COMP1022Q
Introduction to Computing with Excel VBA

Sorting in Excel

David Rossiter and Eddie Chan

What is Sorting?

- Sorting helps you to arrange your data in a desired order
- With sorting, you can quickly organize the data that you want
- For example, if you want to know the top 5 students with the highest GPA in your class then you need to sort the GPA column

Sorting in Excel

- An overview of sorting:

- Excel can sort almost anything
- It can sort the data by more than one column or row

What Will You Learn

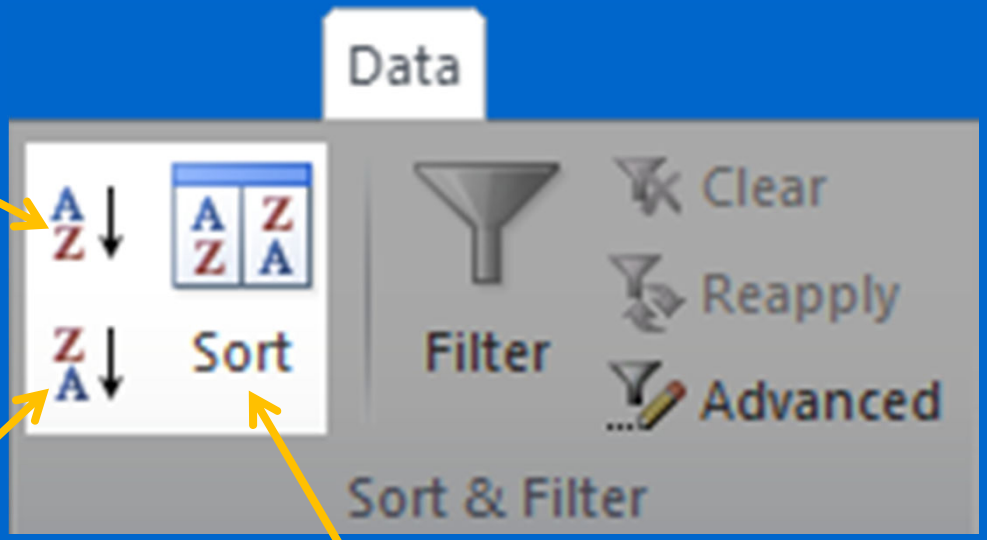
- We will learn the following techniques in this presentation
  - Sort by a column
  - Sort by more than one column
  - Sort one column without affecting the other columns
  - Sort by a row
  - Sort by a custom list

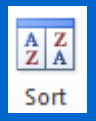
Overview - Sort & Filter Group

- On the Excel Ribbon, select the 'Data' tab; in the 'Sort & Filter' group, you will see these:

- Click  to sort the data in ascending order

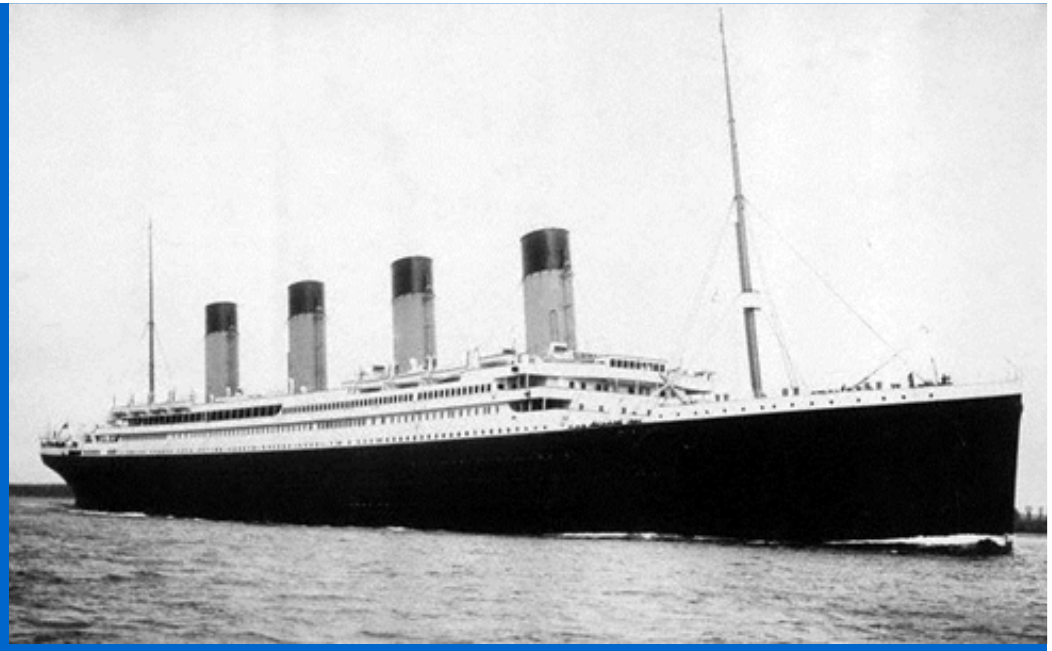
- Click  to sort the data in descending order



- Click  to sort the data using advanced options

The Titanic

- We will use the Titanic's death data as an example to show how to apply sorting



- The Titanic was the largest, most advanced and safest ship ever to sail at that time
- In 1911, Shipbuilder magazine published an article and concluded that the Titanic was *unsinkable*
- The Titanic set sail for New York City on 10 April 1912 with 2,201 people on board - but only 20 lifeboats

The Titanic

- 4 days later she hit an iceberg and sank
- 1,517 people died

- The majority of deaths were caused by hypothermia (freezing to death) in the −2 °C water

The Titanic's Death Data

- Let's look at the Titanic's death data

Economic status

Number of passengers in all classes

Number of deaths in all classes

Percentage of death

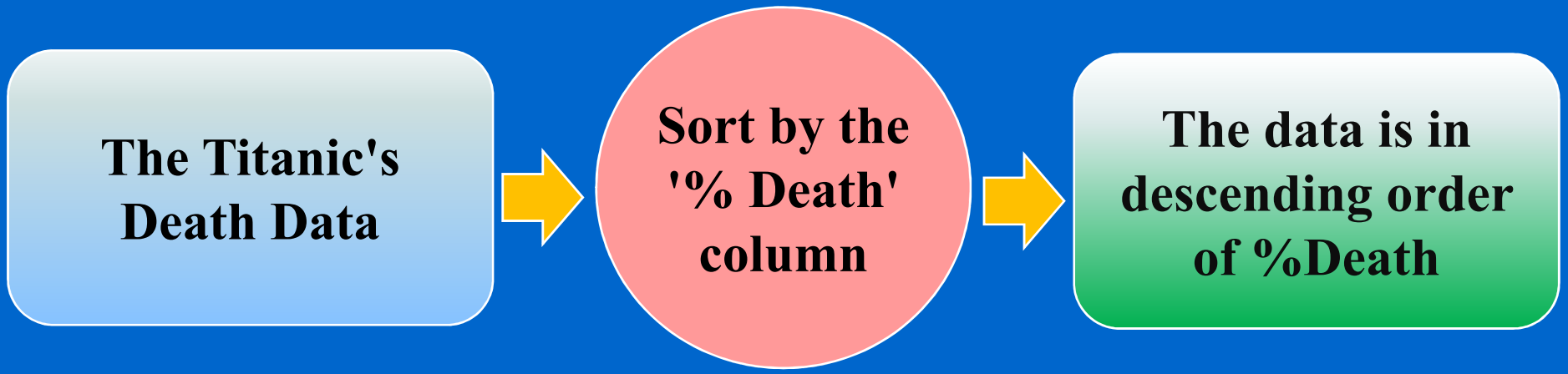
	People on the Ship			Number of Deaths			Percentage of Death
Economic Status	**Male**	**Female**	**Both**	**Male**	**Female**	**Both**	**% Death**
First Class	179	145	324	119	4	123	38.0%
Second Class	178	106	284	154	10	164	57.7%
Third Class	512	196	708	422	96	518	73.2%
Unknown	862	23	885	709	3	712	80.5%
Grand Total	1731	470	2201	1404	113	1517	68.9%

'Unknown' includes the crew and stowaways (= people on the ship without tickets)

These are the column headers

Sort by a Column 1/3

- Here is an example to show how to sort the entire data using the '% Death' column in descending order
- An overview of the sorting in this example:



Sort by a Column 2/3

- Select any cell in the '% Death' column as follows:

	A	B	C	D	E	F	G	H
4		People on the Ship			Number of Deaths			Percentage of Death
5	Economic Status	Male	Female	Both	Male	Female	Both	% Death
6	First Class	179	145	324	119	4	123	38.0%
7	Second Class	178	106	284	154	10	164	57.7%
8	Third Class	512	196	708	422	96	518	73.2%
9	Unknown	862	23	885	709	3	712	80.5%
10								
11	Grand Total	1731	470	2201	1404	113	1517	68.9%

- Then, click  in the 'Data' tab

Sort by a Column 3/3

- Excel will now display rows in order of the percentage that died, with the highest percentage shown at the top:

	People on the Ship			Number of Deaths			Percentage of Death
Economic Status	Male	Female	Both	Male	Female	Both	% Death
Unknown	862	23	885	709	3	712	80.5%
Third Class	512	196	708	422	96	518	73.2%
Second Class	178	106	284	154	10	164	57.7%
First Class	179	145	324	119	4	123	38.0%
Grand Total	1731	470	2201	1404	113	1517	68.9%

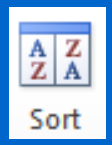
- Excel is smart enough not to mess up the final row (because of the gap above the row)
- You can see that the poorer people had the most deaths

Alternative Method: Column Sort 1/6

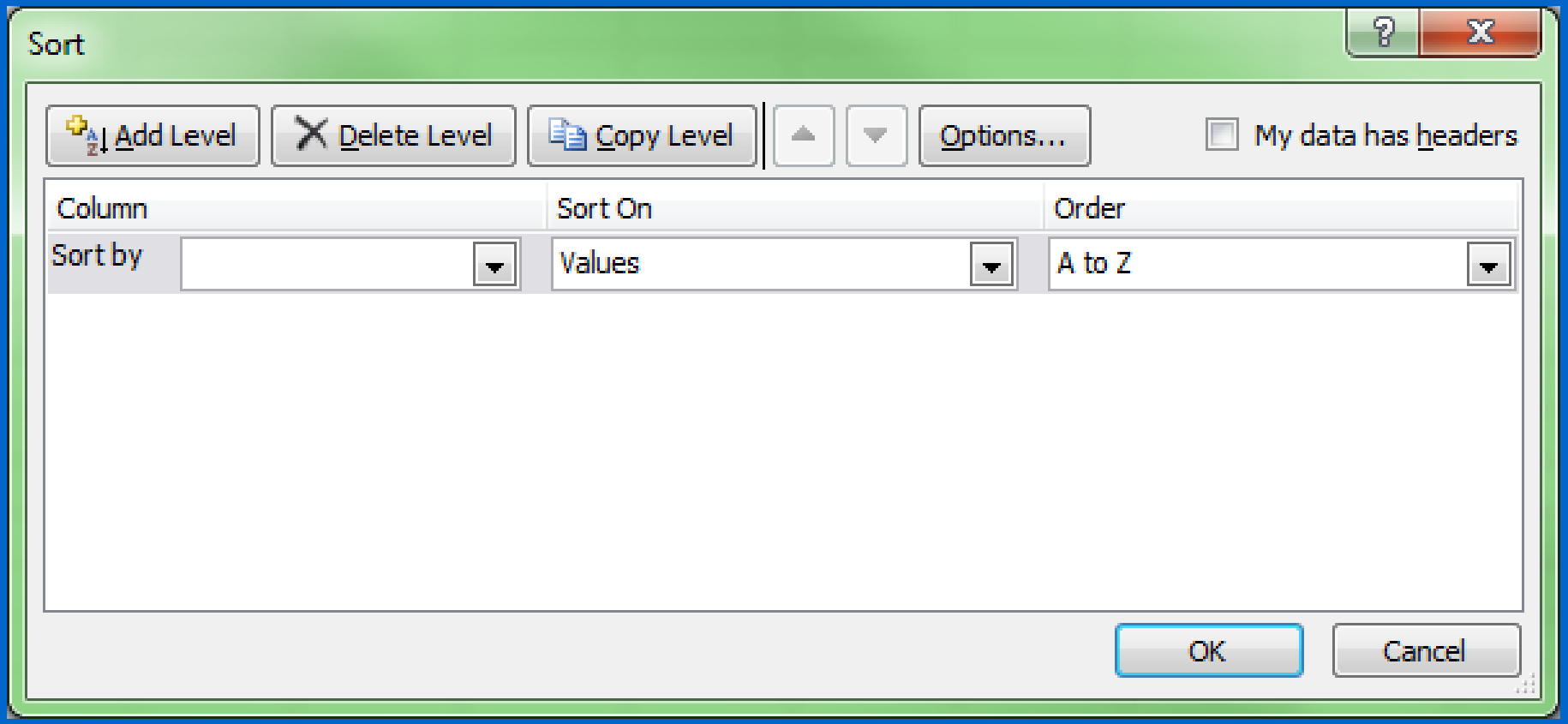
- Now, we will explain another method which does the same thing but gives you much more control
- First, select any cell in the '% Death' column:

	A	B	C	D	E	F	G	H
4		People on the Ship			Number of Deaths			Percentage of Death
5	Economic Status	Male	Female	Both	Male	Female	Both	% Death
6	First Class	179	145	324	119	4	123	38.0%
7	Second Class	178	106	284	154	10	164	57.7%
8	Third Class	512	196	708	422	96	518	73.2%
9	Unknown	862	23	885	709	3	712	80.5%
10								
11	Grand Total	1731	470	2201	1404	113	1517	68.9%

Alternative Method: Column Sort 2/6

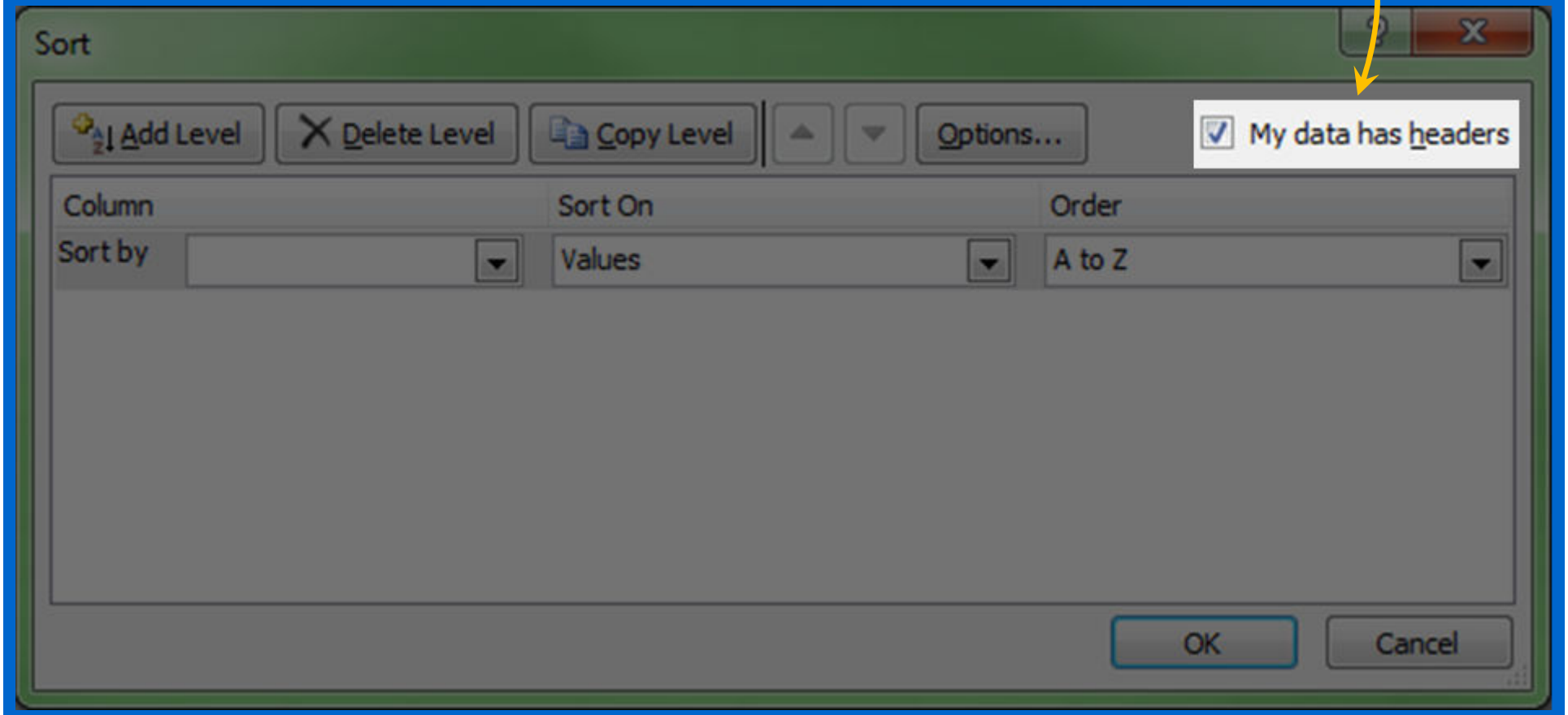
- Click 

to open the 'Sort' window:



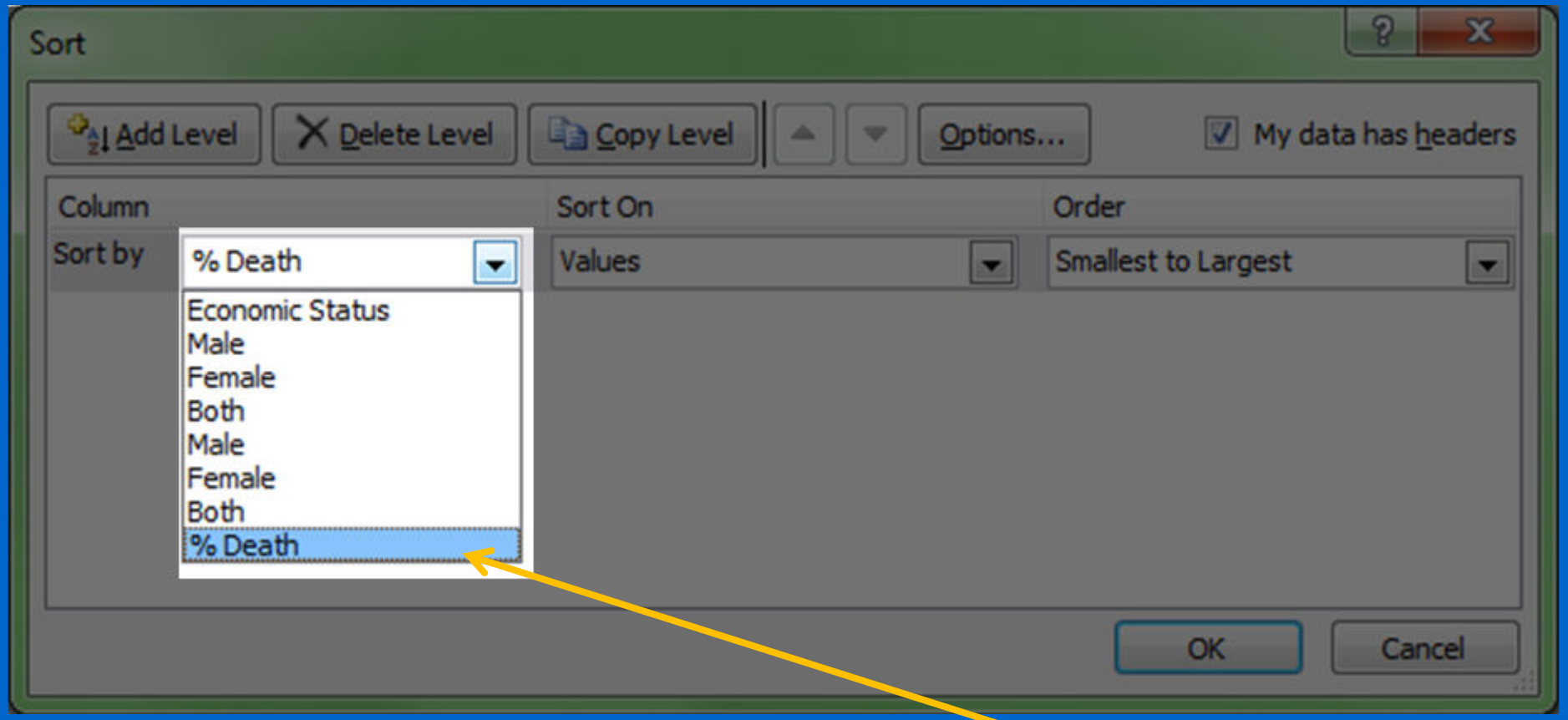
Alternative Method: Column Sort 3/6

- Our data has headers which are useful for sorting, so we select 'My data has headers':



Alternative Method: Column Sort 4/6

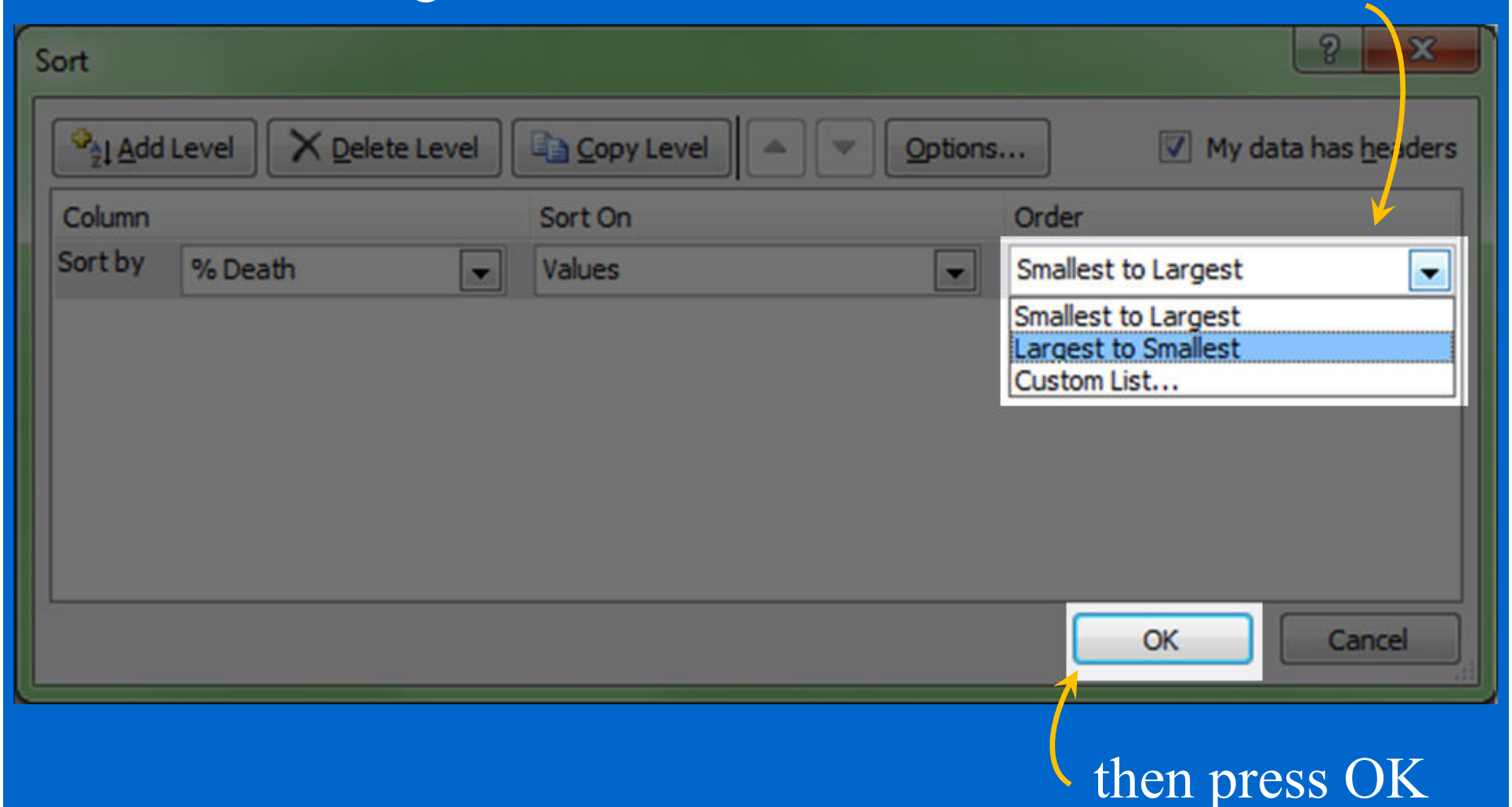
- Select '% Death' in 'Sort by' as shown below:



Economic Status	Population Exposed to Risk			Number of Deaths			Percentage of Death
	Male	Female	Both	Male	Female	Both	% Death
First Class	179	145	324	119	4	123	38.0%

Alternative Method: Column Sort 5/6

- Select Largest to Smallest in 'Order' as shown below:



Alternative Method: Column Sort 6/6

- This will give the same result as before:

	People on the Ship			Number of Deaths			Percentage of Death
Economic Status	Male	Female	Both	Male	Female	Both	% Death
Unknown	862	23	885	709	3	712	80.5%
Third Class	512	196	708	422	96	518	73.2%
Second Class	178	106	284	154	10	164	57.7%
First Class	179	145	324	119	4	123	38.0%
Grand Total	1731	470	2201	1404	113	1517	68.9%

- However, this method gives you much more control, as we will see in later examples

Top 10 Richest People in Asia

- In the following example, we will use a set of data about the top 10 richest people in Asia
- The survey is conducted annually by Forbes.com

Forbes

Top 10 Richest People in Asia Data

- This is what the data looks like

Nationality of the billionaire

Name of the billionaire

The money that the billionaire possesses

Nationality	Name	Net Worth (US Dollar)
India	Mukesh Ambani	29 billion
India	Lakshmi Mittal	28.7 billion
Hong Kong	Li Ka-shing	21 billion
Hong Kong	Lee Shau Kee	18.5 billion
Hong Kong	Kwok Family	17.3 billion
India	Azim Premji	17 billion
Malaysia	Robert Kuok	14.5 billion
India	Anil Ambani	13.7 billion
India	Shashi & Ravi Ruia	13 billion
India	Savitri Jindal	12.2 billion

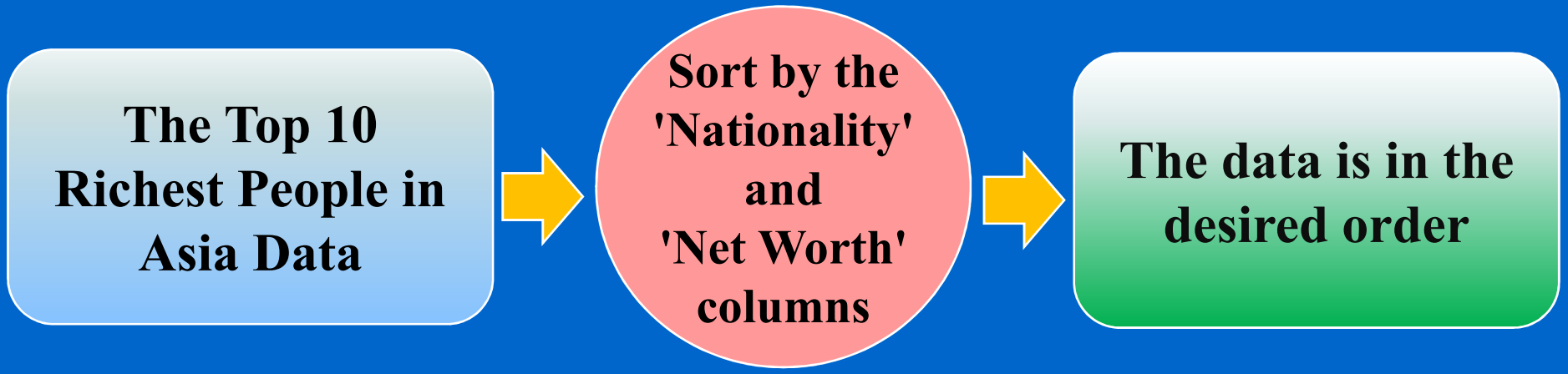
Mukesh Ambani, India

Lee Shau Kee, HK

Before we begin, the data is given to us sorted by the 'Net Worth' column in descending order

Sorting by Two Columns 1/7

- Here is an example to show how to sort the entire data using the 'Nationality' and 'Net Worth' columns
- An overview of the sorting in this example:



Sorting by Two Columns 2/7

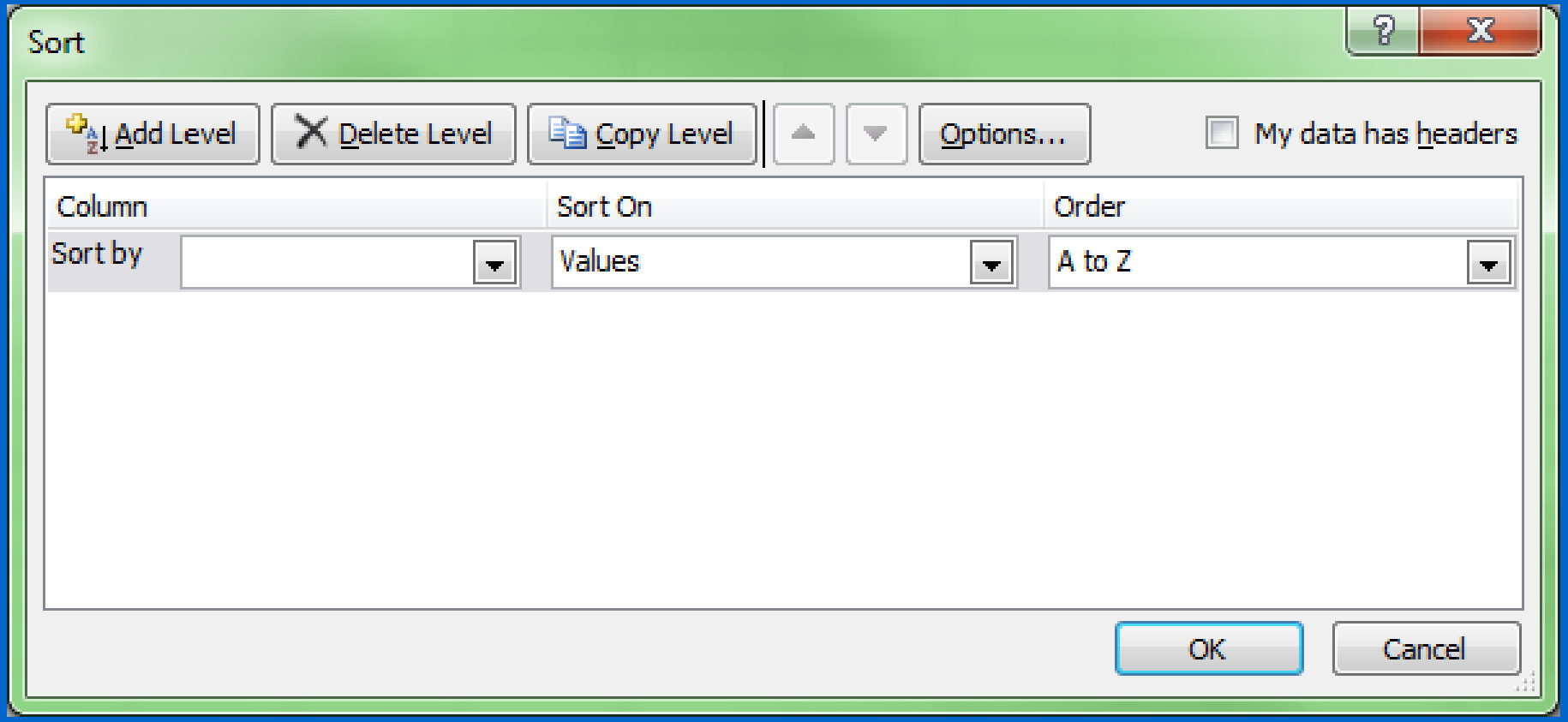
- Select any cell in the data:

	A	B	C
4	Nationality	Name	Net Worth (US Dollar)
5	India	Mukesh Ambani	29 billion
6	India	Lakshmi Mittal	28.7 billion
7	Hong Kong	Li Ka-shing	21 billion
8	Hong Kong	Lee Shau Kee	18.5 billion
9	Hong Kong	Kwok Family	17.3 billion
10	India	Azim Premji	17 billion
11	Malaysia	Robert Kuok	14.5 billion
12	India	Anil Ambani	13.7 billion
13	India	Shashi & Ravi Ruia	13 billion
14	India	Savitri Jindal	12.2 billion

Sorting by Two Columns 3/7

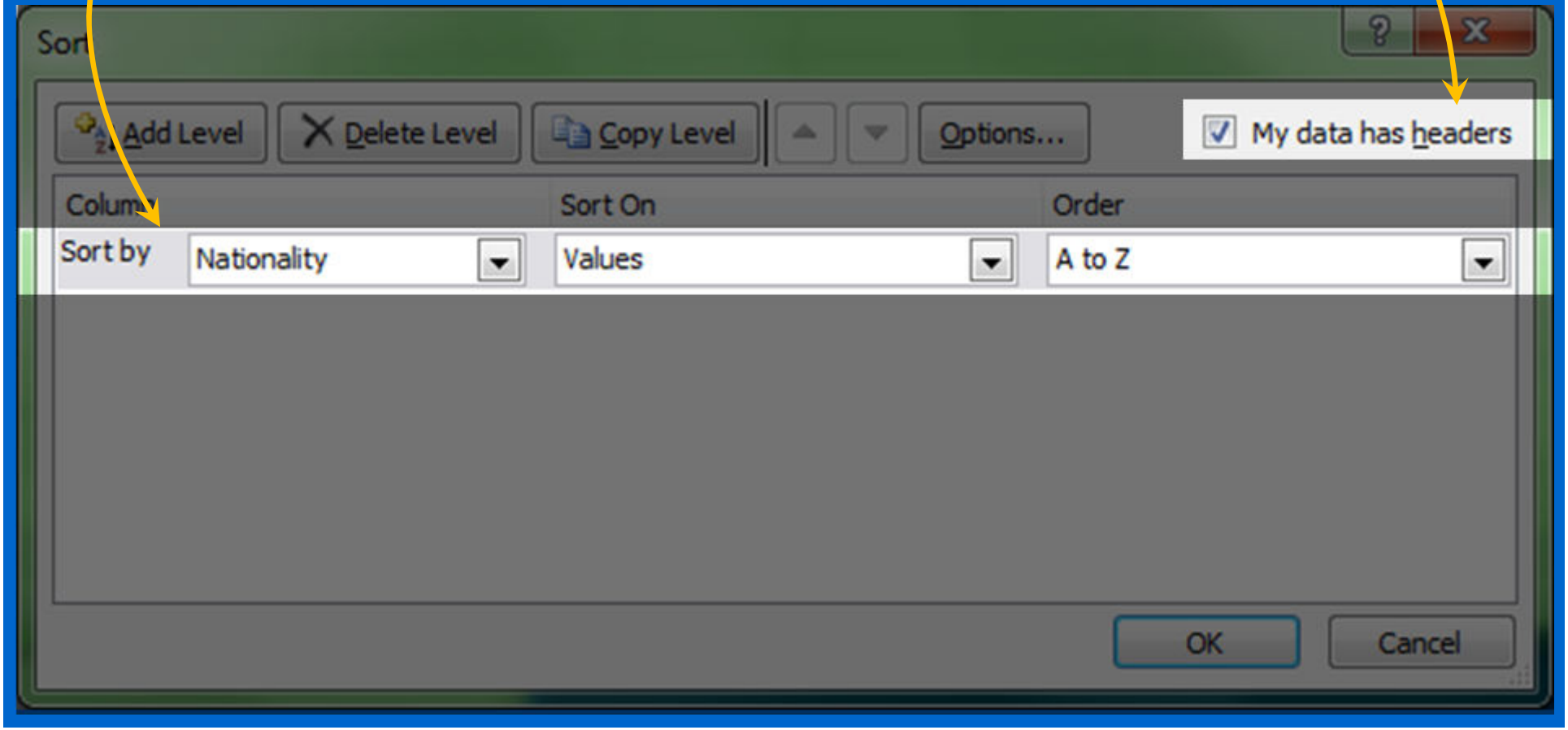
- Click 

to open the 'Sort' window:



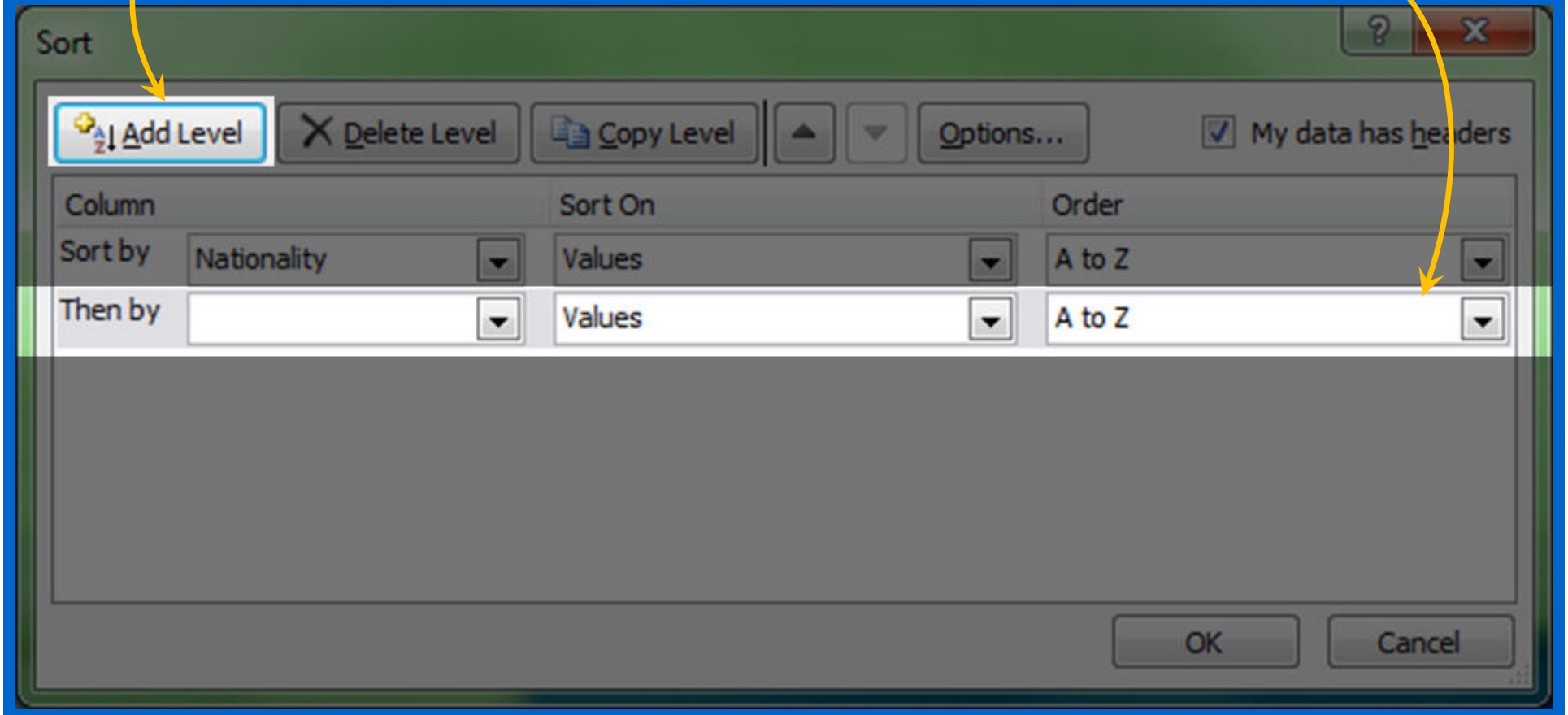
Sorting by Two Columns 4/7

- Like before, select 'My data has headers'
- Sort the 'Nationality' column with the following settings:



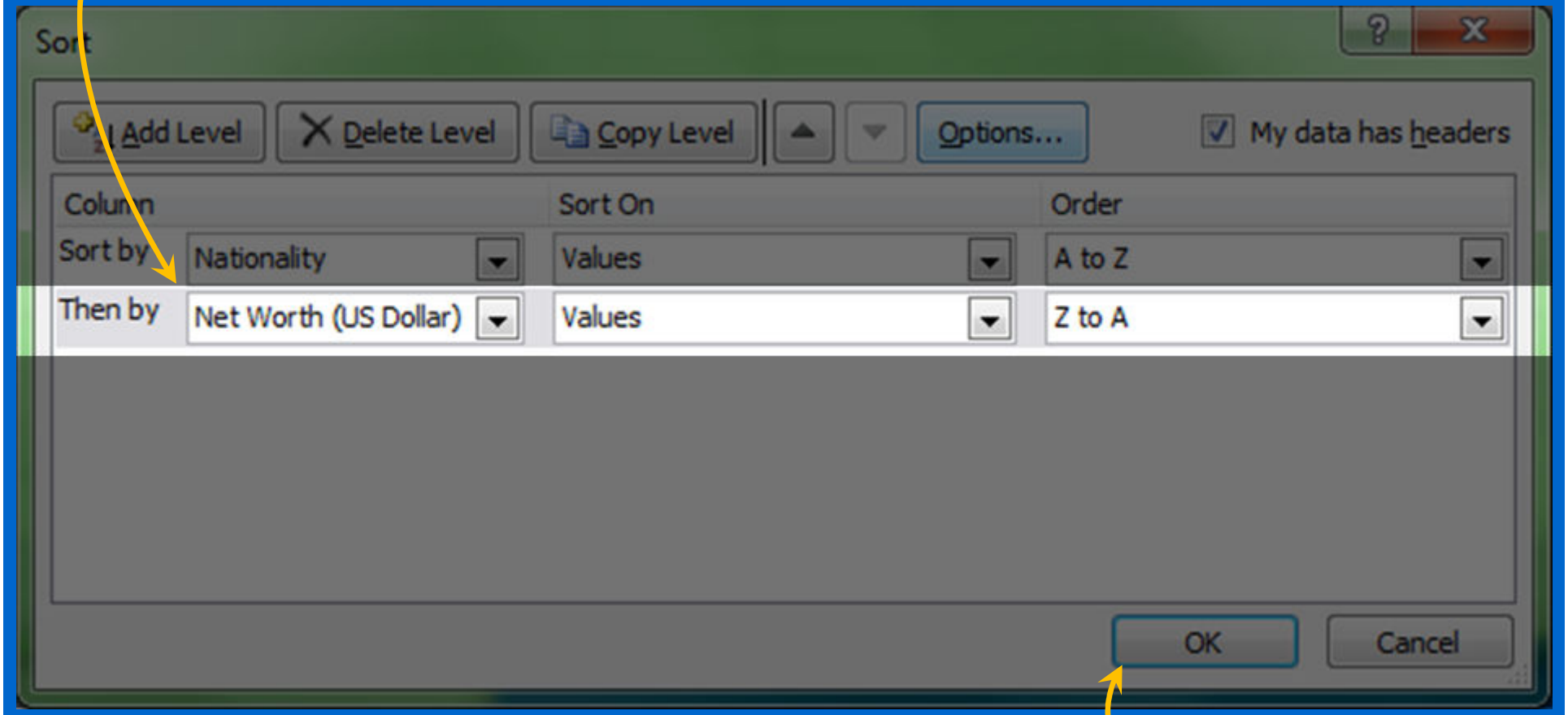
Sorting by Two Columns 5/7

- Click 'Add Level'; it will show a new row:



Sorting by Two Columns 6/7

- Then sort the 'Net Worth (US Dollar)' column with the following settings:



then press OK

Sorting by Two Columns 7/7

- Excel will now sort the data:

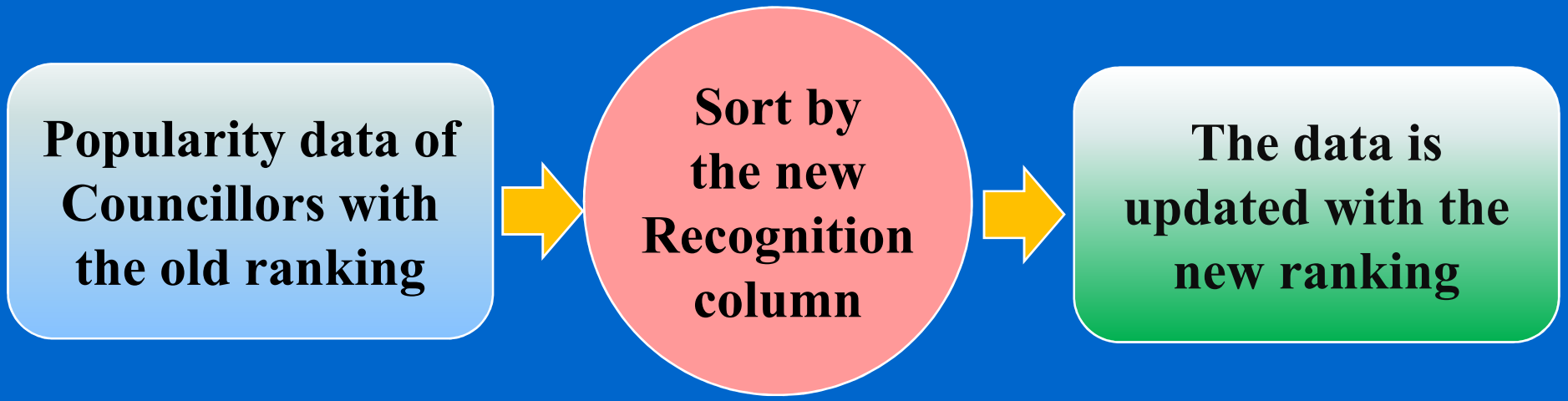
Nationality	Name	Net Worth (US Dollar)
Hong Kong	Li Ka-shing	21 billion
Hong Kong	Lee Shau Kee	18.5 billion
Hong Kong	Kwok Family	17.3 billion
India	Mukesh Ambani	29 billion
India	Lakshmi Mittal	28.7 billion
India	Azim Premji	17 billion
India	Anil Ambani	13.7 billion
India	Shashi & Ravi Ruia	13 billion
India	Savitri Jindal	12.2 billion
Malaysia	Robert Kuok	14.5 billion

Li Ka-shing, HK

- The data has been sorted so that all the people from a particular country are shown together, in descending wealth
- Now, we can easily see that Li Ka-shing is the richest person in Hong Kong

Sorting Some Columns Without Affecting Other Columns

- Here is an example to show how to sort some columns in the order of one column without affecting the first column
- An overview of the sorting in this example:



Hong Kong Legislative Councillors

- In the following example, we will use a set of data about the popularity of Hong Kong Legislative Councillors
- The survey is conducted every three months by the Hong Kong University

香港大學民意網站 HKU POP Site

Recognition of HK Legislative Councillors

- Let's look at the rating data of the top 10 Hong Kong Legislative Councillors in January 2012

Ranking in Jan 2012

How well-known the councillor is

Wong Yuk-man is the most recognized councillor in Jan 2012

Ranking	Chinese Name	English Name	Percentage of Recognition in Jan 2012
1	黃毓民	Wong Yuk-man	10.8%
2	梁國雄	Leung Kwok-hung	8.3%
3	何俊仁	Albert Ho	7.8%
4	劉慧卿	Emily Lau	5.6%
5	余若薇	Audrey Eu	5.6%
6	葉劉淑儀	Regina Ip	5.2%
7	陳偉業	Albert Chan	3.9%
8	陳淑莊	Tanya Chan	3.8%
9	梁家傑	Alan Leong	3.7%
10	李卓人	Lee Cheuk-yan	3.6%

Recognition of HK Legislative Councillors

- We have a new column of data in April 2012
- However, the ranking is now wrong

Ranking is not updated

A new column in April 2012 is added

Ranking	Chinese Name	English Name	Percentage of Recognition in April 2012	Percentage of Recognition in Jan 2012
1	黃毓民	Wong Yuk-man	7.9%	10.8%
2	梁國雄	Leung Kwok-hung	9.2%	8.3%
3	何俊仁	Albert Ho	6.0%	7.8%
4	劉慧卿	Emily Lau	4.2%	5.6%
5	余若薇	Audrey Eu	5.3%	5.6%
6	葉劉淑儀	Regina Ip	4.5%	5.2%
7	陳偉業	Albert Chan	4.2%	3.9%
8	陳淑莊	Tanya Chan	4.1%	3.8%
9	梁家傑	Alan Leong	4.7%	3.7%
10	李卓人	Lee Cheuk-yan	4.5%	3.6%

- So we need to sort the data in the correct ranking order

Updating the Correct Ranking

- To update the correct ranking, we only need to sort the last four columns using 'Percentage of Recognition in April 2012' column, in descending order

Ranking	Chinese Name	English Name	Percentage of Recognition in April 2012	Percentage of Recognition in Jan 2012
1	黃毓民	Wong Yuk-man	7.9%	10.8%
2	梁國雄	Leung Kwok-hung	9.2%	8.3%
3	何俊仁	Albert Ho	6.0%	7.8%
4	劉慧卿	Emily Lau	4.2%	5.6%
5	余若薇	Audrey Eu	5.3%	5.6%
6	葉劉淑儀	Regina Ip	4.5%	5.2%
7	陳偉業	Albert Chan	4.2%	3.9%
8	陳淑莊	Tanya Chan	4.1%	3.8%
9	梁家傑	Alan Leong	4.7%	3.7%
10	李卓人	Lee Cheuk-yan	4.5%	3.6%

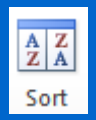
- These columns need to be sorted without affecting the 'Ranking' column

Sorting Some Columns Without Affecting Other Columns 1/4

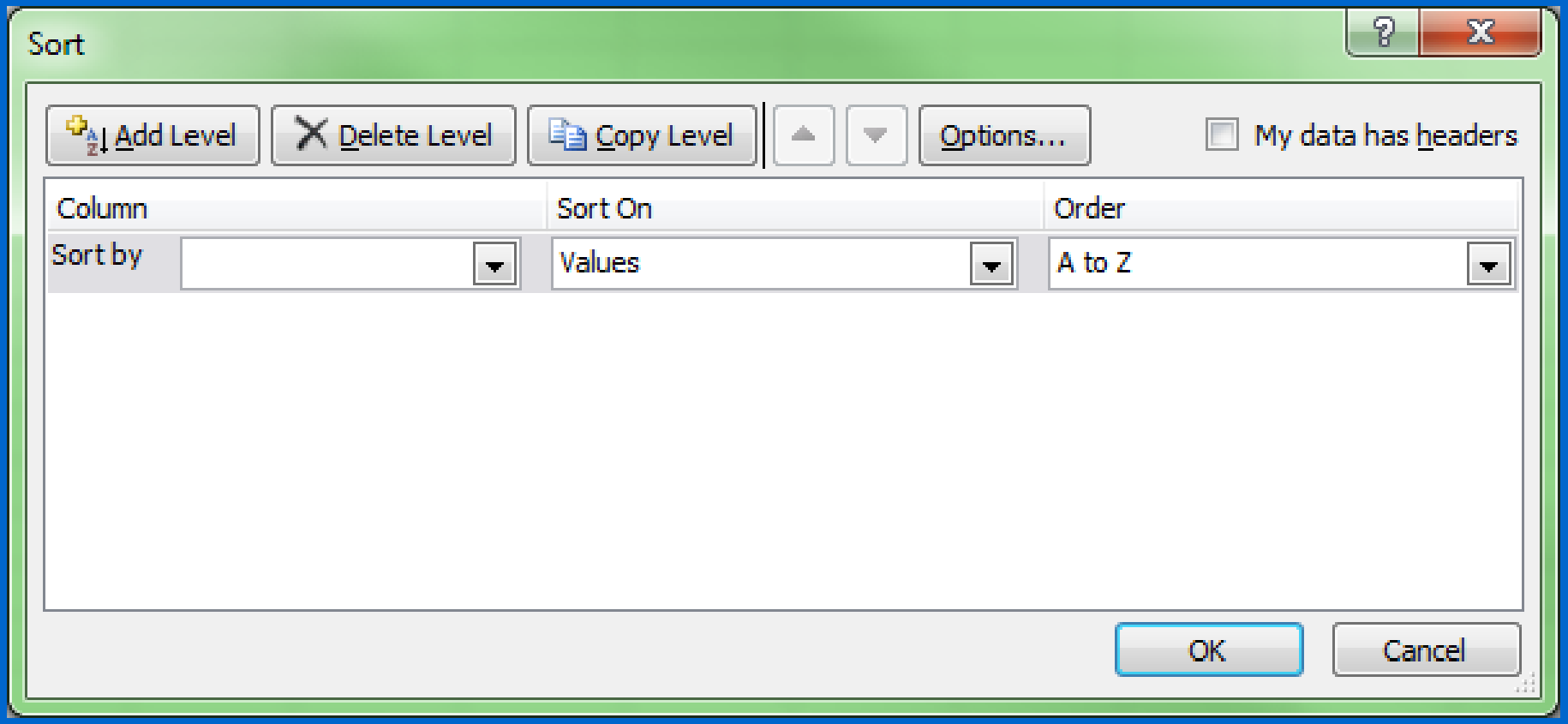
- Select the entire data except the 'Ranking' column

Ranking	Chinese Name	English Name	Percentage of Recognition in April 2012	Percentage of Recognition in Jan 2012
1	黃毓民	Wong Yuk-man	7.9%	10.8%
2	梁國雄	Leung Kwok-hung	9.2%	8.3%
3	何俊仁	Albert Ho	6.0%	7.8%
4	劉慧卿	Emily Lau	4.2%	5.6%
5	余若薇	Audrey Eu	5.3%	5.6%
6	葉劉淑儀	Regina Ip	4.5%	5.2%
7	陳偉業	Albert Chan	4.2%	3.9%
8	陳淑莊	Tanya Chan	4.1%	3.8%
9	梁家傑	Alan Leong	4.7%	3.7%
10	李卓人	Lee Cheuk-yan	4.5%	3.6%

Sorting Some Columns Without Affecting Other Columns 2/4

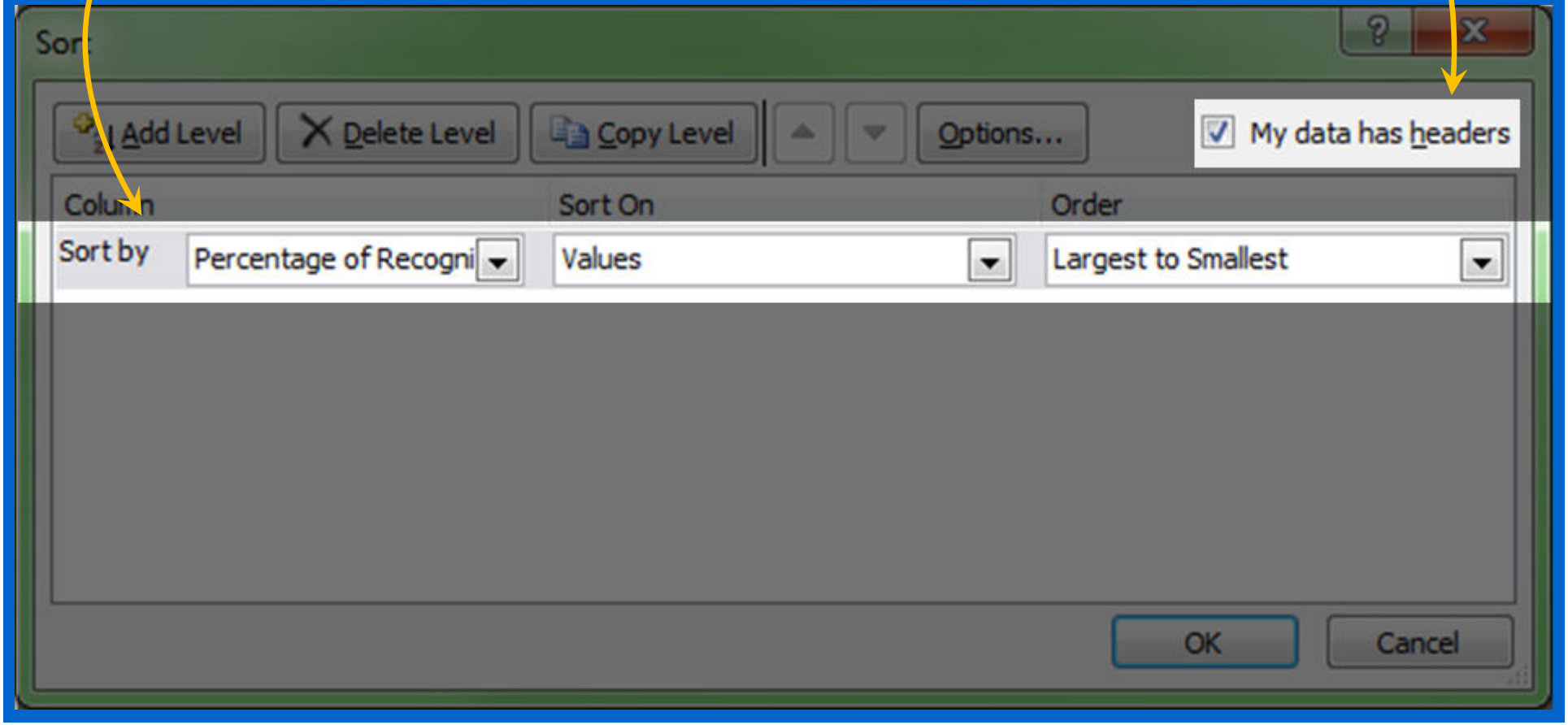
- Click 

to open the 'Sort' window:



Sorting Some Columns Without Affecting Other Columns 3/4

- Click the check box of 'My data has headers'
- Sort the 'Percentage of Recognition in April 2012' column with the following settings and press OK:



Sorting Some Columns Without Affecting Other Columns 4/4

- Excel will now update the new ranking for April 2012:

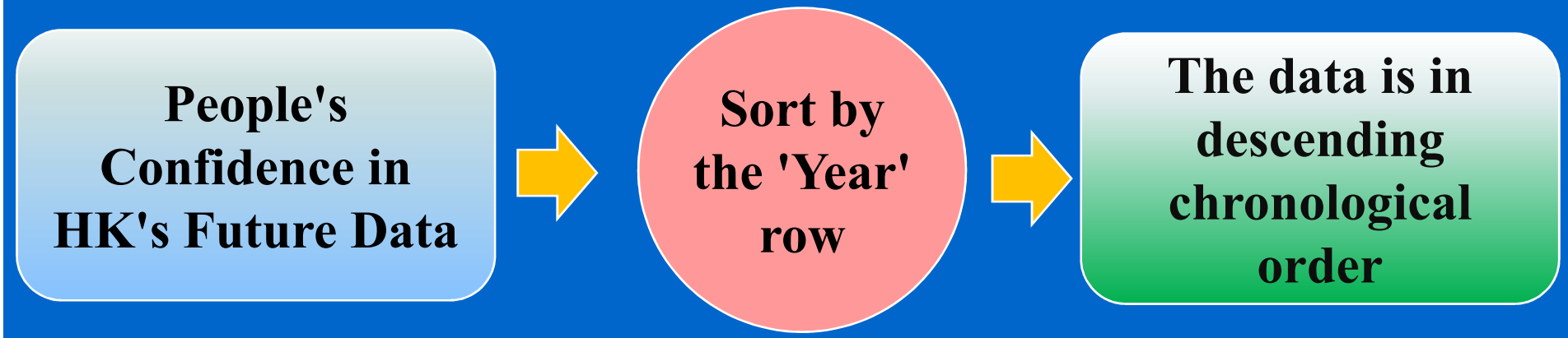
Leung Kwok-hung is now the most recognized councillor

Ranking	Chinese Name	English Name	Percentage of Recognition in April 2012	Percentage of Recognition in Jan 2012
1	梁國雄	Leung Kwok-hung	9.2%	8.3%
2	黃毓民	Wong Yuk-man	7.9%	10.8%
3	何俊仁	Albert Ho	6.0%	7.8%
4	余若薇	Audrey Eu	5.3%	5.6%
5	梁家傑	Alan Leong	4.7%	3.7%
6	葉劉淑儀	Regina Ip	4.5%	5.2%
7	李卓人	Lee Cheuk-yan	4.5%	3.6%
8	劉慧卿	Emily Lau	4.2%	5.6%
9	陳偉業	Albert Chan	4.2%	3.9%
10	陳淑莊	Tanya Chan	4.1%	3.8%

- These columns have been sorted in the correct order, using the fourth column

Sorting by a Row

- All previous examples use ‘column’ structured data
- But sometimes data is structured in rows
- Here is an example of row structured data, which we will sort in descending chronological order
- An overview of the sorting in this example:



- 'Chronological order' means in order of increasing time
- 'Descending chronological order' means in order of decreasing time (the most recent first)

People's Confidence in Future Data

- The data for this example is HK people's confidence in HK's future, for the 1994 to 2011 period
- The HKU survey team phones around 4,000 Hong Kong citizens every year
- They ask the following question to the interviewee, "*Do you have confidence in Hong Kong's future?*"
- The interviewee can answer 3 possible options: *Confident*, *Hard to Say*, and *Not Confident*

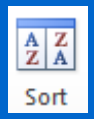
Confidence Level	2003	2002	2001	1994	1995	1996	2004	. . .	2008
Confident	39.9%	51.3%	54.5%	55.6%	58.0%	58.3%	60.0%		78.9%
Hard to Say	12.8%	13.5%	14.7%	23.4%	20.2%	24.0%	14.0%		5.4%
Not Confident	47.3%	35.2%	30.8%	21.0%	21.8%	17.7%	26.0%		15.7%

- When we receive the data, it is not in chronological order

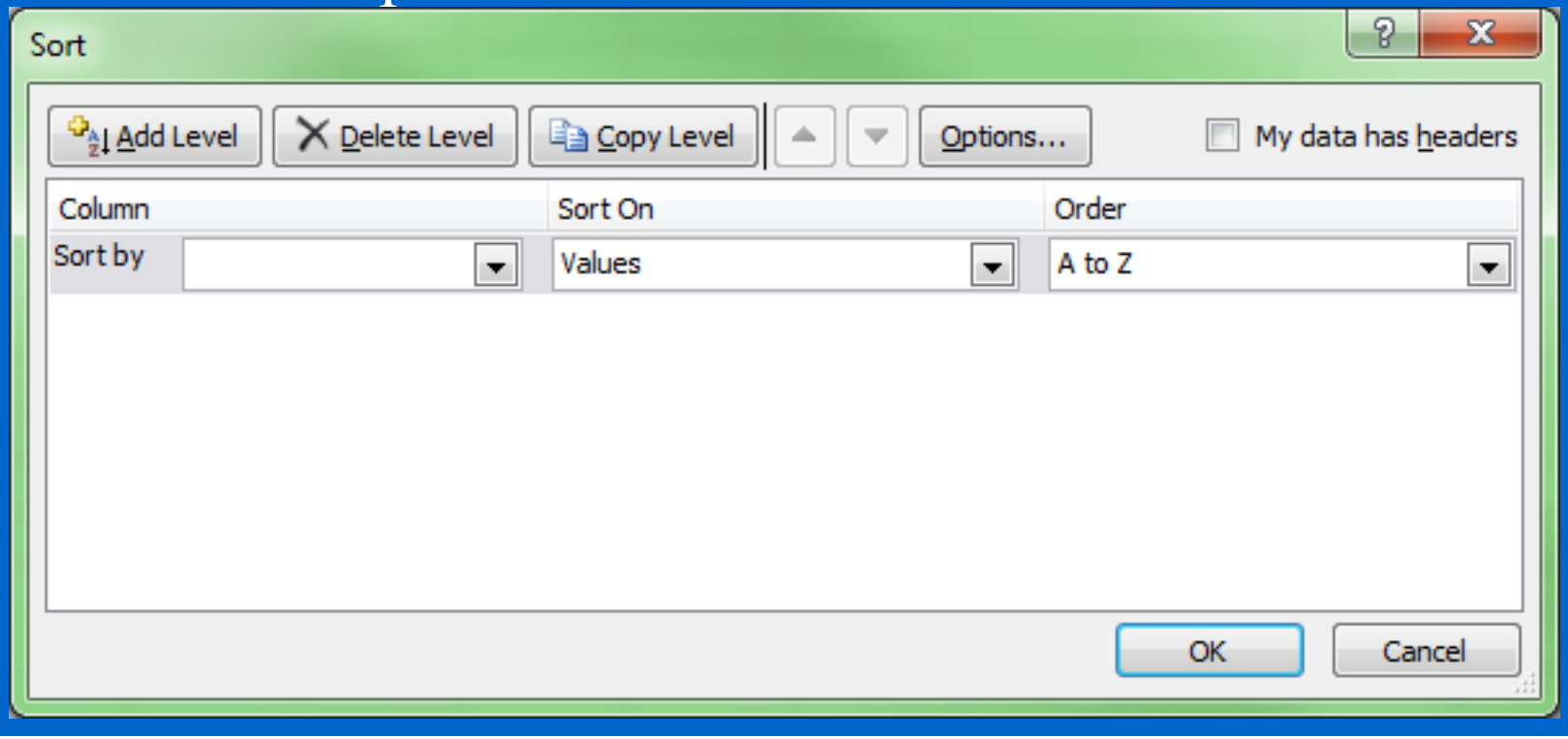
Sorting Row Structured Data 1/5

- Let's re-organise the data in chronological order
- Select the entire table without the row headers:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
4	Confidence Level	2003	2002	2001	1994	1995	1996	2004	1999	2011	1998	2010	2000	2005	2009	1997	2006	2007	2008
5	Confident	39.9%	51.3%	54.5%	55.6%	58.0%	58.3%	60.0%	60.2%	62.8%	63.2%	64.0%	65.2%	69.5%	70.4%	71.5%	77.7%	78.8%	78.9%
6	Hard to Say	12.8%	13.5%	14.7%	23.4%	20.2%	24.0%	14.0%	15.0%	5.5%	17.8%	5.3%	14.1%	13.5%	4.7%	16.9%	6.6%	5.8%	5.4%
7	Not Confident	47.3%	35.2%	30.8%	21.0%	21.8%	17.7%	26.0%	24.8%	31.7%	19.0%	30.7%	20.6%	17.0%	24.9%	11.6%	15.7%	15.4%	15.7%

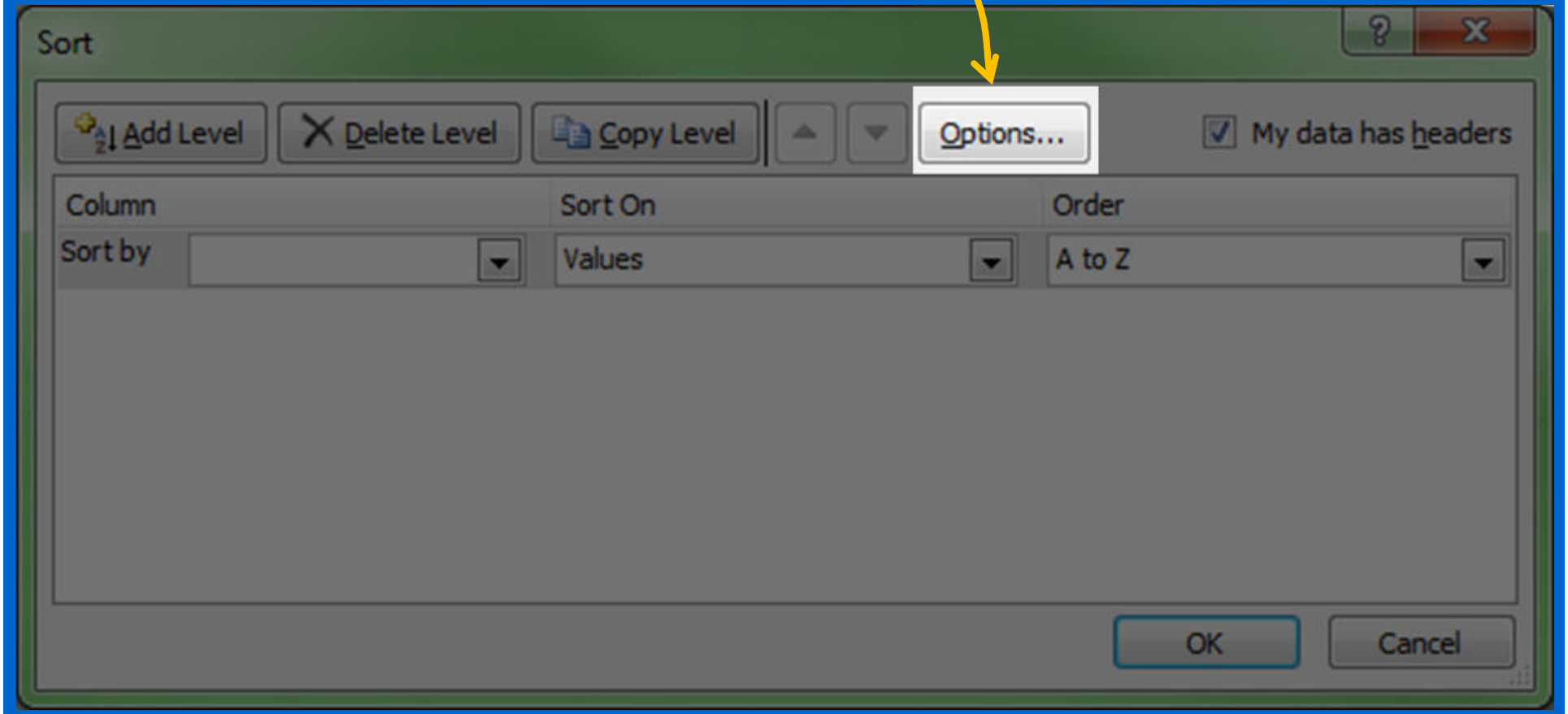
- Click 

to open the 'Sort' window:



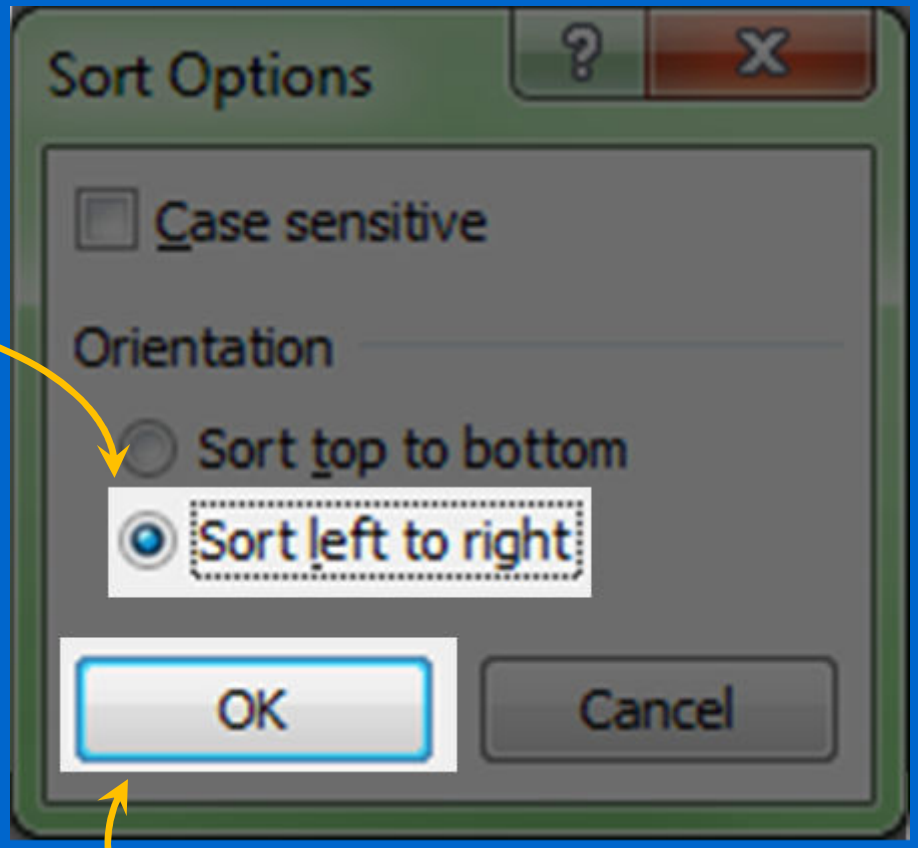
Sorting Row Structured Data 2/5

- Click 'Options…'



Sorting Row Structured Data 3/5

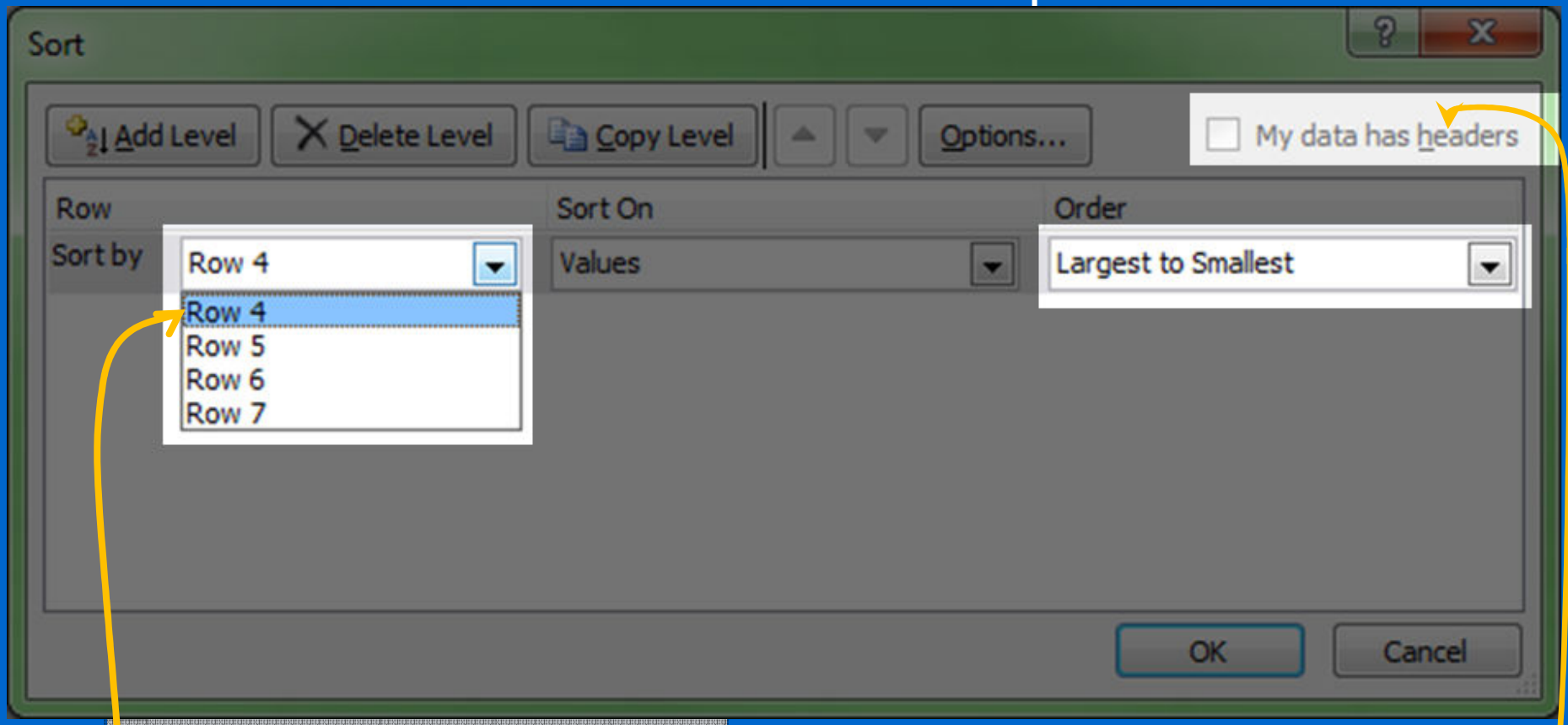
- After clicking 'Options', under 'Orientation', click 'Sort left to right'



then press OK

Sorting Row Structured Data 4/5

- Select Row 4 in 'Sort by' with the Order 'Largest to Smallest' as shown below and press OK:



	A	B	C	D
4	Confidence Level	2003	2002	2011
5	Confident	39.9%	51.3%	62.8%
6	Hard to Say	12.8%	13.5%	5.5%
7	Not Confident	47.3%	35.2%	31.7%

. . .

- Excel is a bit stupid when doing row sorting - it cannot find the row headers, even though you have them

Sorting Row Structured Data 5/5

- Excel will now display records in descending chronological order, as we want:

Confidence Level	2011	2010	2009	2008	2007	2006	2005	...	1994
Confident	62.8%	64.0%	70.4%	78.9%	78.8%	77.7%	69.5%		55.6%
Hard to Say	5.5%	5.3%	4.7%	5.4%	5.8%	6.6%	13.5%		23.4%
Not Confident	31.7%	30.7%	24.9%	15.7%	15.4%	15.7%	17.0%		21.0%

Study about Confidence Data

- After we sort the data in descending chronological order, we can now 'experience' the data better:

A	B	C	D	E	J	K	L	M	P	Q
Confidence Level	**2011**	**2010**	**2009**	**2008**	**2003**	**2002**	**2001**	**2000**	**1997**	**1996**
Confident	62.8%	64.0%	70.4%	78.9%	39.9%	51.3%	54.5%	65.2%	71.5%	58.3%
Hard to Say	5.5%	5.3%	4.7%	5.4%	12.8%	13.5%	14.7%	14.1%	16.9%	24.0%
Not Confident	31.7%	30.7%	24.9%	15.7%	47.3%	35.2%	30.8%	20.6%	11.6%	17.7%

Global economic downturn

The Hang Seng Index falls to 10,000

SARS

300 deaths in Hong Kong

Asia economic downturn

Many IT companies closed

The year before the handover of HK back to China

- We can see moments in time when Hong Kong people became less confident in the future

Custom Sorting

- Let's look at custom sorting
- The data we will use is people's appraisal of the last Chief Executive's job performance for the 2005 to 2011 period
- To obtain this data, a HKU team phone 14,000 people every year and ask them to rate the performance of the last Chief Executive using 5 measures:
  - Very Good
  - Quite Good
  - Half-half
  - Quite Bad
  - Very Bad

People's Appraisal of the Last Chief Executive Data

Total number of interviewees

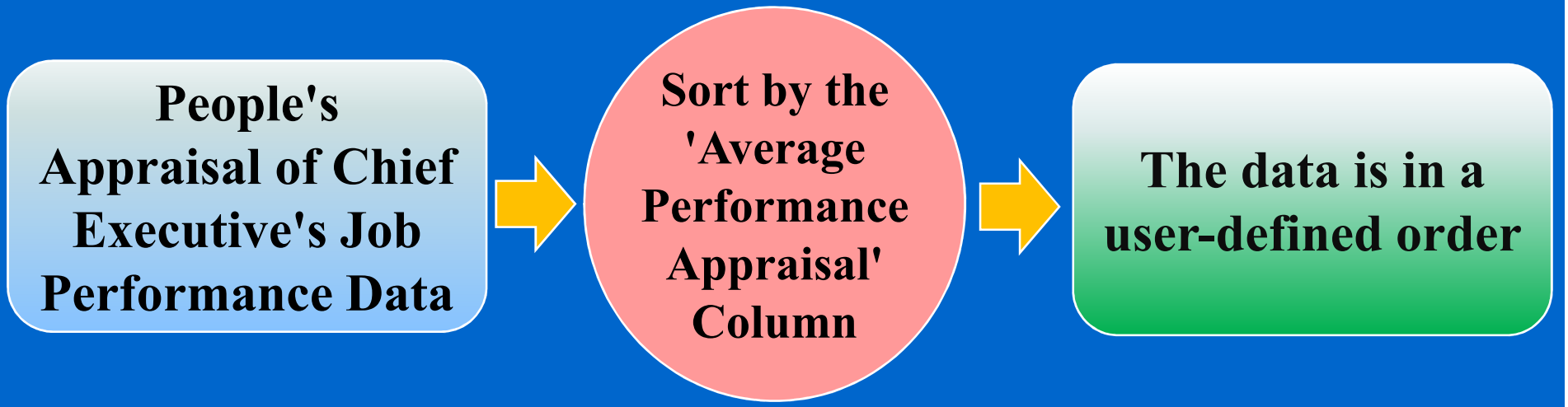
- Here is the data

Year of Survey	Average Performance Appraisal	Total sample
2011	Quite Bad	8176
2010	Half-half	13140
2009	Half-half	13083
2008	Half-half	13276
2007	Quite Good	13177
2006	Half-half	13180
2005	Quite Good	13083

- The data is given to us in descending chronological order

Sorting by a Custom List 1/9

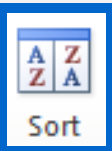
- How can you sort the data so that rows containing Very Good appear first, followed by Quite Good, Half-half, Quite Bad and Very Bad?
- You can use a *custom list* to do this
- An overview of the sorting in this example:



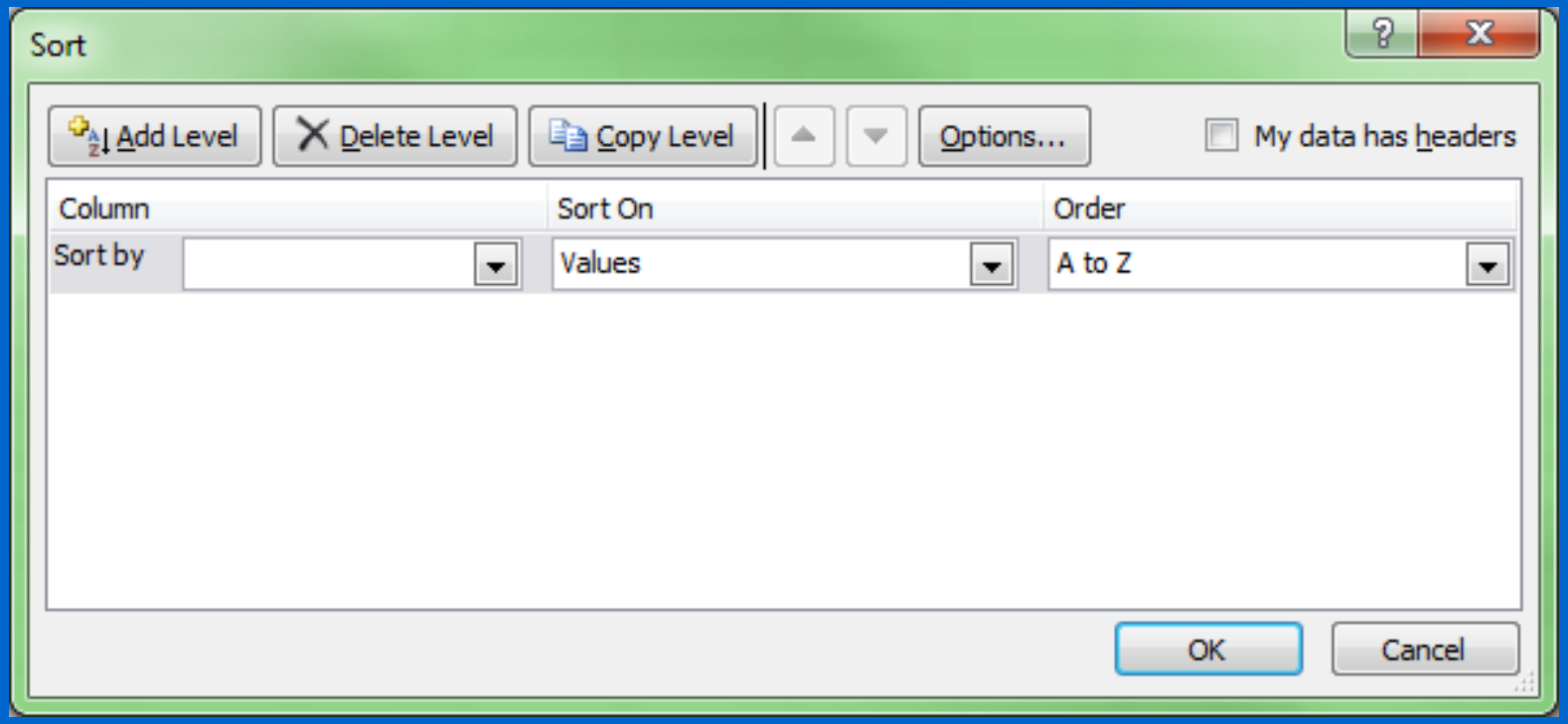
Sorting by a Custom List 2/9

- Select any cell in the data:

	A	B	C
4	Year of Survey	Average Performance Appraisal	Total sample
5	2005	Quite Good	13083
6	2006	Half-half	13180

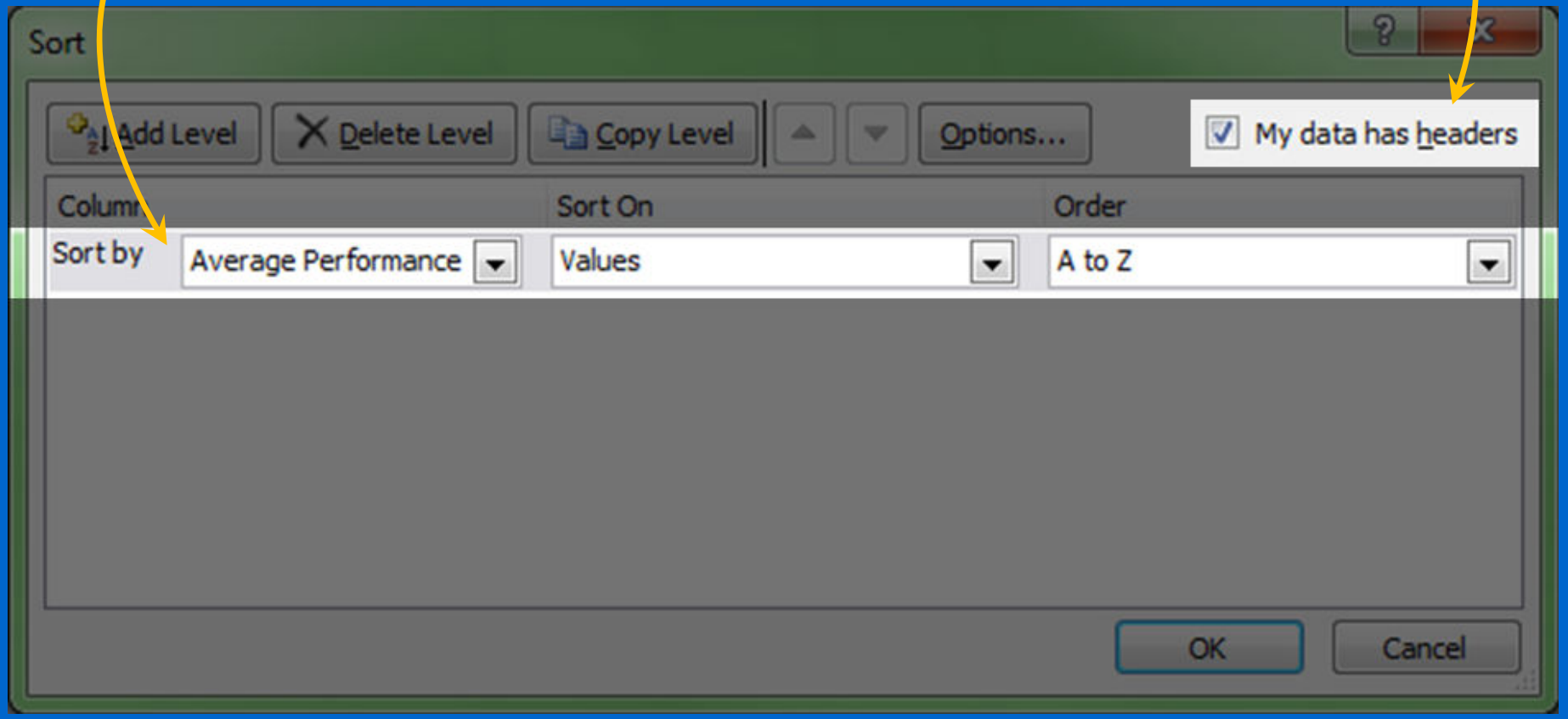
- Click 

to open the 'Sort' window:



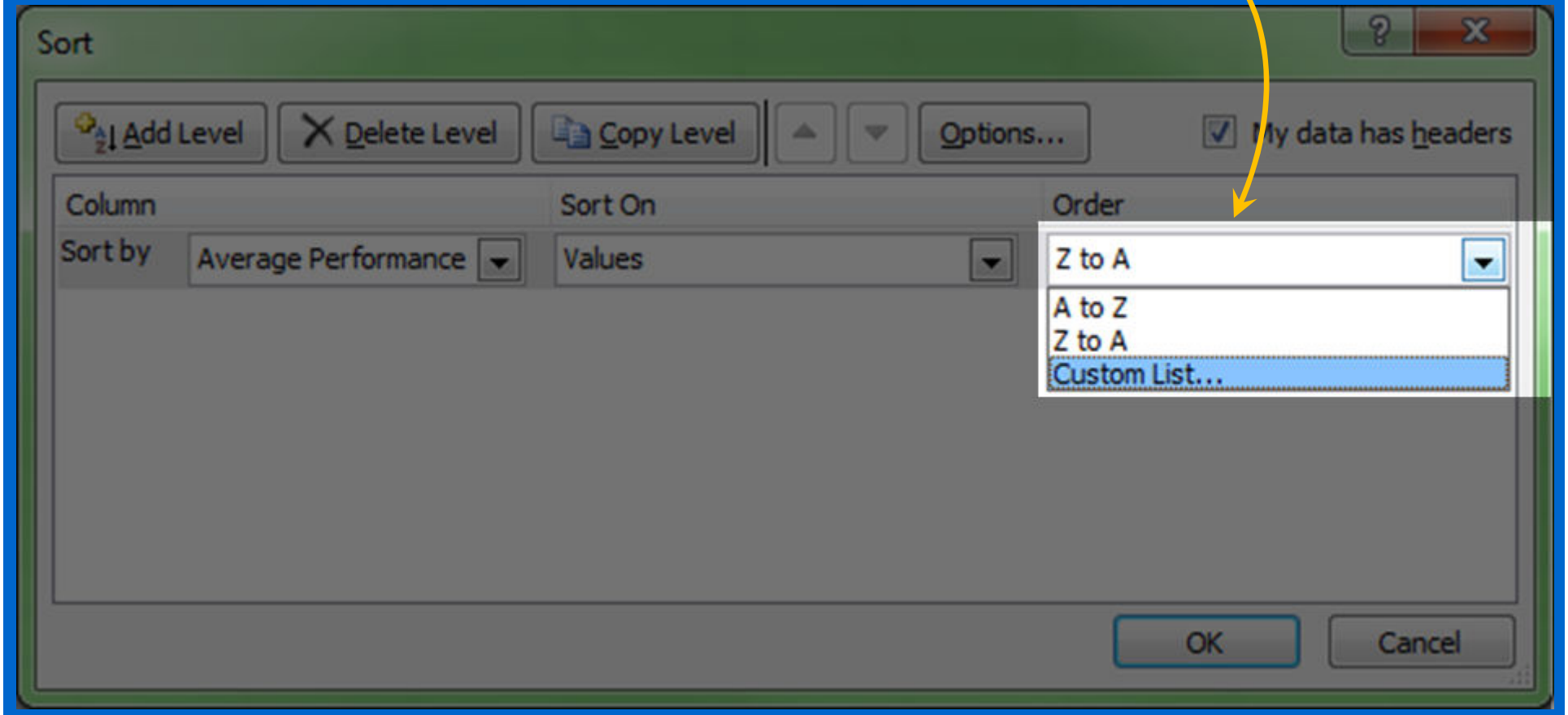
Sorting by a Custom List 3/9

- Click the check box of 'My data has headers'
- Like before, select the 'Average Performance' column in 'Sort by'



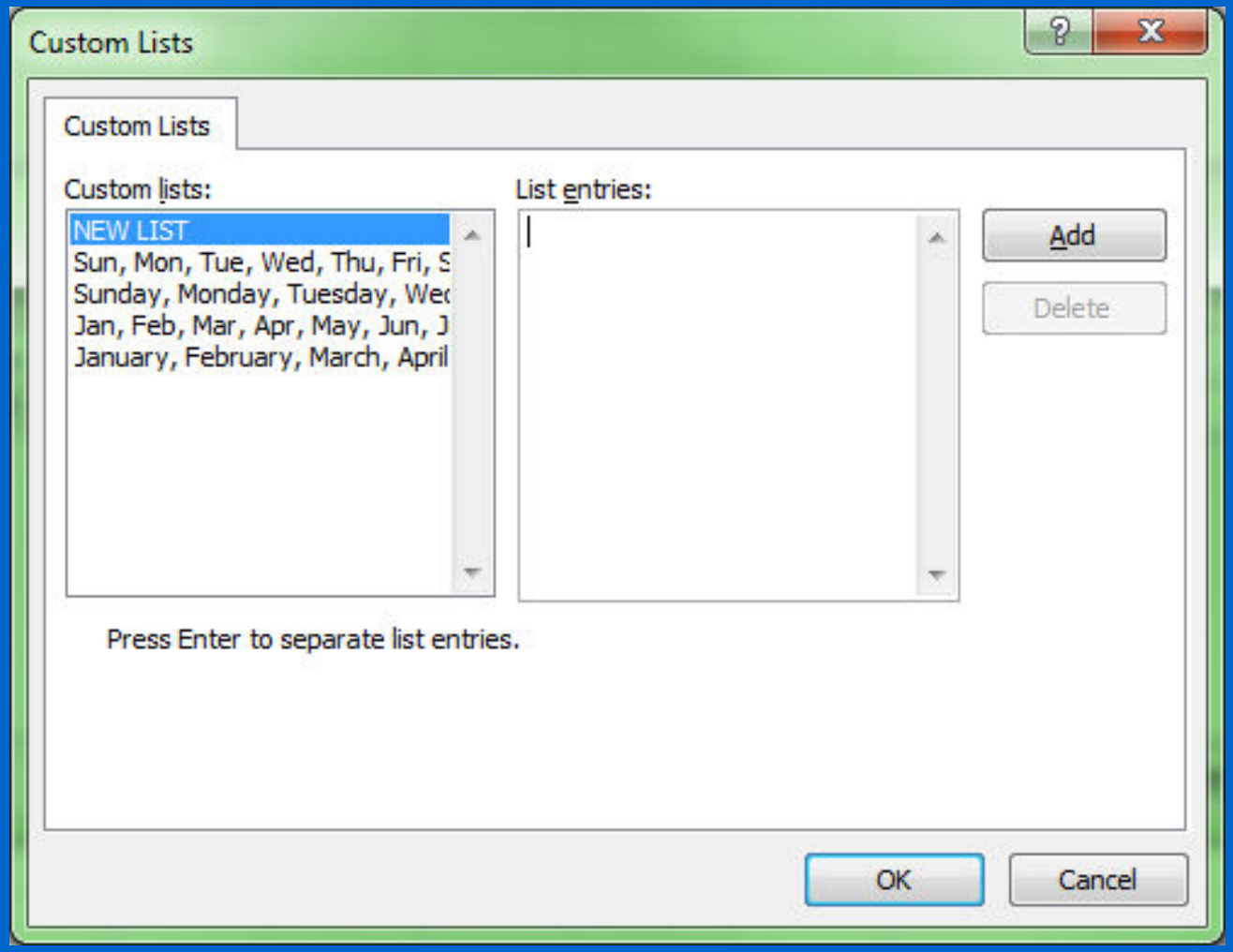
Sorting by a Custom List 4/9

- Select 'Custom List…' in 'Order':



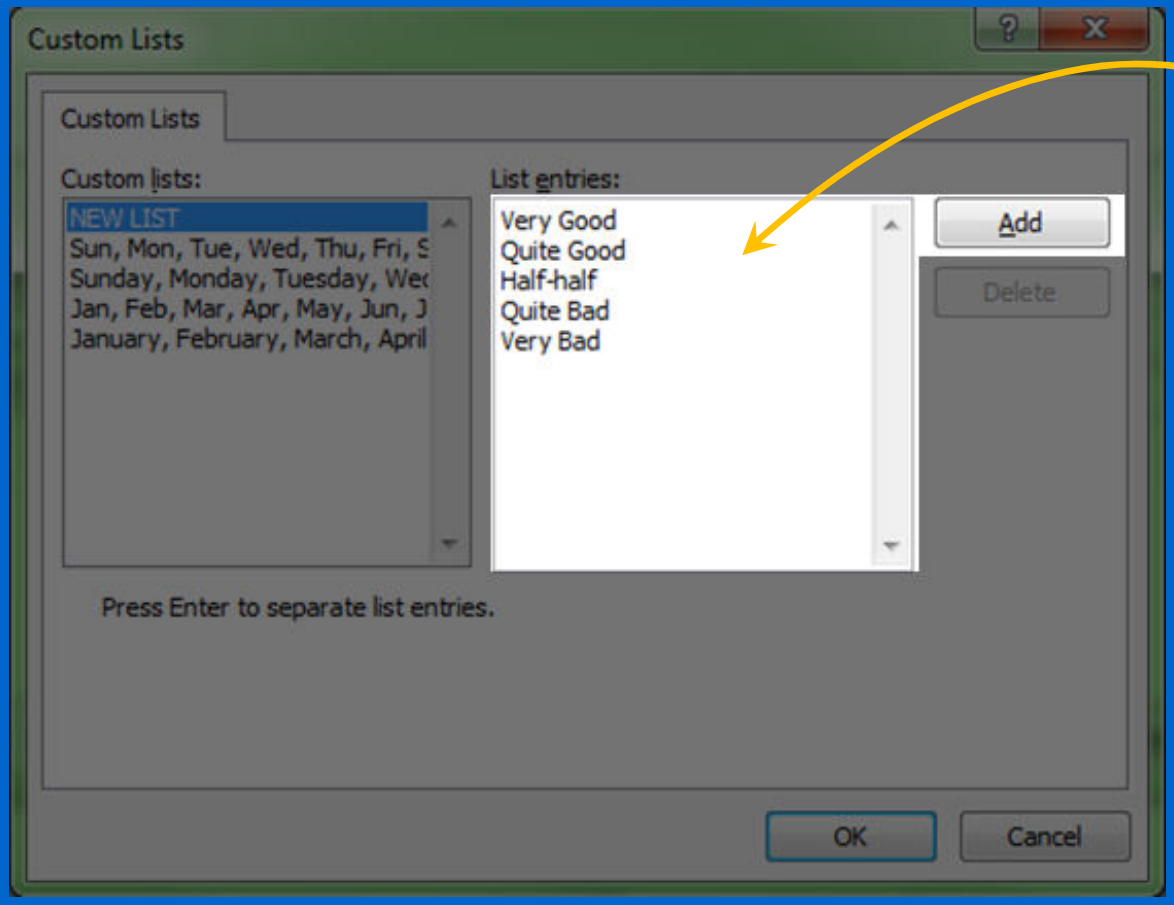
Sorting by a Custom List 5/9

- A 'Custom Lists' window appears:



Sorting by a Custom List 6/9

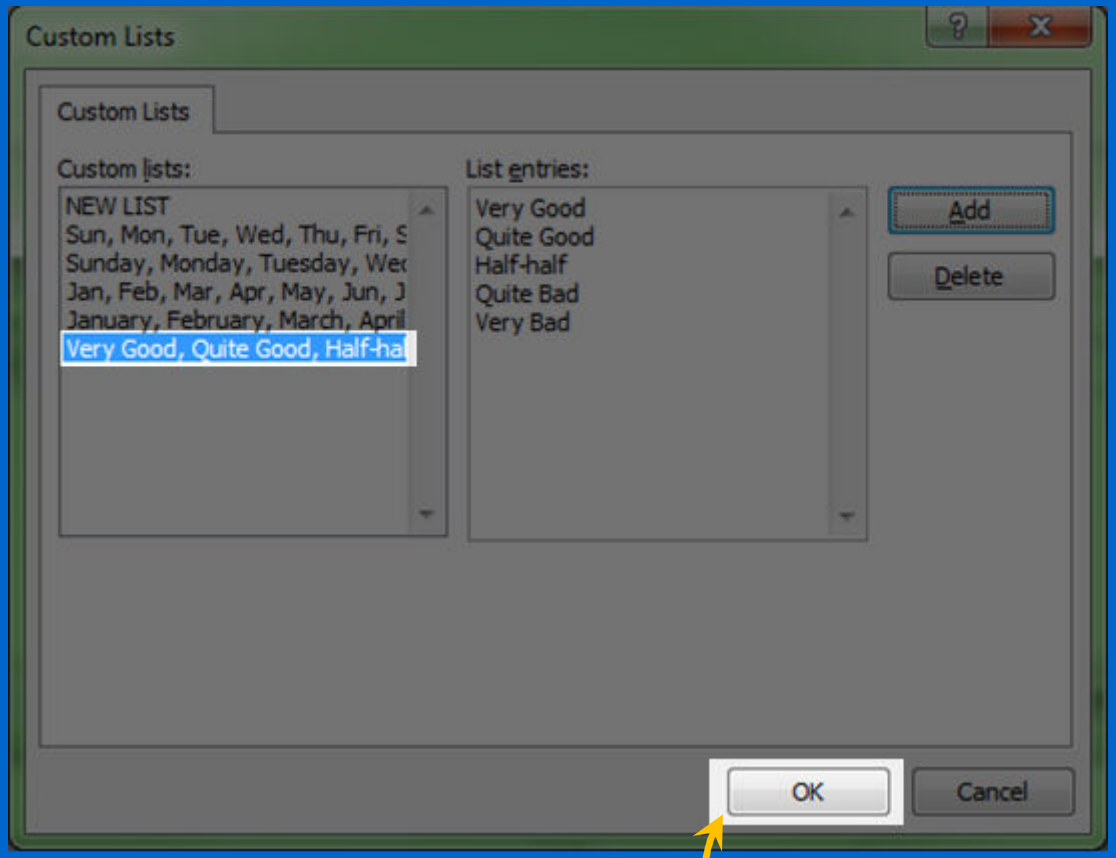
- Type *Very Good*, *Quite Good*, *Half-half* and *Quite Bad* directly into the 'List entries:', then click 'Add':



- The text must be exactly same as the text in the cells

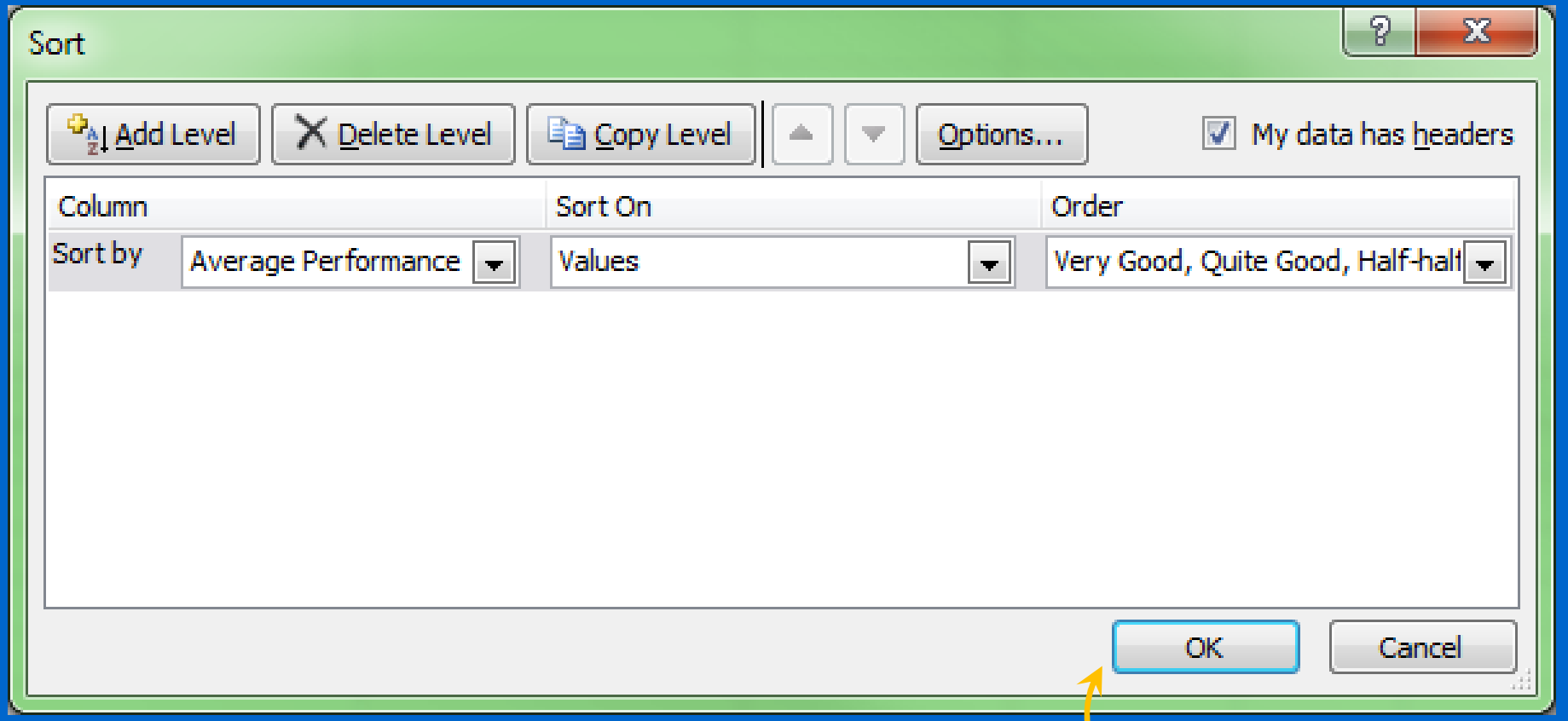
Sorting by a Custom List 7/9

- After clicking 'Add', the 'Custom Lists' will have a new entry in the left hand side:



Sorting by a Custom List 8/9

- A Sort window will be shown:



Sorting by a Custom List 9/9

- After pressing 'OK' twice, Excel will now sort the records using the custom list:

Year of Survey	Average Performance Appraisal	Total sample
2005	Quite Good	13083
2007	Quite Good	13177
2006	Half-half	13180
2008	Half-half	13276
2009	Half-half	13083
2010	Half-half	13140
2011	Quite Bad	8176

- Now, we can easily see the most successful year for the Chief Executive, followed by the average years, then the not very popular years