

Using Excel as Journalists

Hong Kong Baptist University, Hong Kong

zhangtianyi@hkbu.edu.hk / ydu@hkbu.edu.hk

Learning Objectives

By the end of this tutorial, you will learn how to

1. **Import data** from different resources and of different format using Excel;
2. **Analysis data** using simple operations and formulas in Excel;
3. **Summarize data** using pivot tables in Excel.

Preparation steps

1. You need to build up your own mind-map of the Excel toolbar.

Here is an example of the mind-map: <http://goo.gl/VWK7wQ>

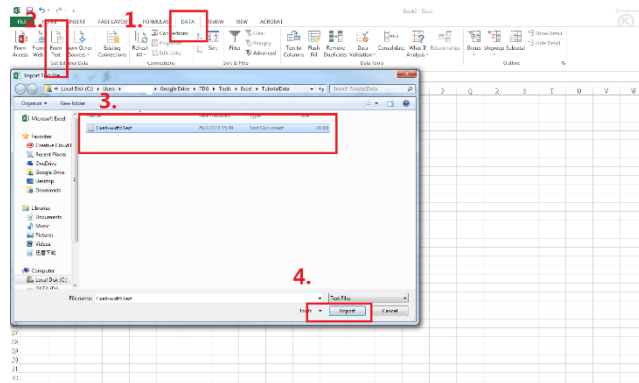
PART 1: Importing Data

In this part, you are going to import data from:

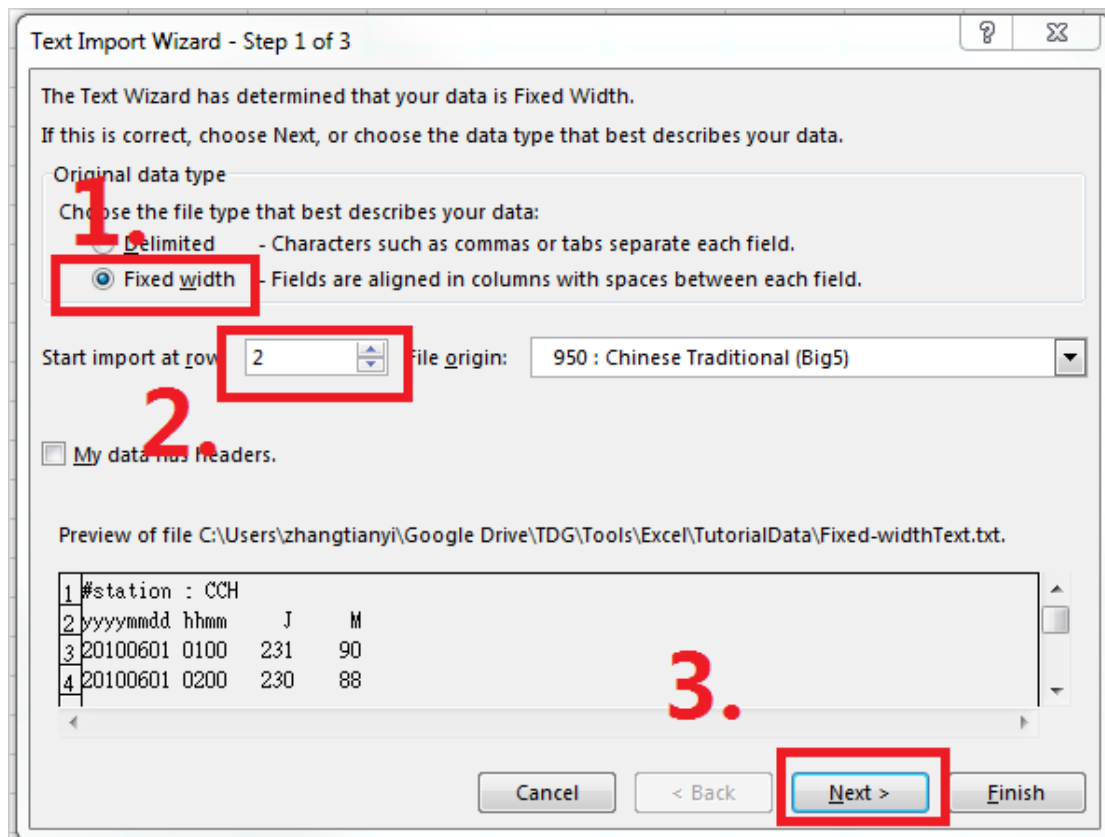
1. Text file (Fixed-width text)
2. CVS file (Delimited text)
3. HTML tables
4. Tables from PDF.

Step 1.1: **Download** the fixed-width text file: <https://goo.gl/Av1Fd9>, you will see the data is separated by fixed-width if you open it.

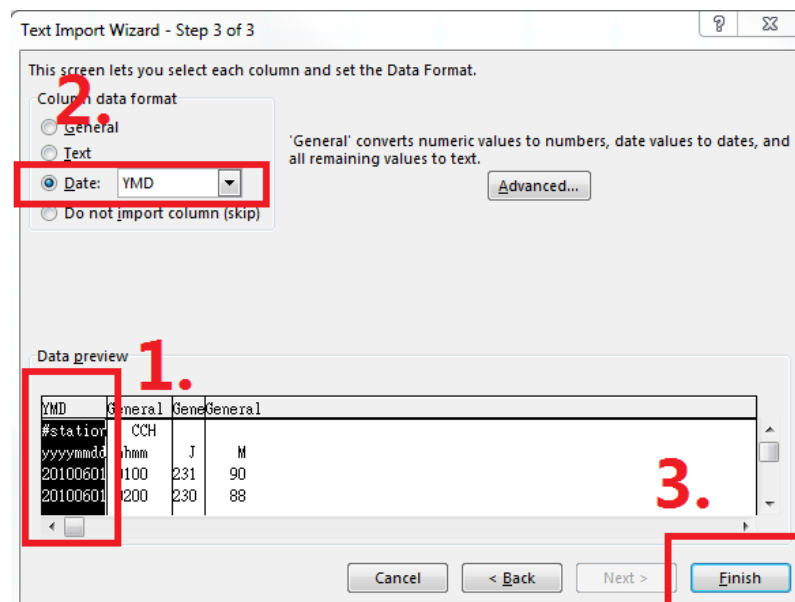
Then **open** a blank spreadsheet, go to tab **DATA->From Text**, choose the file you just downloaded, click **Import**.



Then you will see the window below, choose “**Fixed width**”, and because the 1st row of data is useless, we choose to start import at row “**2**”, and then click “**Next**”. Click “**Next**” again.



You will see that the 1st column of the data is “yyyymmdd” (Year+Month+Day), so choose the 1st column of the data, and choose it as “**Date**”->YMD. Finally, click “**Finish**”. Click “**OK**” if there is no problem.



The result is like below:

yyyyymmdd	hhmm	J	M						
1/5/2010	100	231	90						
1/5/2010	200	230	88						
1/5/2010	300	234	88						
1/5/2010	400	233	88						
1/5/2010	500	233	86						
1/5/2010	600	233	86						
1/5/2010	700	235	86						
1/5/2010	800	236	87						
1/5/2010	900	241	86						
1/5/2010	1000	242	85						
1/5/2010	1100	241	82						
1/5/2010	1200	244	80						
1/5/2010	1300	250	77						
1/5/2010	1400	240	80						
1/5/2010	1500	222	97						
1/5/2010	1600	219	98						
1/5/2010	1700	218	98						
1/5/2010	1800	212	98						
1/5/2010	1900	205	95						
1/5/2010	2000	211	89						
1/5/2010	2100	219	89						
1/5/2010	2200	223	89						
1/5/2010	2300	227	93						
1/5/2010	2400	228	94						
2/5/2010	100	229	95						
2/5/2010	200	229	95						
2/5/2010	300	229	94						
2/5/2010	400	229	94						
2/5/2010	500	228	94						
2/5/2010	600	225	98						
2/5/2010	700	221	98						
2/5/2010	800	220	98						
2/5/2010	900	216	97						

Step 1.2: **Download** the CSV file: <https://goo.gl/NRmrMg>

Choose tab **DATA->From Text**, and then choose your downloaded data set, click **“Import”**.

Choose **“Delimited”** this time because the data is separated by tabs, and start import at row **“8”**, because the first seven rows are useless to us. Click **“Next”**.

Text Import Wizard - Step 1 of 3

The **1.** Text Wizard has determined that your data is Fixed Width.
If this is **1.** correct, choose Next, or choose the data type that best describes your data.

Original data type

Choose the file type that best describes your data:

☒ **Delimited** - Characters such as commas or tabs separate each field.

☐ Fixed width - Fields are aligned in columns with spaces between each field.

Start import at row: **2.** 8 File origin: 950 : Chinese Traditional (Big5)

☐ My data has headers.

Preview of file C:\Users\zhangtianyi\Google Drive\TDG\Tools\Excel\TutorialData\CSVFile.csv.

6 "4. In case of station or equipment suspension due to maintenance, the data collection
7
8 Date\Hour\Central\Western\Eastern\Kwun Tong\Sham Shui Po\Kwai Chung\Tsuen Wan\Yuen Lo
9 2014-08-01\01\21213121313*1212131312131313

3.

Cancel < Back **Next >** Finish

As the data is separated by “|”, we need to choose “**Other**”, and type “|” into the box, click “**Next**”. Similarly, like step 1.1, choose the first column and make it **Date**->**YMD** format, and then “**Finish**”. Then “**OK**”.

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

☐ Tab
☐ Semicolon
☐ Comma
☐ Space
☒ **Other:** |

☐ Treat consecutive delimiters as one

Text qualifier: " "

Data preview

Date	Hour	Central/Western	Eastern	Kwun Tong	Sham Shui Po	Kwai Chung	Tsuen Wan	Yue
2014-08-01	01	2	2	3	2	3	3*	2
	02	2	2	2	2	3	2*	2

Buttons: Cancel, < Back, **Next >**, Finish

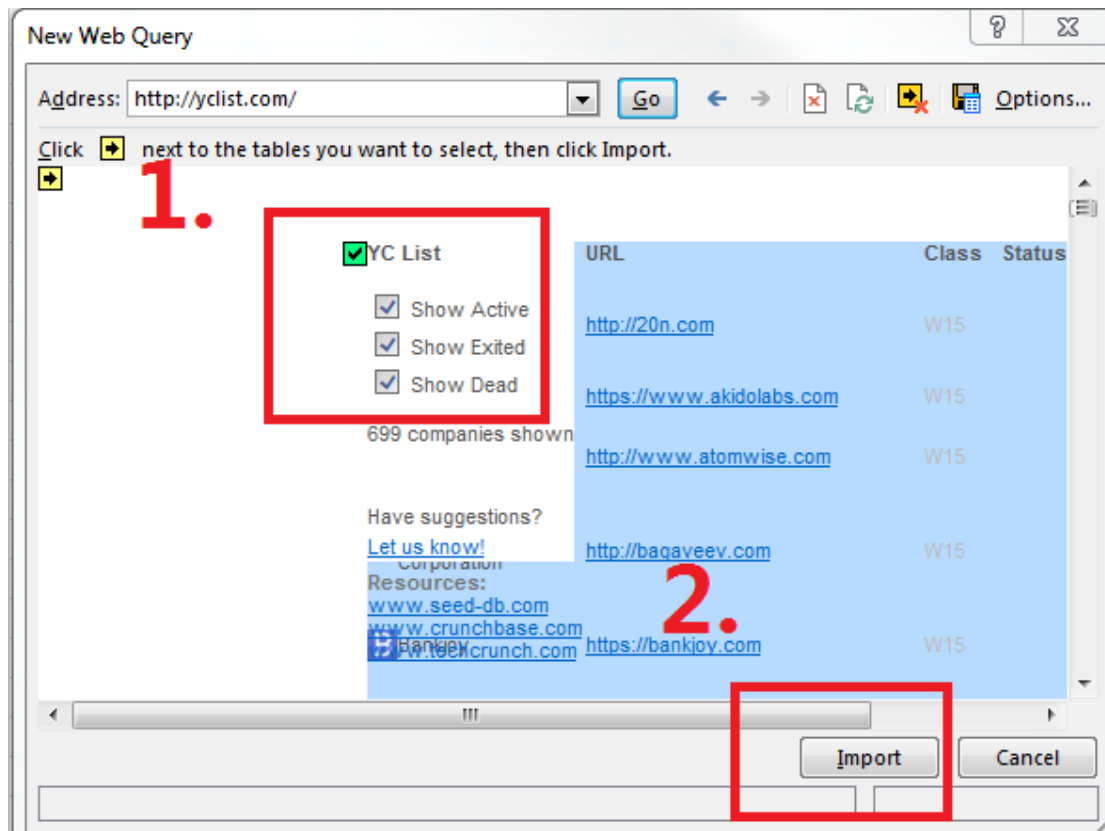
The result is like below:

Date	Hour	Central/Western	Eastern	Kwun Tong	Sham Shui Po	Kwai Chung	Tsuen Wan	Yuen Long	Tuen Mun	Tung Chung	Tai Po	Sha Tin	Tap Mun	Causeway Bay	Central	Mong Kok
1/8/2014	1	2	2	3	2	3 3*		2	2	2	3	2	3		3	3
	2	2	2	2	2	3 2*		2	2	2	3	2	2		3	2
	3	2	2	2	2	3 2*		2	2	2	3	2	2		2	2
	4	2	2	2	2	3 2*		2	2	2	2	2	2		2	2
	5	1	2	2	2	2 2*		2	2	2	2	2	2		2	2
	6	1	2	2	2	2 2*		2	2	2	2	2	2		2	2
	7	1	2	2	2	2 2*		2	2	2	2	2	2		2	2
	8	2	2	2	2	2 2*		2	2	2	2	2	2		3	2
	9	2	2	2	2	3 2*		2	2	2	2	2	2		4	3
	10	2	2	3	2	3 3*		3	3	3	3	3	3		5	3
	11	3	3	4	3	3 4*		3	3	3	3	3	3		5	4
	12	3	4	4	3	4	4	4	4	4	4	4	4		5	4
	13	4	4	4	4	4	4	5	5	4	5	4	4		6	5
	14	5	5	5	5	6	6	7	6	5	6	5	4		7	5
	15	6	6	7	6	7	8	9	8	7	7	7	5		8	6
	16	6	6	8	6	7	8	10	9	7	8	7	6		8	7
	17	5	6	8	6	6	8	10	8	6	8	7	6		6	7
	18	4	5	7	5	5	6	8	7	5	7	6	6		5	6
	19	4	5	6	5	5	5	7	6	5	6	6	5		6	6
	20	5	5	6	5	5	5	6	6	5	5	5	5		7	6
	21	5	5	6	5	5	5	6	6	5	5	5	4		7	6
	22	5	5	6	5	5	5	6	6	4	5	5	4		7	6
	23	5	5	5	5	5	5	5	5	4	4	5	4		7	6
	24	5	5	5	5	5	5	5	5	4	4	4	4		7	6
	Daily Max	6	6	8	6	7	8	10	9	7	8	7	6		8	7
2/8/2014	1	5	4	5	5	5	4	5	5	4	4	4	4		6	6
	2	5	4	5	5	4	4	5	4	3	4	4	3		6	5
	3	4	4	5	5	4	4	5	4	3	4	4	3		5	5

Step 1.3: Go to link: <http://yclist.com/>, you will see at HTML table on the web page.

In Excel, go to **DATA->From Web**, paste the link to the “**Address**” and click “**Go**” button beside it.

When you hover on the yellow arrow beside the table, you will find it becomes green, **click on the green arrow** to choose the table, then click “**Import**”. Then “**OK**”.



You will see the result:

A	B	C	D	E	F
Name	URL	Class	Status	Description	
20n	http://20n.com	W15		A computational synthetic biology company	
Akido Labs	https://www.akidolabs.com	W15		The standard API for hospitals	
Atomwise	http://www.atomwise.com	W15		Bold funding for big discoveries	
Bagaveev Corporation	http://bagaveev.com	W15		3D printed rocket engines will power a nanosatellite launcher rocket	
Bankjoy	https://bankjoy.com	W15		Mobile banking for credit unions	
BlueSmart	http://bluesmart.com	W15		The world's first smart, connected luggage	
Bonfire	http://gobonfire.com	W15		Bonfire is an RFx tool that lets purchasing teams accept and evaluate supplier submissions quickly	
Booktrope	http://booktrope.com	W15		Booktrope has pioneered an entirely new choice in book publishing	
Bright	http://www.brightnetwork.co.uk	W15		An exclusive careers network	
BuildScience	http://buildscience.com	W15		Building management at your fingertips	
Chariot	https://www.ridechariot.com	W15		The world's first crowdfunded network of commuter routes	
Cinder	http://www.cindercooks.com	W15		Notifications when food is done. All in a countertop electric grill.	
Cleanly	http://www.getcleanly.com	W15		Laundry & dry-cleaning delivered at the tap of a button	
CloudMedX	http://www.cloud-medx.com	W15		Improving patient outcomes through data	
CribSpot	https://www.cribspot.com	W15		Search 1000s of local houses and apartments	
Dealzye	http://dealzye.com	W15		Punch cards for the 21st century	
DemocracyOS	http://democracyos.org	W15		Get from zero to democracy in a click. Build proposals and be the change you want to see	
Detroit Water Project	https://www.detroitwaterproject.org	W15		You can help keep a family's water on	

Step 1.4: Download the PDF file: <https://goo.gl/cseypl>

Choose any table you want in the file, **right lick->Copy with formatting->Paste in Excel**, then you will get the table in Excel.

PART 2: Analyze Data with Basic Operations and Formulas

Download the data here: <https://goo.gl/Dao1rV>

Operations:

Sorting & Filtering

Step 2.1: (Sort by 1 variable) **Open** the downloaded data, find sheet “**ItalyCrimes**”, put the cursor in column “**Province**”, click “**Sort & Filter**”, choose “**A->Z**”, you will get the results with the alphabetic order.

Notice: Remember to put the cursor in the column, but **not** choose the whole column, or you will get the result of only this “Province” to be ordered, which totally disorder your data. Please see if your result is the same as below.

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rapes	Thefts	Robberies	Arsons	Drugs
Sicilia	Agrigento	454,002	14,124	3	43	5,059	133	66	114
Piemonte	Alessandria	440,613	19,213	1	54	8,896	184	82	155
Marche	Ancona	481,028	18,603	1	27	7,774	136	20	299
Valle d'Aosta/Vallée d'Aoste	Aosta	128,230	4,946	0	10	2,045	17	8	51
Toscana	Arezzo	349,651	12,436	4	24	5,356	94	56	159
Marche	Ascoli Piceno	214,068	13,427	1	26	6,706	138	68	181
Piemonte	Asti	221,687	9,116	0	16	3,928	67	14	66
Campania	Avellino	439,137	11,515	4	16	4,191	98	124	131
Puglia	Bari	1,258,706	65,808	15	107	35,789	1,235	592	597
Veneto	Belluno	213,474	5,857	1	11	2,087	8	22	61
Campania	Benevento	287,874	7,294	4	18	2,583	74	218	118
Lombardia	Bergamo	1,098,740	45,360	8	96	23,716	515	148	586
Piemonte	Biella	185,768	7,236	1	14	2,503	46	33	108
Emilia-Romagna	Bologna	991,924	70,225	8	144	39,383	770	126	888
Trentino-Alto Adige	Bolzano/Bozen	507,657	13,790	1	59	6,794	67	37	275
Lombardia	Brescia	1,256,025	64,106	8	126	31,451	777	131	883
Puglia	Brindisi	403,229	15,864	3	32	7,096	170	121	166
Sardegna	Cagliari	563,180	25,497	4	63	10,532	213	105	453
Sicilia	Caltanissetta	271,729	10,610	2	14	4,209	111	119	84
Molise	Campobasso	231,086	6,414	1	10	2,755	17	83	62
Campania	Caserta	916,467	33,134	28	52	15,977	1,672	223	404
Sicilia	Catania	1,090,101	54,618	16	65	35,203	2,003	218	526
Calabria	Catanzaro	368,597	14,917	14	24	5,834	71	374	210
Abruzzo	Chieti	397,123	12,861	0	15	5,667	101	98	117
Lombardia	Como	594,988	19,358	4	44	9,945	186	86	487
Calabria	Cosenza	734,656	25,163	7	51	9,867	254	834	302
Lombardia	Cremona	363,606	13,027	2	31	6,158	105	23	103
Calabria	Crotone	174,605	5,777	11	11	1,625	28	219	86
Piemonte	Cuneo	592,303	18,322	5	38	7,792	122	59	144
Sicilia	Enna	172,485	4,137	3	16	1,518	38	66	40
Emilia-Romagna	Ferrara	359,994	16,712	3	25	8,754	108	40	199
Toscana	Firenze	998,098	57,222	8	104	30,167	561	160	876
Puglia	Foggia	640,836	28,397	16	54	13,636	403	331	253
Emilia-Romagna	Forlì/Cesena	395,489	16,200	0	30	8,196	127	46	245
Lazio	Frosinone	498,167	13,824	1	25	5,206	150	152	196
Liguria	Genova	882,718	58,626	17	84	28,536	635	162	912
Friuli-Venezia Giulia	Gorizia	142,407	4,868	2	14	2,068	26	11	116
Toscana	Grosseto	228,157	9,940	5	15	4,841	36	42	185
Liguria	Imperia	222,648	12,754	1	18	5,768	54	141	254
Molise	Isernia	88,694	2,974	1	3	762	15	70	52
Liguria	La Spezia	223,516	10,826	1	12	5,128	98	84	190
Abruzzo	L'Aquila	309,820	8,944	0	17	3,334	60	49	170

(Sort by 2 variables): Go to **DATA->Sort**, let's sort with “**Territorio**” from “**A->Z**”, with “**Total crimes**” from “largest to smallest”. Click “OK”.

1. DATA tab, 2. Sort button, 3. Sort dialog box, 4. Add Level button, 5. Sort by dropdown, 6. OK button

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rapes	Thefts	Robberies	Arsons	Drugs
Lombardia	Milano	3,156,694	280,809	39	502	168,057	5,508	532	2,567
Lazio	Roma	4,194,068	236,636	40	325	144,511	4,271	381	3,032
Piemonte	Torino	2,302,353	154,536	16	208	74,889	2,691	264	1,660
Campania	Napoli							549	1,971
Emilia-Romagna	Bologna							126	888
Puglia	Bari							592	597
Lombardia	Brescia							131	883
Liguria	Genova							162	912
Toscana	Firenze							160	876
Sicilia	Catania							218	526
Sicilia	Palermo							391	474
Lombardia	Bergamo							148	586
Veneto	Padova							79	1,072
Veneto	Venezia							37	408
Veneto	Verona							127	318
Emilia-Romagna	Modena							100	384
Campania	Salerno	1,109,705	34,747	12	68	16,092	541	511	392
Lombardia	Varese	883,285	34,413	6	61	16,154	318	109	620
Campania	Caserta	916,467	33,134	28	52	15,977	1,672	223	404
Puglia	Foggia	640,836	28,397	16	54	13,636	403	331	253
Veneto	Vicenza	870,740	27,770	5	47	14,852	230	82	187
Lombardia	Pavia	548,307	26,565	3	41	14,343	254	47	278

The result is below:

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rapes	Thefts	Robberies	Arsons	Drugs
Abruzzo	Pescara	323,184	17,635	1	33	8,256	202	55	217
Abruzzo	Chieti	397,123	12,861	0	15	5,667	101	98	117
Abruzzo	Teramo	312,239	12,365	2	25	5,740	108	52	237
Abruzzo	L'Aquila	309,820	8,944	0	17	3,334	60	49	170
Basilicata	Potenza	383,791	9,303	2	29	2,567	39	72	219
Basilicata	Matera	203,726	4,272	0	12	1,486	19	16	63
Calabria	Cosenza	734,656	25,163	7	51	9,867	254	834	302
Calabria	Reggio di Calabria	566,977	19,919	34	33	7,766	331	448	275
Calabria	Catanzaro	368,597	14,917	14	24	5,834	71	374	210
Calabria	Vibo Valentia	166,560	6,466	10	12	1,882	76	78	75
Calabria	Crotone	174,605	5,777	11	11	1,625	28	219	86
Campania	Napoli	3,080,873	141,705	63	184	69,868	11,152	549	1,971
Campania	Salerno	1,109,705	34,747	12	68	16,092	541	511	392
Campania	Caserta	916,467	33,134	28	52	15,977	1,672	223	404
Campania	Avellino	439,137	11,515	4	16	4,191	98	124	131
Campania	Benevento	287,874	7,294	4	18	2,583	74	218	118
Emilia-Romagna	Bologna	991,924	70,225	8	144	39,383	770	126	888
Emilia-Romagna	Modena	700,913	37,489	6	53	22,357	366	100	384
Emilia-Romagna	Reggio nell'Emilia	530,343	22,896	3	47	12,507	214	38	178
Emilia-Romagna	Rimini	329,302	22,438	4	47	13,031	329	44	316
Emilia-Romagna	Ravenna	392,458	21,557	2	31	12,230	207	58	268
Emilia-Romagna	Parma	442,120	19,929	3	38	10,211	159	54	185
Emilia-Romagna	Ferrara	359,994	16,712	3	25	8,754	108	40	199
Emilia-Romagna	Forlì-Cesena	395,489	16,200	0	30	8,196	127	46	245
Emilia-Romagna	Piacenza	289,875	10,444	1	13	5,510	99	23	130
Friuli-Venezia Giulia	Udine	541,522	16,409	1	36	8,169	71	26	207
Friuli-Venezia Giulia	Trieste	236,556	10,121	0	22	4,642	61	16	134
Friuli-Venezia Giulia	Pordenone	315,323	9,709	2	29	5,291	37	21	58
Friuli-Venezia Giulia	Gorizia	142,407	4,868	2	14	2,068	26	11	116

Filtering: Choose the header row, go to **DATA** tab, and click **“Filter”**, the arrows will appear

automatically. Suppose you only would like to see the records of **Abruzzo** from **Territorio**, click the small arrow, and choose **“Abruzzo”** only. Then you can only see four records.

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rape	Thefts	Robberies	Arson	Drug
Abruzzo	Pescara	323,184	17,635	1	33	8,256	202	55	217
Abruzzo	Chieti	397,123	12,861	0	15	5,667	101	98	117
Abruzzo	Teramo	312,239	12,365	2	25	5,740	108	52	237
Abruzzo	L'Aquila	309,820	8,944	0	17	3,334	60	49	170

Or you want to see the total crimes that are larger than 5000 and smaller than 10000, you can clear the filter first from **Territorio**: click the small arrow on **Territorio**, then click **“Clear Filter from Territorio”**.

Then click the arrow on **“Total crimes”->Number Filters->Custom Filter**

Book1 - Excel

1. 2. 3.

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rape	Thefts	Robberies	Arson	Drug
Abruzzo	Pescara			1	33	8,256	202	55	217
Abruzzo	Chieti			0	15	5,667	101	98	117
Abruzzo	Teramo			2	25	5,740	108	52	237
Abruzzo	L'Aquila			0	17	3,334	60	49	170
Basilicata	Potenza			2	29	2,567	39	72	219
Basilicata	Matera			0	12	1,486	19	16	63
Calabria	Cosenza			1,867	254	834	302		
Calabria	Reggio di Calabria			7,766	331	448	275		
Calabria	Catanzaro			3,834	71	374	210		
Calabria	Vibo Valentia			1,882	76	78	75		
Calabria	Crotone			1,625	28	219	86		
Campania	Napoli			1,868	11,152	549	1,971		
Campania	Salerno			1,092	541	511	392		
Campania	Caserta			1,977	1,672	223	404		
Campania	Avellino			1,191	98	124	131		
Campania	Benevento			1,583	74	218	118		
Emilia-Romagna	Bologna			7,383	770	126	888		
Emilia-Romagna	Modena			357	366	100	384		
Emilia-Romagna	Reggio nell'Emilia			507	214	38	178		
Emilia-Romagna	Rimini			4	47	13,031	329	44	316
Emilia-Romagna	Ravenna			2	31	12,230	207	58	268
Emilia-Romagna	Parma			3	38	10,211	159	54	185
Emilia-Romagna	Ferrara			3	25	8,754	108	40	199
Emilia-Romagna	Forlì/Cesena			0	30	8,196	127	46	245
Emilia-Romagna	Piacenza			1	13	5,510	99	23	130
Emilia-Romagna	Udine			1	36	8,169	71	26	207
Friuli-Venezia Giulia	Trieste			0	22	4,642	61	16	134
Friuli-Venezia Giulia	Pordenone			2	29	5,291	37	21	58
Friuli-Venezia Giulia	Gorizia			2	14	2,068	26	11	116
Lazio	Roma			40	325	144,511	4,271	381	3,032
Lazio	Latina			6	52	10,660	284	269	401
Lazio	Frosinone			1	25	5,206	150	152	196
Lazio	Viterbo			3	17	4,978	68	118	183
Lazio	Rieti			0	8	1,593	22	42	89
Liguria	Genova			17	84	28,536	635	162	912
Liguria	Savona			4	22	7,029	134	69	202
Liguria	Imperia			1	18	5,768	54	141	254
Liguria	La Spezia			1	12	5,128	98	84	190
Lombardia	Milano			39	502	168,057	5,508	532	2,567
Lombardia	Brescia			8	126	31,451	777	131	883
Lombardia	Bergamo			8	96	23,716	515	148	586
Lombardia	Varese			6	61	16,154	318	109	620

Then set the window like below, click “OK”.

Custom AutoFilter

Show rows where:

Total crimes

is greater than 4,946

And Or

is less than 10,121

Use ? to represent any single character
Use * to represent any series of characters

OK Cancel

The result is below:

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rape	Thefts	Robberies	Arson	Drug
Abruzzo	L'Aquila	309,820	8,944	0	17	3,334	60	49	170
Basilicata	Potenza	383,791	9,303	2	29	2,567	39	72	219
Calabria	Vibo Valentia	166,560	6,466	10	12	1,882	76	78	75
Calabria	Crotone	174,605	5,777	11	11	1,625	28	219	86
Campania	Benevento	287,874	7,294	4	18	2,583	74	218	118
Friuli-Venezia Giulia	Pordenone	315,323	9,709	2	29	5,291	37	21	58
Lombardia	Lodi	227,655	7,374	3	17	4,132	101	19	46
Lombardia	Sondrio	183,169	5,296	0	13	1,560	20	28	74
Molise	Campobasso	231,086	6,414	1	10	2,755	17	83	62
Piemonte	Asti	221,687	9,116	0	16	3,928	67	14	66
Piemonte	Biella	185,768	7,236	1	14	2,503	46	33	108
Piemonte	Vercelli	179,562	6,428	1	9	2,740	26	40	98
Piemonte	Verbano-Cusio-Oss	163,247	5,504	0	13	1,766	24	16	159
Sardegna	Nuoro	160,677	7,753	20	6	2,893	73	39	61
Toscana	Grosseto	228,157	9,940	5	15	4,841	36	42	185
Toscana	Massa-Carrara	203,901	8,836	1	18	4,537	67	9	154
Toscana	Siena	272,638	8,730	0	27	3,749	39	54	84
Umbria	Terni	234,665	8,501	1	17	3,778	78	52	118
Veneto	Rovigo	247,884	8,735	0	27	4,165	49	38	68
Veneto	Belluno	213,474	5,857	1	11	2,087	8	22	61

Separating Columns

Step2.2: Let's separate a whole name into 2 columns: First name and Last name. Go to sheet

"SeparatingColumns", choose column "PLAYER", go to DATA->Text to Columns, choose "Delimited", then "Next".

1. 2. 3. 4. 5.

PLAYER	SALARY	POSITION	Team
Alex Rodriguez	\$ 30,000,000	Third Baseman	Yankees
Mark Teixeira	\$ 23,125,000	First Baseman	Yankees
CC Sabathia	\$ 23,000,000	Pitcher	Yankees
Derek Jeter	\$ 15,729,364	Shortstop	Yankees
Mariano Rivera	\$ 14,940,025	Pitcher	Yankees
Robinson Cano	\$ 14,000,000	Second Baseman	Yankees
Rafael Soriano	\$ 11,000,000	Pitcher	
Nick Swisher	\$ 10,250,000	Outfielder	
Curtis Granderson	\$ 10,000,000	Outfielder	
Hiroki Kuroda	\$ 10,000,000	Pitcher	
Russell Martin	\$ 7,500,000	Catcher	
Freddy Garcia	\$ 4,000,000	Pitcher	
Pedro Feliciano	\$ 3,750,000	Pitcher	
Phil Hughes	\$ 3,200,000	Pitcher	
Brett Gardner	\$ 2,800,000	Outfielder	
Andrew Jones	\$ 2,000,000	Outfielder	
Boone Logan	\$ 1,875,000	Pitcher	
Joba Chamberlain	\$ 1,675,000	Pitcher	
David Robertson	\$ 1,600,000	Pitcher	
Raul Ibanez	\$ 1,100,000	Designated Hitter	
Eric Chavez	\$ 900,000	Infielder	
Michael Pineda	\$ 528,475	Pitcher	
Ivan Nova	\$ 527,200	Pitcher	
Clayton Kershaw	\$ 525,000	Pitcher	
Eduardo Nunez	\$ 523,800	Shortstop	
Cory Wade	\$ 508,925	Pitcher	Yankees
David Aardsma	\$ 500,000	Pitcher	Yankees
Chris Stewart	\$ 482,500	Catcher	Yankees

Choose "Space", then click "Next", in order not to replace other data in the sheet, change Destination to "E1", then click "Finish". You will see the result like below.

1. 2.

Delimiters

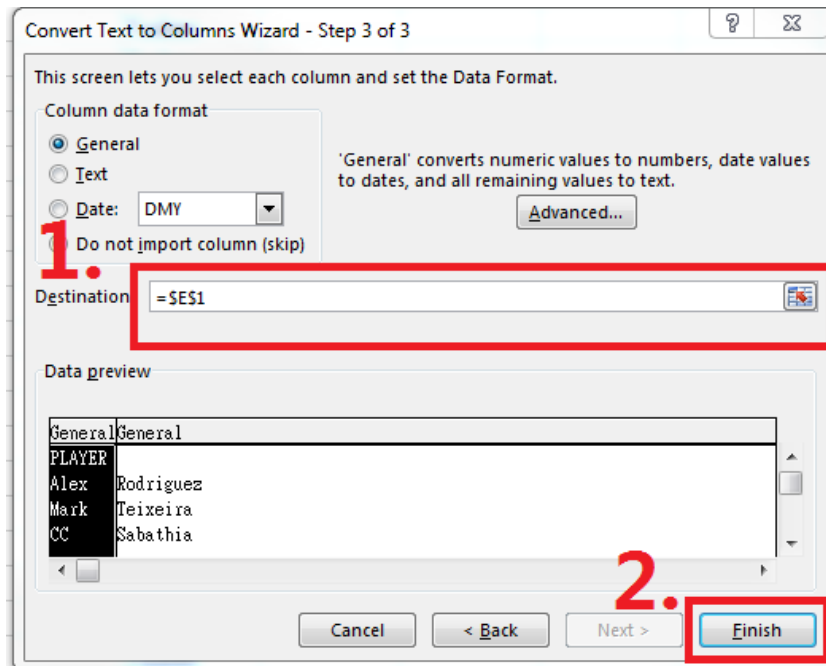
- ☐ Tab
- ☐ Semicolon
- ☐ Comma
- ☒ Space
- ☐ Other:

☒ Treat consecutive delimiters as one

Text qualifier: *

Data preview

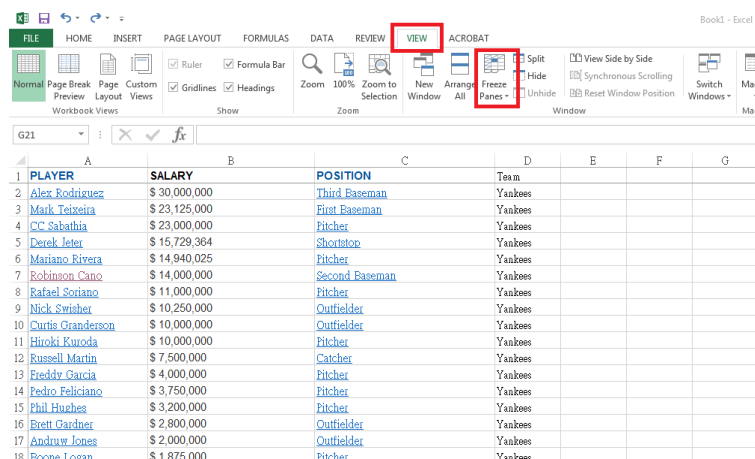
PLAYER
Alex Rodriguez
Mark Teixeira
CC Sabathia



PLAYER	SALARY	POSITION	Team	PLAYER	
Alex Rodriguez	\$ 30,000,000	Third Baseman	Yankees	Alex	Rodriguez
Mark Teixeira	\$ 23,125,000	First Baseman	Yankees	Mark	Teixeira
CC Sabathia	\$ 23,000,000	Pitcher	Yankees	CC	Sabathia
Derek Jeter	\$ 15,729,364	Shortstop	Yankees	Derek	Jeter
Mariano Rivera	\$ 14,940,025	Pitcher	Yankees	Mariano	Rivera
Robinson Cano	\$ 14,000,000	Second Baseman	Yankees	Robinson	Cano
Rafael Soriano	\$ 11,000,000	Pitcher	Yankees	Rafael	Soriano
Nick Swisher	\$ 10,250,000	Outfielder	Yankees	Nick	Swisher
Curtis Granderson	\$ 10,000,000	Outfielder	Yankees	Curtis	Granderson
Hiroki Kuroda	\$ 10,000,000	Pitcher	Yankees	Hiroki	Kuroda
Russell Martin	\$ 7,500,000	Catcher	Yankees	Russell	Martin
Freddie Garcia	\$ 4,000,000	Pitcher	Yankees	Freddie	Garcia
Pedro Feliciano	\$ 3,750,000	Pitcher	Yankees	Pedro	Feliciano
Phil Hughes	\$ 3,200,000	Pitcher	Yankees	Phil	Hughes
Brett Gardner	\$ 2,800,000	Outfielder	Yankees	Brett	Gardner
Andruw Jones	\$ 2,000,000	Outfielder	Yankees	Andruw	Jones
Boone Logan	\$ 1,875,000	Pitcher	Yankees	Boone	Logan
Joba Chamberlain	\$ 1,675,000	Pitcher	Yankees	Joba	Chamberlain
David Robertson	\$ 1,600,000	Pitcher	Yankees	David	Robertson
Raul Ibanez	\$ 1,100,000	Designated Hitter	Yankees	Raul	Ibanez

Freeze Panes:

Step 2.3: Still in sheet "SeparatingColumns", go to "VIEW", then click "Freeze Panes", then you can choose whatever row you would like to fix when you scroll down or up.



Formulas

Add & Concatenate

We all know that Excel can do some easy calculations within cells. For example, the “=A1+B1” formula represents that it add the numbers in A1 and B1 cell together. But what if we have different data type in the cells? **Data type:** Cells can hold text (also known as strings), percentages, dates, time durations, currency and more. Imagine if we changed the cell type of A1 and B1 from “number” to “text”, and now we need to change the formula to “=A1&B1”, if A1=8, and B1=4, instead of getting the result of “12”, we will get “84”. That’s because Excel is now “**concatenating**” the cells.

Go to sheet “**Concatenate**”, let’s concatenate the code value and definition together.

Step 2.4: Go to column “Account Payable Call Center Number”, you will find that the number is separated into 3 columns, so insert a new column beside it. Make sure these 3 columns are “text”, then click “P3” cell, input: =M3&“-”&N3&“-”&O3

Then you will get the result like this:

Accounts Payable Call Center Number			Accounts Payable Call Center Number
8620	8518	6600	8620-8518-6600
8620	8518	6600	8620-8518-6600
8620	8518	6600	8620-8518-6600
8620	8518	6600	8620-8518-6600

“\$” sign: Relative & Absolute Cell References

The **cell reference** is the cell’s name when we refer to it in the formula. When we use the fill handle to copy the formula to the new row of the table, the cell reference will plus 1 normally. This is called the **relative cell reference**. When we add “\$” before any number or letter, the formula will keep the number or letter locked, this is called **absolute cell reference**.

Step 2.5: Go to sheet “ItalyCrimes” again, insert a new column “Abruzzo Total Crimes” on the right of the table, in cell K2, input: =SUM(D2: D5) to calculate the total crimes of Territorio Abruzzo.

Insert another column named “Crime Rate by Province” to compare the crime rate in different provinces in Abruzzo, so in cell L2, input: =D2/K2, click “Enter”, and double click when the mouse hover on the right bottom corner of cell L2. You will find errors like below:

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rapes	Thefts	Robberies	Arsons	Drugs	Abruzzo Total Crimes	Crime Rate by Province
Abruzzo	Pescara	323,184	17635.00	1	33	8,256	202	55	217	51805.00	#DIV/0!
Abruzzo	Chieti	397,123	12861.00	0	15	5,667	101	98	117		#DIV/0!
Abruzzo	Teramo	312,239	12365.00	2	25	5,740	108	52	237		#DIV/0!
Abruzzo	L'Aquila	309,820	8944.00	0	17	3,334	60	49	170		#DIV/0!
Basilicata	Potenza	383,791	9303.00	2	29	2,567	39	72	219		#DIV/0!
Basilicata	Matera	203,726	4272.00	0	12	1,486	19	16	63		#DIV/0!
Calabria	Cosenza	734,656	25163.00	7	51	9,867	254	834	302		#DIV/0!
Calabria	Reggio di Calabria	566,977	19919.00	34	33	7,766	331	448	275		#DIV/0!
Calabria	Catanzaro	368,597	14917.00	14	24	5,834	71	374	210		#DIV/0!
Calabria	Vibo Valentia	166,560	6466.00	10	12	1,882	76	78	75		#DIV/0!
Calabria	Crotone	174,605	5777.00	11	11	1,625	28	219	86		#DIV/0!

That’s because, in every cell, the number of the cell adds one by default, to make K2 cell fixed, we need to add “\$” sign before both K and 2. So in cell L2, input: =D2/\$K\$2, and double click on the right bottom corner, you will see the right answer. For this case, only the first 5 rows are useful.

Territorio	Province	TOTAL POPULATION 2010	Total crimes	Homicides	Rapes	Thefts	Robberies	Arsons	Drugs	Abruzzo Total Crimes	Crime Rate by Province
Abruzzo	Pescara	323,184	17635.00	1	33	8,256	202	55	217	51805.00	=D2/\$K\$2
Abruzzo	Chieti	397,123	12861.00	0	15	5,667	101	98	117		25%
Abruzzo	Teramo	312,239	12365.00	2	25	5,740	108	52	237		24%
Abruzzo	L'Aquila	309,820	8944.00	0	17	3,334	60	49	170		17%

Common Formulas

You can add, subtract, multiply or divide by using the symbols +, -, *, /

Average – Calculate the arithmetic mean of a column or row of numbers

Median – Finds the middle value of a column or row of numbers

Count – Tells you how many items there are in a column or row

Max – Tells you the largest value in a column or row

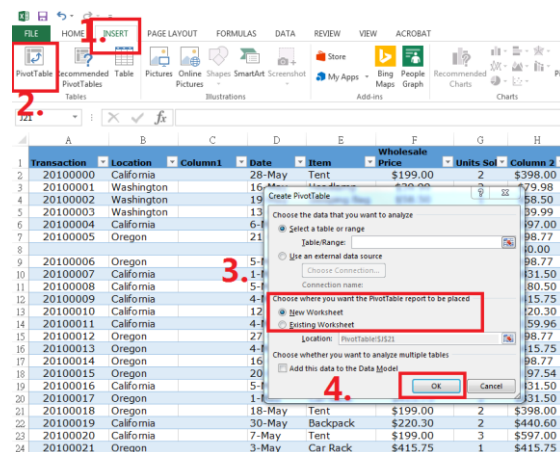
Min – Tells you the smallest value in a column or row

For more information, you can refer to official document about Excel: <https://goo.gl/RT1USJ>

PART 3 Pivot Tables

A pivot table is a powerful function in Excel that helps summarize data in categories. It helps you analyze data in a much clearer way. To create a pivot table, every column of your data must have a variable label.

Step 3.1: Go to sheet “PivotTable”, INSERT->Pivot Table. Choose “New Worksheet”, and then click “OK”. (Remember to choose the whole table before you INSERT)



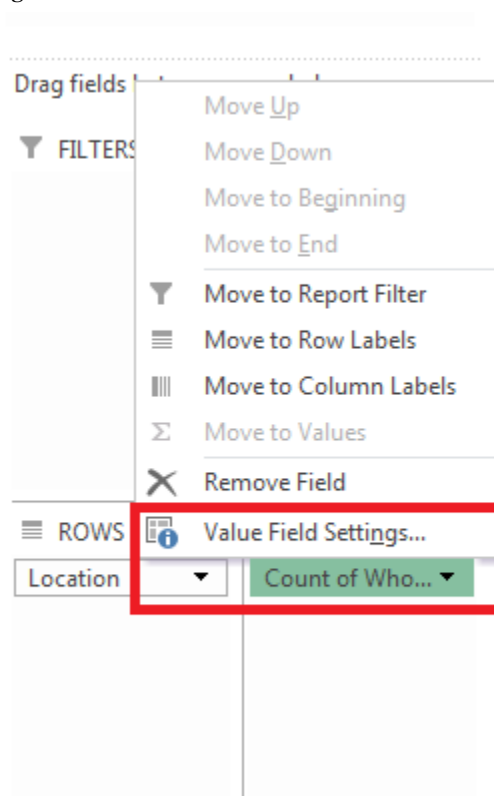
Notice: when your cursor is clicked in the pivot table, the wizard “PivotTable Fields” on the right side would appear, or it would disappear.

To better analyze the data, we first figure out the questions below:

1. Which state has the most number of sale?
 - 1) Drag **Location** to **ROWS**;
 - 2) Drag **WholesalePrice** to **VALUES**;
 - 3) You will find that the count of WholesalePrice below: so Oregon has the most sale number.

Row Labels	Count of WholesalePrice
California	258
Oregon	271
Washington	259
(blank)	
Grand Total	788

- 4) Click small down arrow beside “Count of WholesalePrice” on the right, choose “**Value Field Settings**”, and change “Count” to “**Sum**”. Click “**OK**”.



- 5) You could also see that the highest total amount of sale is also from Oregon.

A	B
Row Labels	Sum of Wholesale
California	55191.78
Oregon	61732.23
Washington	55834.64
(blank)	
Grand Total	172758.65

2. Which item is the most sold?

- 1) Drag **"Item"** to **COLUMNS**;
- 2) You could see the result below:

Sum of Wholesale	Column Labels							
Row Labels	Backpack	Car Rack	Headlamp	Sleeping Bag	Tent	(blank)	Grand Total	
California	11642.73	17680.87	3091.79	8036.39	14740		55191.78	
Oregon	8348.16	28322.97	2199.66	8234.44	14627		61732.23	
Washington	9162.04	21479.95	2589.73	8849.92	13753		55834.64	
(blank)								
Grand Total	29152.93	67483.79	7881.18	25120.75	43120		172758.65	

3. How to see the sale on a specific date?

- 1) Drag **“Date”** to **FILTERS**;
- 2) You will see the **“Date”** row in the pivot table, and click the down arrow beside it;
- 3) Then you will see the specific sale on 1/5/2006.

[illegible]

Tips: If we go back to our original table and change some data, we could refresh our pivot table by go to **“Option”** tab, then click **“Refresh”** to refresh the pivot table.