.83...
7	'6. Mean absolute devia...	87.9310	28.5223	31.6631	34.9975
8	'7. Median'	631	333	401	327
9	'8. Minimum'	241	207	222	189
10	'9. Range'	977	472	367	601
11	'10. Root mean square'	666.65...	342.81...	397.65...	333.22...

Step 1 – Import Excel File

- ▶ Please compare the raw2 and num2. What's the difference?

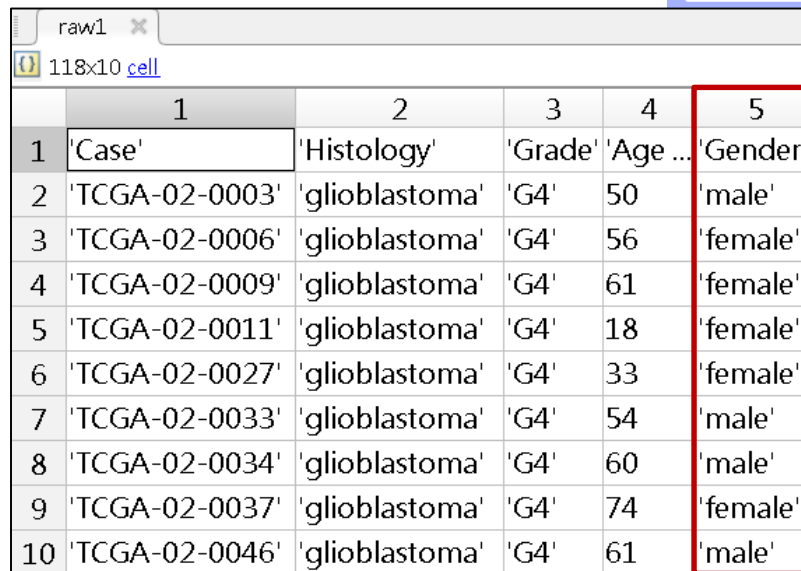
num2 22x117 double
raw2 23x118 cell

	1	2	3	4	5	6
1	Subjec...	'TCGA-...	'TCGA-...	'TCGA-...	'TCGA-...	'TCGA-...
2	1. Ener...	1.8651...	90326...	1.6028...	6.4784...	4.0471...
3	2. Entr...	8.6829	7.0652	7.2625	7.5195	11.3252
4	3. Kurt...	8.7226...	1.3652...	1.9723...	6.7903...	1.6591...
5	4. Max...	1218	679	589	790	7488
6	5. Mea...	655.47...	339.75...	395.72...	329.83...	3.5748...
7	6. Mea...	87.9310	28.5223	31.6631	34.9975	561.84...
8	7. Me...	631	333	401	327	3481
9	8. Mini...	241	207	222	189	1518
10	9. Ran...	977	472	367	601	5970
11	10. Ro...	666.65...	342.81...	397.65...	333.22...	3.6497...

	1	2	3	4	5	6
1	1.8651...	90326...	1.6028...	6.4784...	4.0471...	7.0471...
2	8.6829	7.0652	7.2625	7.5195	11.3252	8.6829
3	8.7226...	1.3652...	1.9723...	6.7903...	1.6591...	2.1659...
4	1218	679	589	790	7488	1218
5	655.47...	339.75...	395.72...	329.83...	3.5748...	403.57...
6	87.9310	28.5223	31.6631	34.9975	561.84...	87.9310
7	631	333	401	327	3481	631
8	241	207	222	189	1518	241
9	977	472	367	601	5970	977
10	666.65...	342.81...	397.65...	333.22...	3.6497...	413.65...

Step 2 – Find Targets from Excel Data

- ▶ Extract data of female subjects from all 117 subjects.
- ▶ `female_check=ismember(raw1(:,5),'female');`
- ▶ `female_ind=find(female_check);`



	1	2	3	4	5
1	'Case'	'Histology'	'Grade'	'Age ...'	'Gender'
2	'TCGA-02-0003'	'glioblastoma'	'G4'	50	'male'
3	'TCGA-02-0006'	'glioblastoma'	'G4'	56	'female'
4	'TCGA-02-0009'	'glioblastoma'	'G4'	61	'female'
5	'TCGA-02-0011'	'glioblastoma'	'G4'	18	'female'
6	'TCGA-02-0027'	'glioblastoma'	'G4'	33	'female'
7	'TCGA-02-0033'	'glioblastoma'	'G4'	54	'male'
8	'TCGA-02-0034'	'glioblastoma'	'G4'	60	'male'
9	'TCGA-02-0037'	'glioblastoma'	'G4'	74	'female'
10	'TCGA-02-0046'	'glioblastoma'	'G4'	61	'male'

Step 2 – Find Targets from Excel Data

```
5 %% Import data from two sheets
6 [num1,txt1,row1]=xlsread('TCGAGBM_data.xlsx','TCGA_data');
7 [num2,txt2,row2]=xlsread('TCGAGBM_data.xlsx','features');
8
9 %% Find all female subjects
10 female_check=ismember(row1(:,5),'female');
11 female_ind=find(female_check==1);
12 female_data=row1(female_ind,:);
```

	1	2	3	4	5
1	'TCGA-02-0006'	'glioblastoma'	'G4'	56	'female'
2	'TCGA-02-0009'	'glioblastoma'	'G4'	61	'female'
3	'TCGA-02-0011'	'glioblastoma'	'G4'	18	'female'
4	'TCGA-02-0027'	'glioblastoma'	'G4'	33	'female'
5	'TCGA-02-0037'	'glioblastoma'	'G4'	74	'female'
6	'TCGA-02-0054'	'glioblastoma'	'G4'	44	'female'
7	'TCGA-02-0060'	'glioblastoma'	'G4'	66	'female'
8	'TCGA-02-0085'	'glioblastoma'	'G4'	63	'female'
9	'TCGA-02-0086'	'glioblastoma'	'G4'	45	'female'
10	'TCGA-06-0119'	'glioblastoma'	'G4'	81	'female'

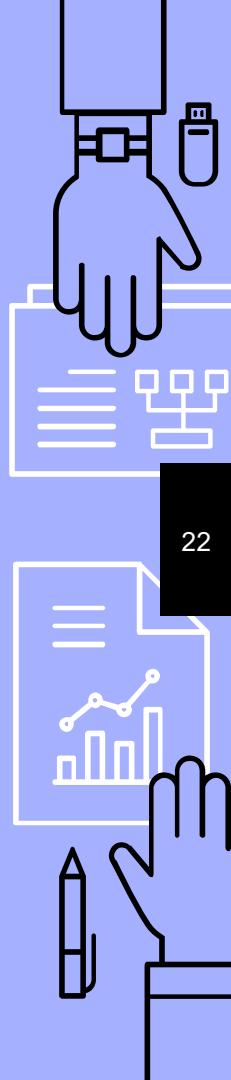


Step 3 – Export Data as Excel Sheet/File

xlsxwrite Write to Microsoft Excel spreadsheet file.

xlsxwrite(FILE,ARRAY,SHEET) writes ARRAY to the specified worksheet in the Excel file named FILE, starting at cell A1.

xlsxwrite(FILE,ARRAY,SHEET,RANGE) writes to the specified SHEET and RANGE.



Step 3 – Export Data as Excel Sheet/File

- ▶ Export the extracted female data to the **female** sheet.
- ▶ `xlswrite('TCGAGBM_data.xlsx', female_data, 'female');`

Command Window

```
>> ExcellEx06
```

```
Error using xlswrite (line 224)
```

```
The file D:\CloudStation\YM\10701Matlab程式語言\Lesson6_異質矩陣-Excel資料讀寫\MImaterials_L6\TCGAGBM_data.xlsx is not writable. It might be locked by another process.
```

Do not open **TCGAGBM_data.xlsx** in Excel while exporting data to it!

Step 3 – Export Data as Excel Sheet/File

```

9      %% Find all female subjects
10     female_check=ismember(raw1(:,5),'female');
11     female_ind=find(female_check==1);
12     female_data=raw1(female_ind,:);
13
14     %% Export the extracted female data to the female sheet
15     xlswrite('TCGAGBM_data.xlsx', female_data, 'female');
```

Command Window

Warning: Added specified worksheet.
 > In `xlswrite`> `activate sheet` (line 298)
 In `xlswrite/ExecuteWrite` (line 264)
 In `xlswrite` (line 218)
 In `ExcelEx06` (line 15)

26	TCGA-08-0355	glioblastoma	G4	30	female	24.542
27	TCGA-08-0356	glioblastoma	G4	59	female	31.080
28	TCGA-08-0359	glioblastoma	G4	59	female	3.3840

TCGA_data | features | female (+)

Only works on Microsoft OS with Excel installed.

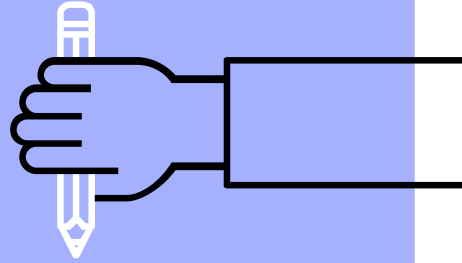


Homework

- ▶ Calculate mean values of 22 features for all female subjects.
- ▶ HINTS:
 - The subjects are arranged in row fashion in the **TCGA_data** sheet; the subjects are arranged in a **column** fashion in the **features** sheet.
 - Do remember the differences between **raw1** and **num2**.

- **Modify from ExcelEx06.m**

	meanfeature
	22x1 double
	1
1	2.9335e+10
2	8.6521
3	2.4037e+08
4	1.5051e+03
5	663.3037
6	152.3429
7	616.5367
8	238.9703
9	1.2662e+03
10	694.5514
11	1.0621
12	192.3036
13	0.0045
14	8.4790e+04
15	68.2478
16	0.2074
17	0.0000



THE END

alvin4016@nycu.edu.tw

