REL MID-ATLANTIC
Regional Educational Laboratory
at ICF International

District Data Coordinator Toolbox: Implementing Database Connections in Excel

Jason Schoeneberger, Ph.D.
Senior Researcher & Task Lead

Data, data, everywhere

The volume of and the push to make use educational data is growing:

- More people must become data savvy (teachers, coordinators, etc.)
- Leadership may request cyclical reporting to establish and monitor trends
- Little time to document business rules or standardize data storage practices
- Quality control can take time or be difficult to manage

Teachers, principals, administrators and analysts often have difficulty keeping pace.

REL MID-ATLANTIC
Regional Educational Laboratory at ICF International

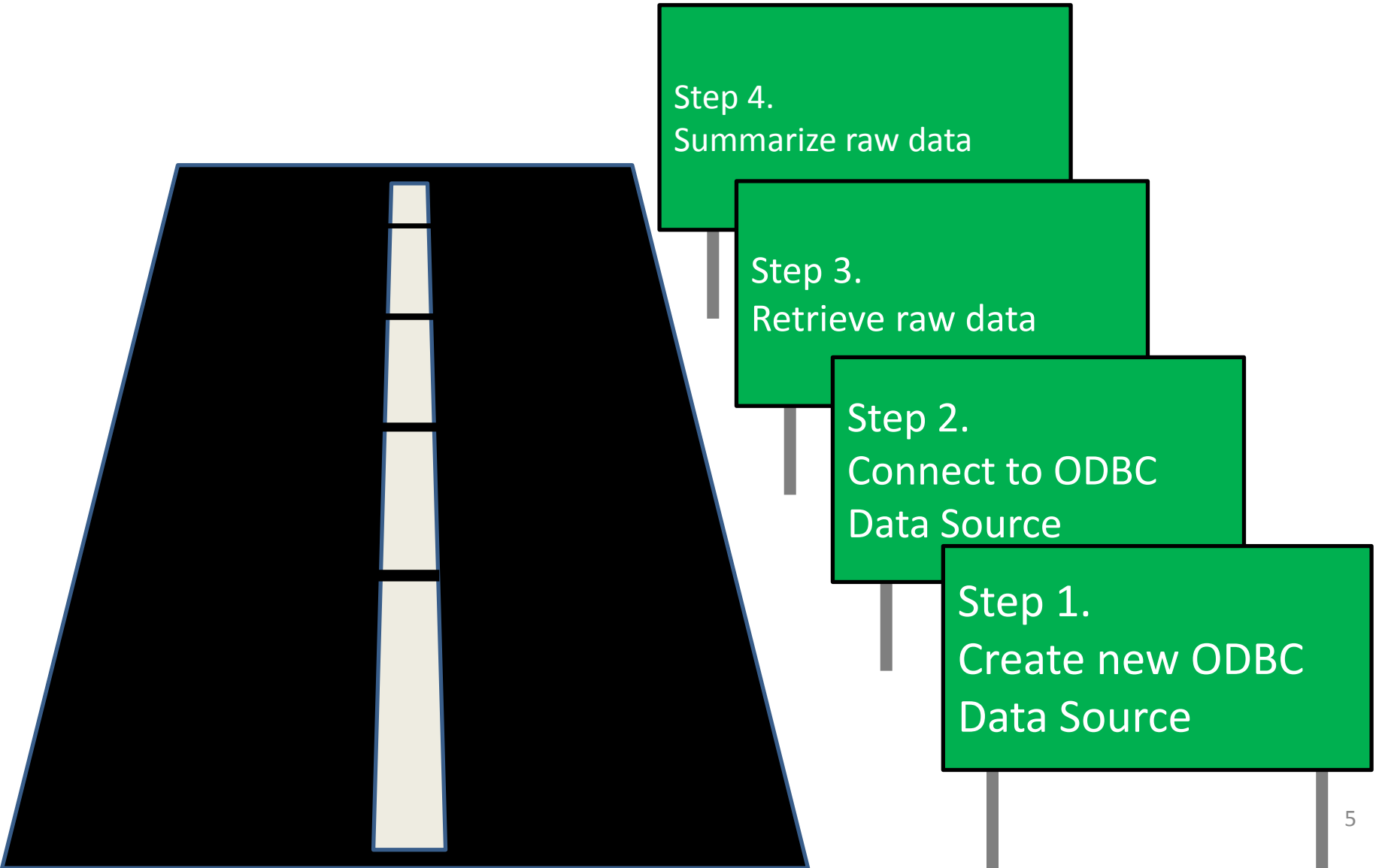
Some familiar scenarios
(using data stored in SQL, Oracle, Access, etc.)

- The same data points are necessary across reporting cycles
- Process to acquire and report data is repetitive across reporting cycles
- A non-technical person may be tasked with reporting responsibility
- Lack of documentation
- Analysts report shortage of storage space on network or external hard drives
- Analysts are maintaining idiosyncratic versions of various data elements (e.g. test score files, student attendance files, etc.)
- Idiosyncratic versions have commonalities across analyst versions
- Separate data requests completed by different analysts yield conflicting results (e.g. a school mean test score)

Database connections

- Databases (e.g., SQL, Oracle, Access, etc.) allow for basic data base connectivity:
  - Open Database Connectivity (ODBC)
  - Object Linking and Embedding Database (OLEDB)
  - These are often standard on computers
- ODBC/OLEDB connections are frameworks to allow data manipulation software (e.g. Excel, SPSS, SAS) to communicate with databases

Road map to data connectivity



Data connectivity example

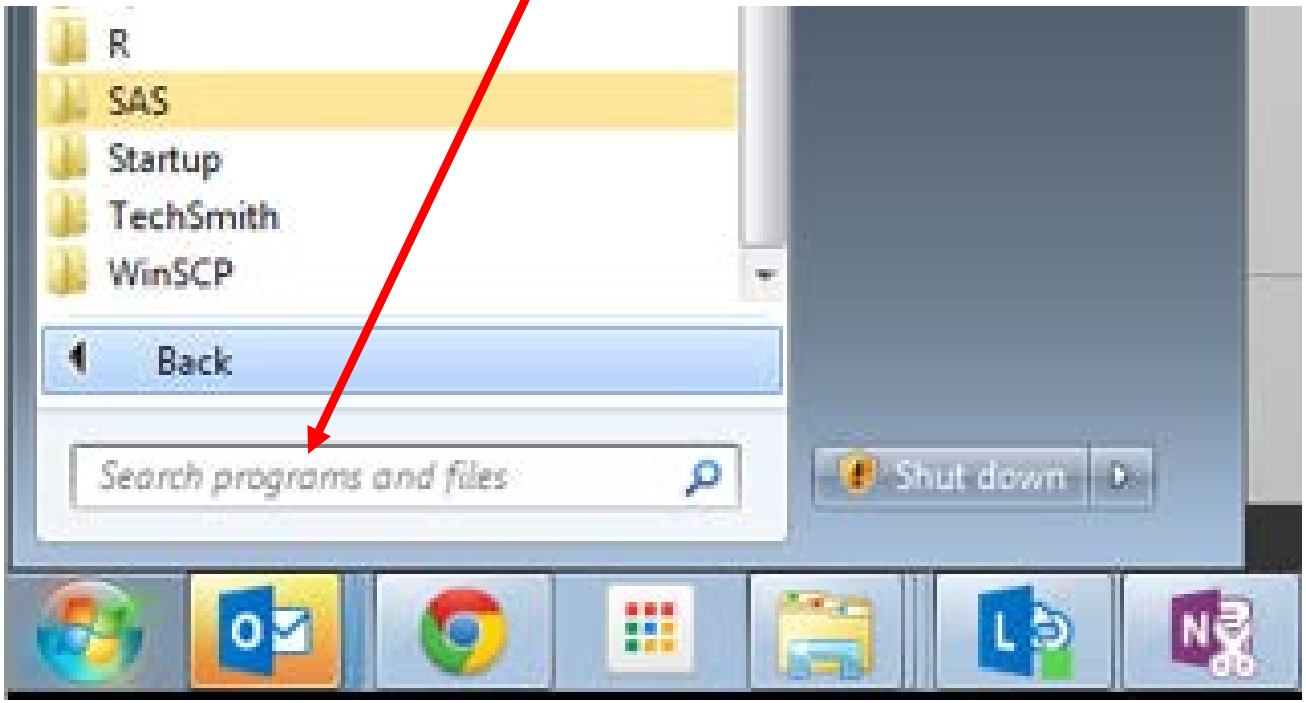
- To follow the steps in our road map to connectivity, let's assume the following example:
  - District leadership has asked us to examine reading achievement as measured by reading assessment achievement levels
  - Leadership is specifically interested in 6th grade student performance
  - They want to examine performance by student Limited English Proficient (LEP) status.
  - The data we need to obtain are stored in an Access database

Creating an ODBC data source

- The first step is to create an ODBC Data Source centered on an existing database such as Access, SQL, or Oracle. ODBC Data Sources are frameworks, or linkages for software packages such as Excel to communicate with databases

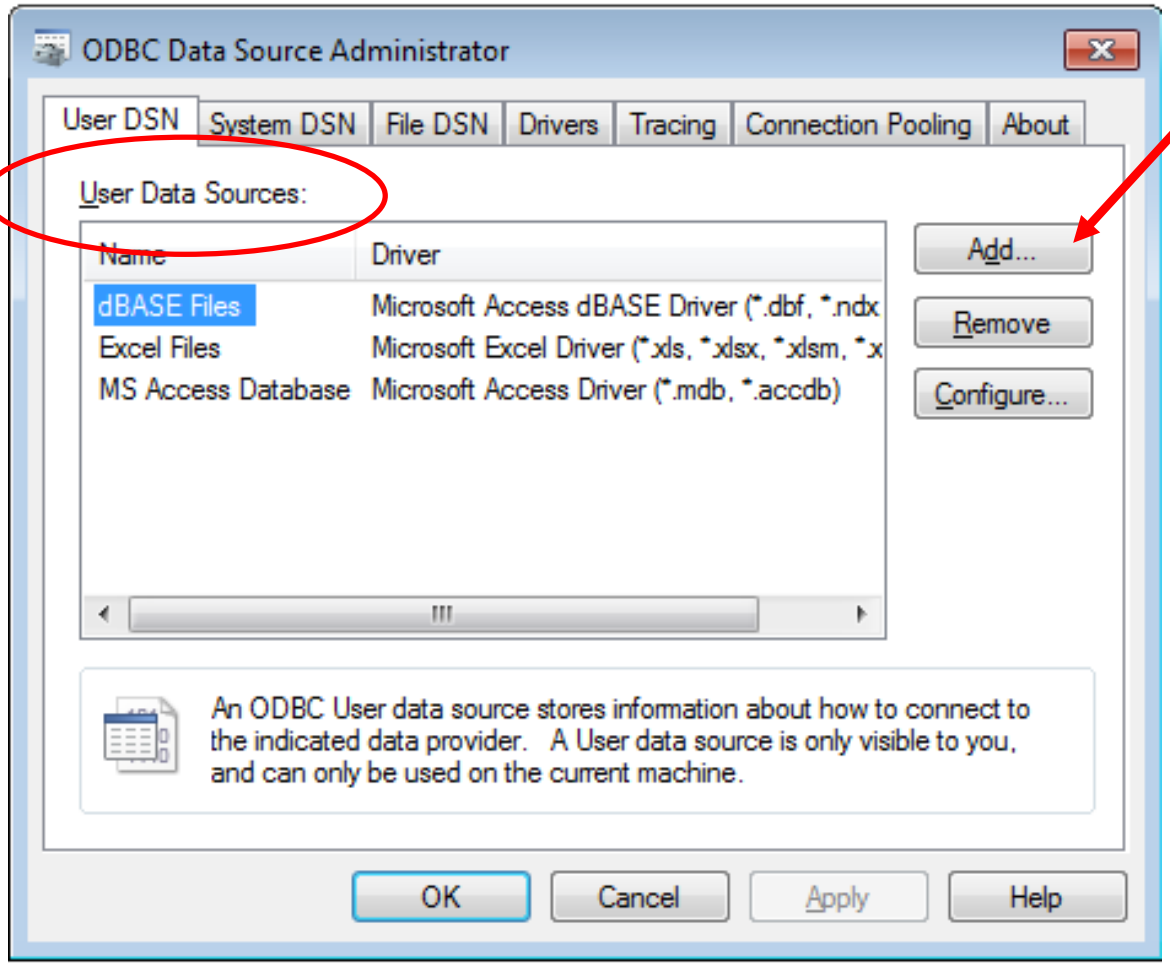
Open ODBC administrator window

- Type ‘ODBC’ in Search Box and press Enter



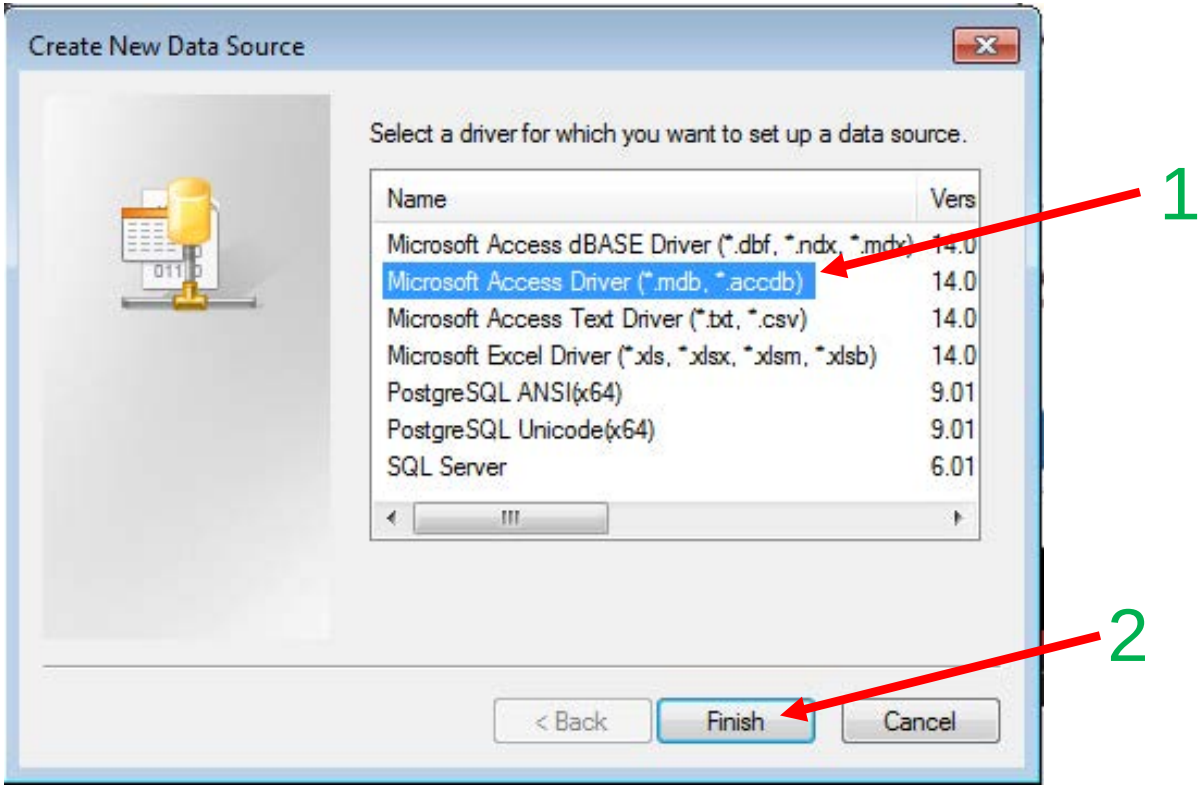
Add a new data source

- Click ‘Add’ to begin adding a new ODBC data source



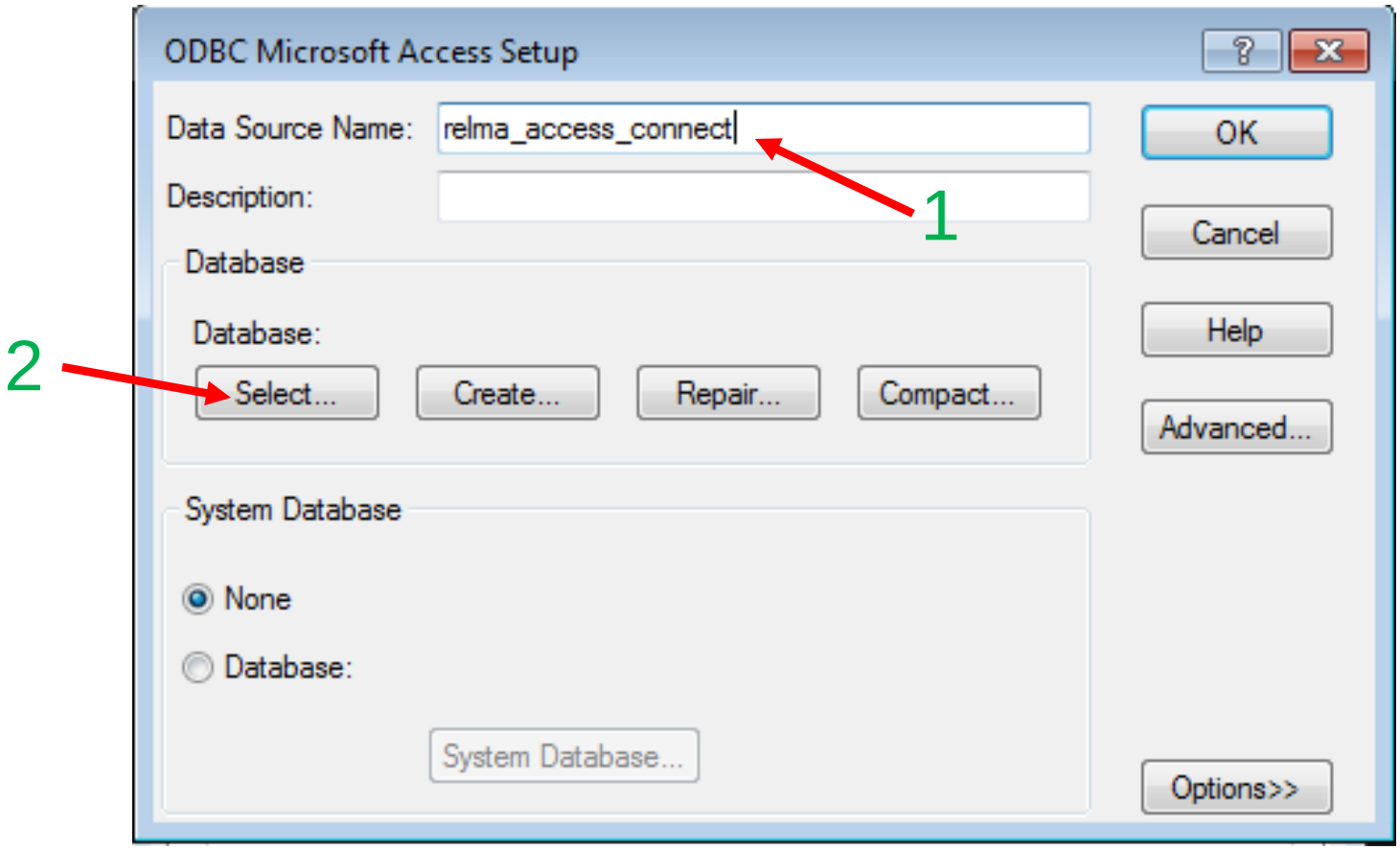
Choose a driver for the data source

1. Choose driver for connection to a source (in this example, we connect to an Access database)
2. Click Finish



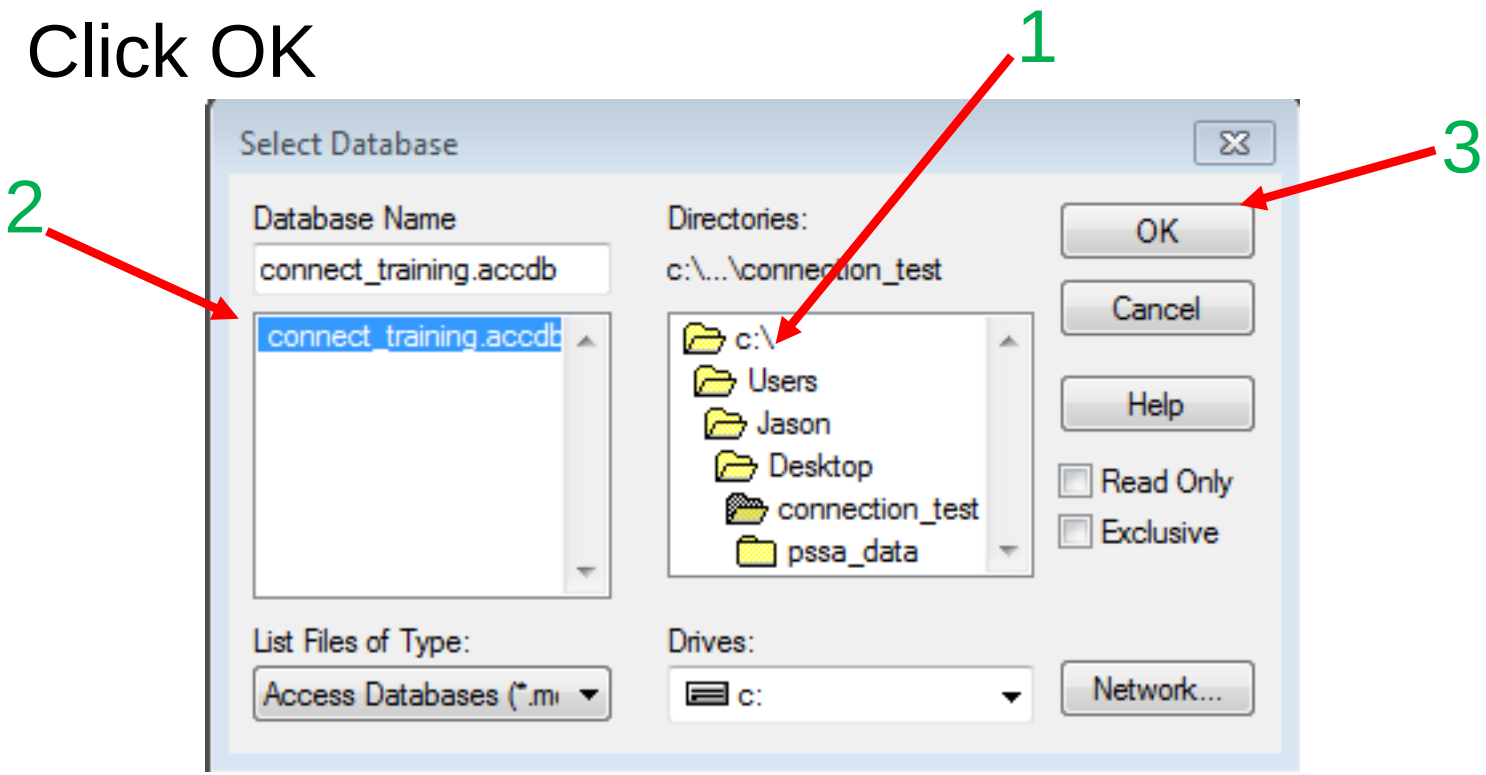
Name the database connection

1. Name the connection to the database
2. Click Select button under Database



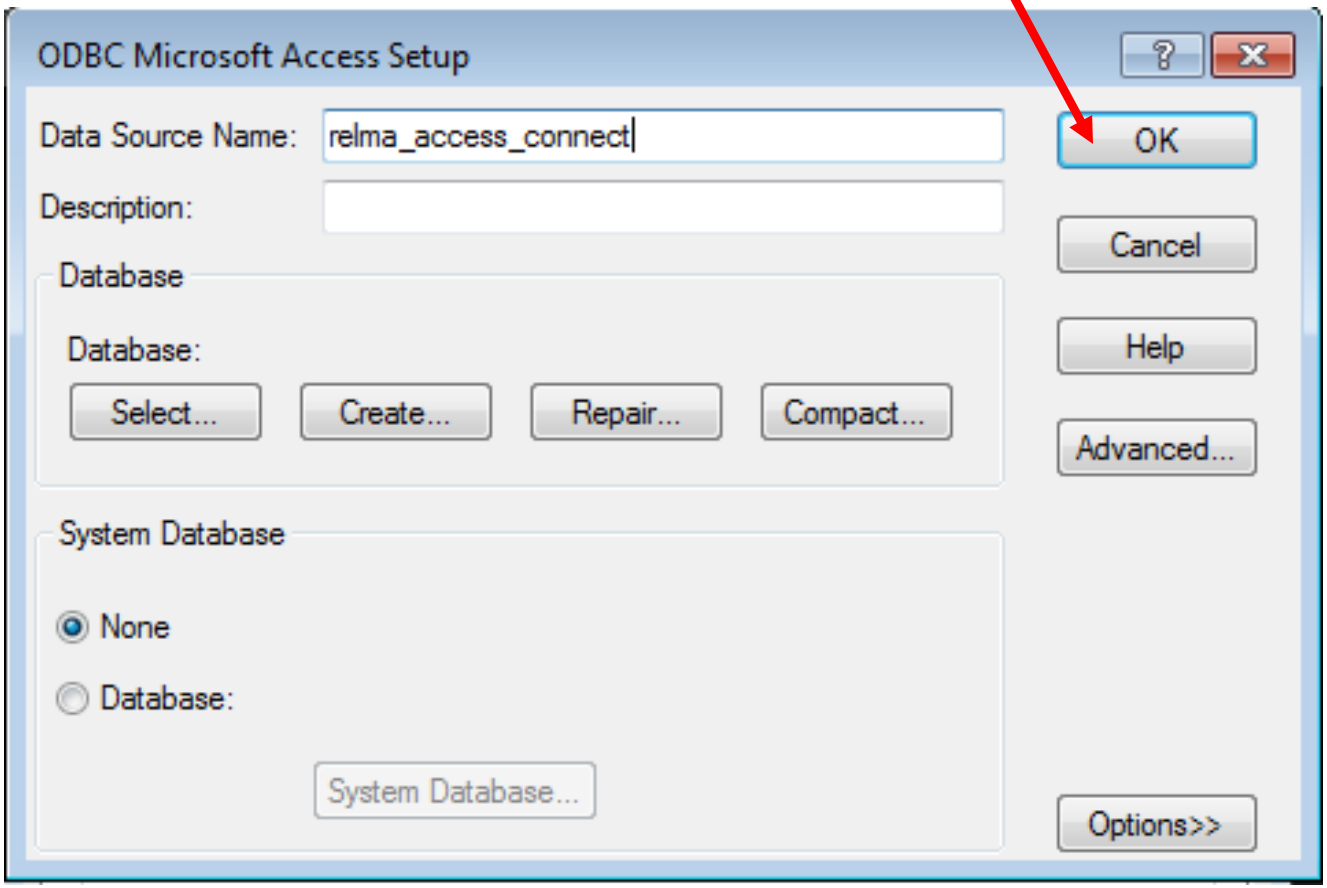
Link to the source database

1. Navigate to location of the database (the Access database we want to connect to in this example)
2. Select source database
3. Click OK



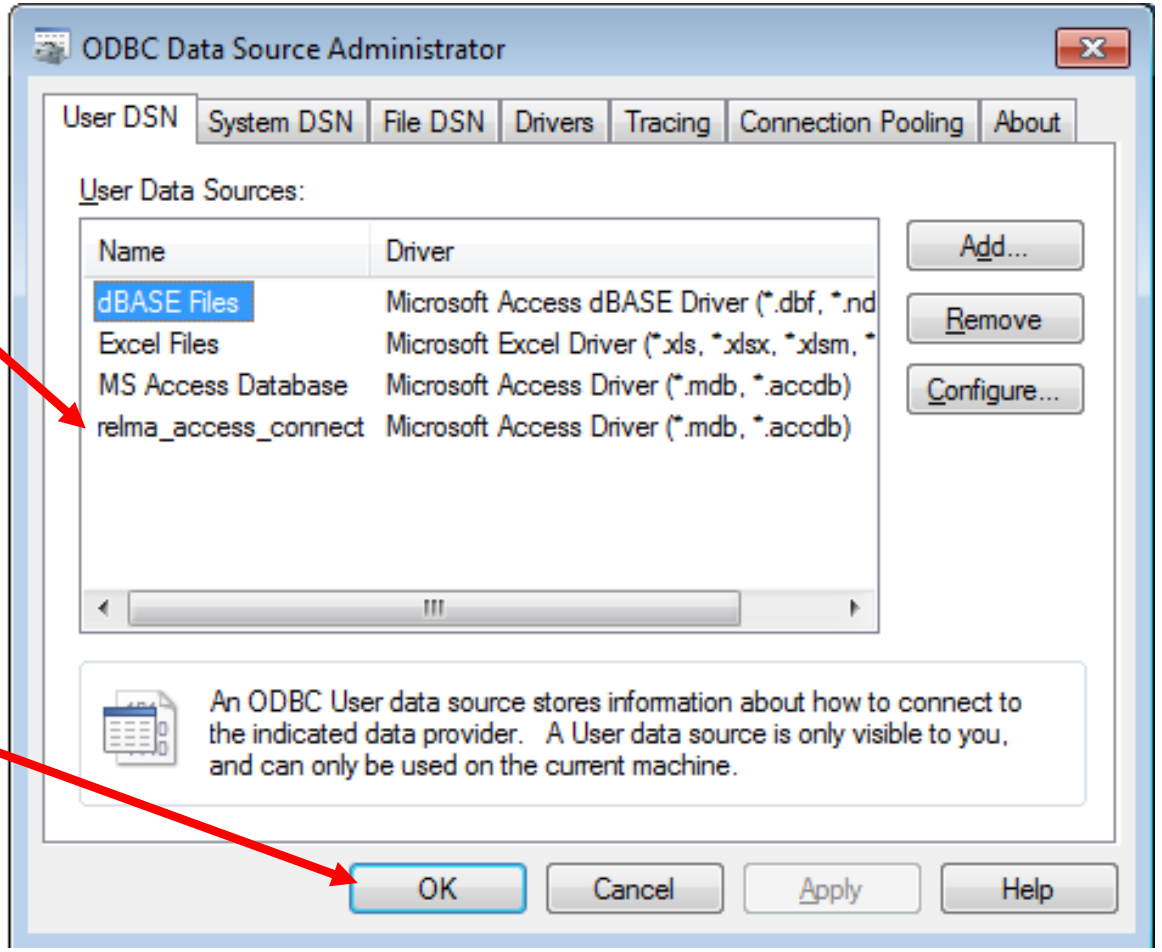
Click OK for data source name

- Click OK button under Database



New data source appears in ODBC directory

1. New data source listed in ODBC directory
2. Click OK

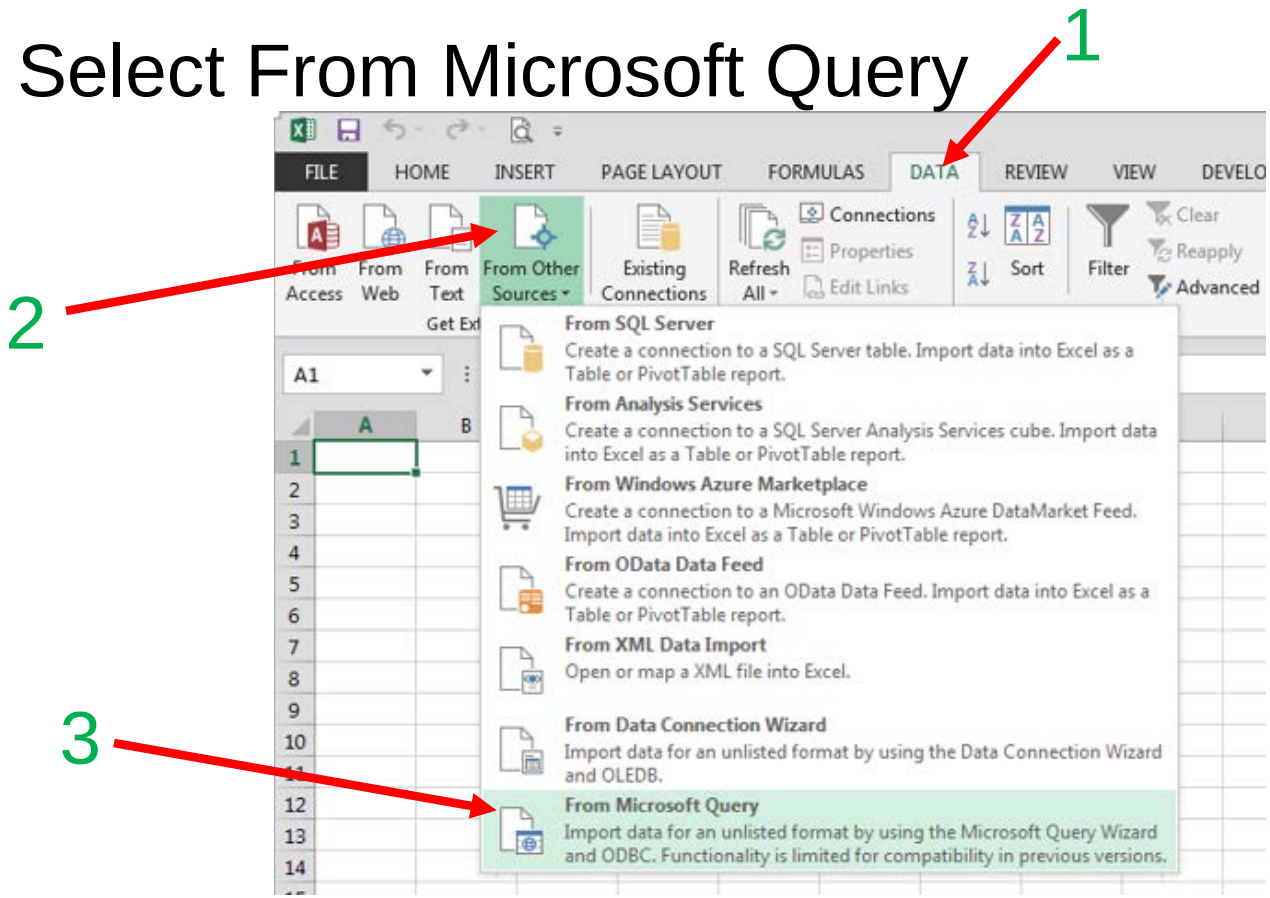


Connecting to a Database

- Now that our ODBC data source exists for communicating with the database, the information in the database can be extracted directly into other software packages (e.g. Excel) for further manipulation

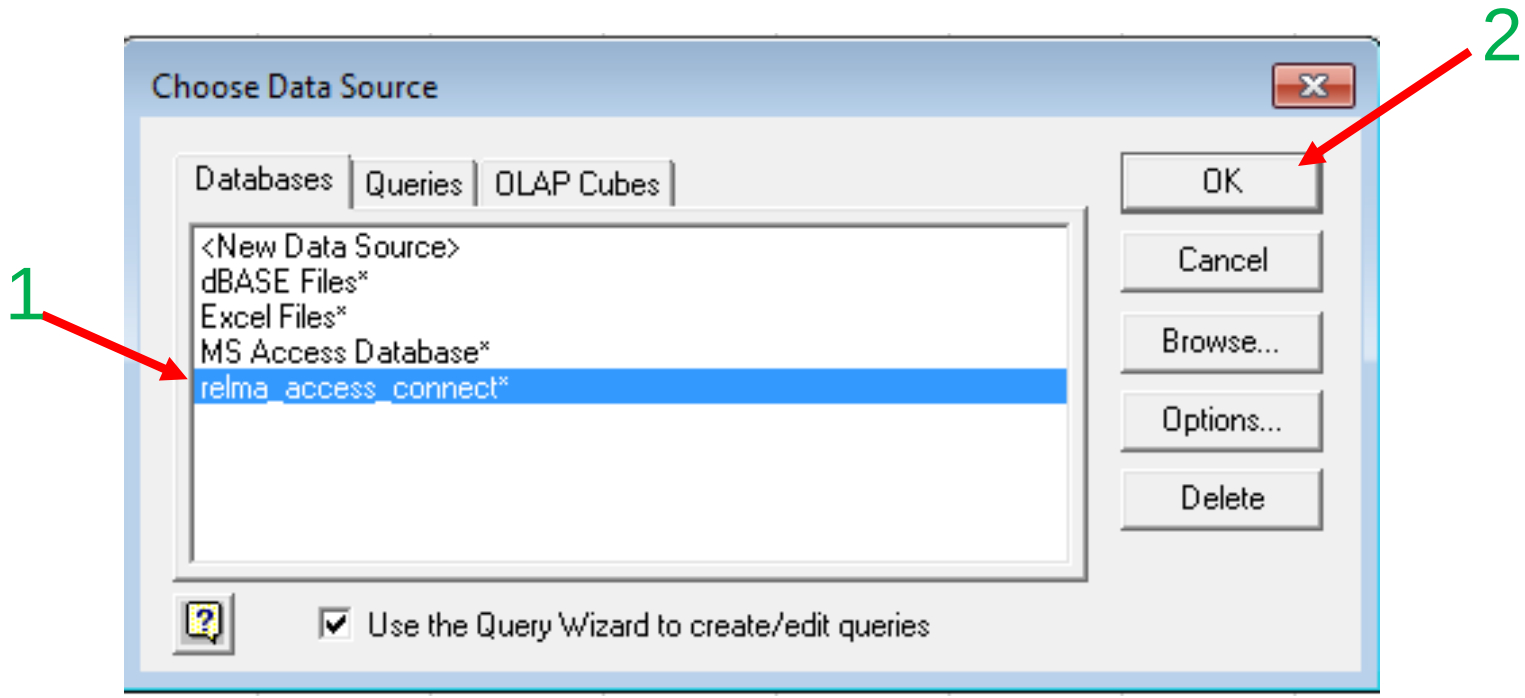
Connect to database using Excel

1. Navigate to the Data tab in Excel
2. Click on From Other Sources
3. Select From Microsoft Query



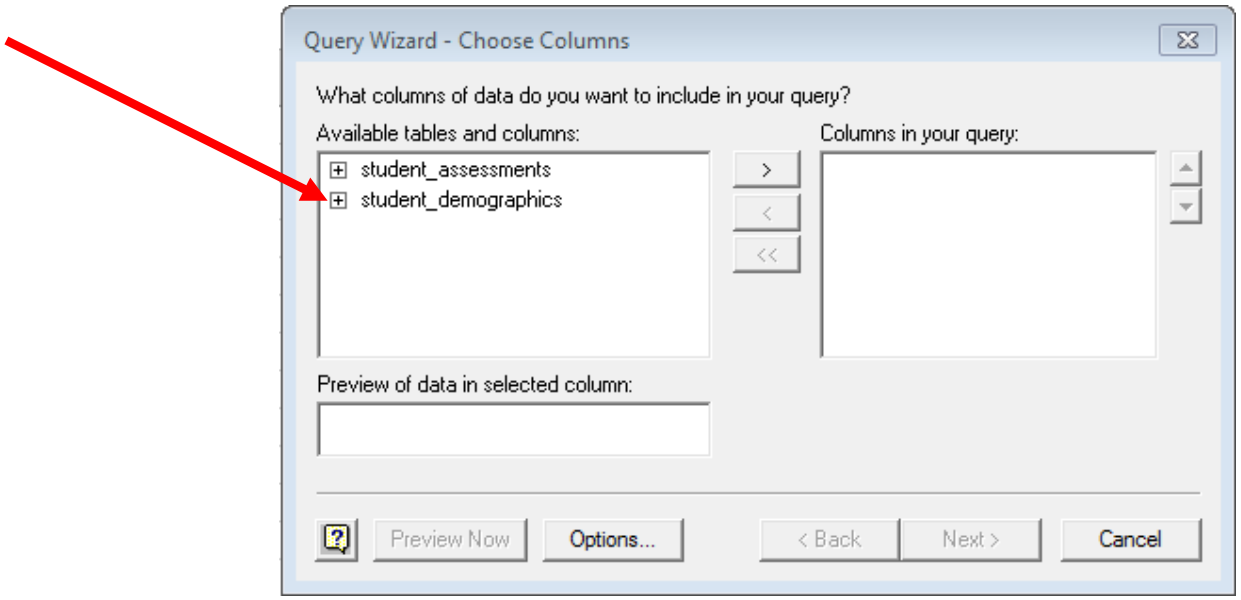
Choose your data source

1. Select the data source of interest (in this example, the relma_access_connect data source)
2. Click OK



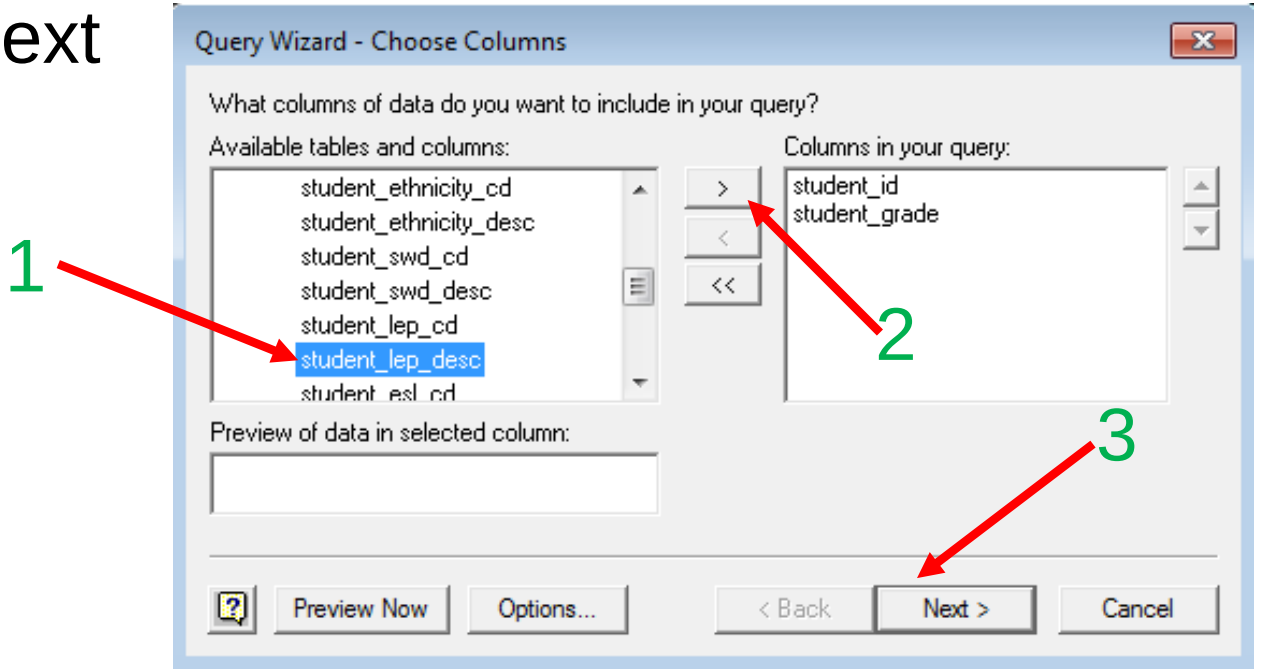
View available tables in data source

- Now we can view data tables in the database
  - student_assessments table
  - student_demographics table
- Press '+' button to list variables in student_demographics table



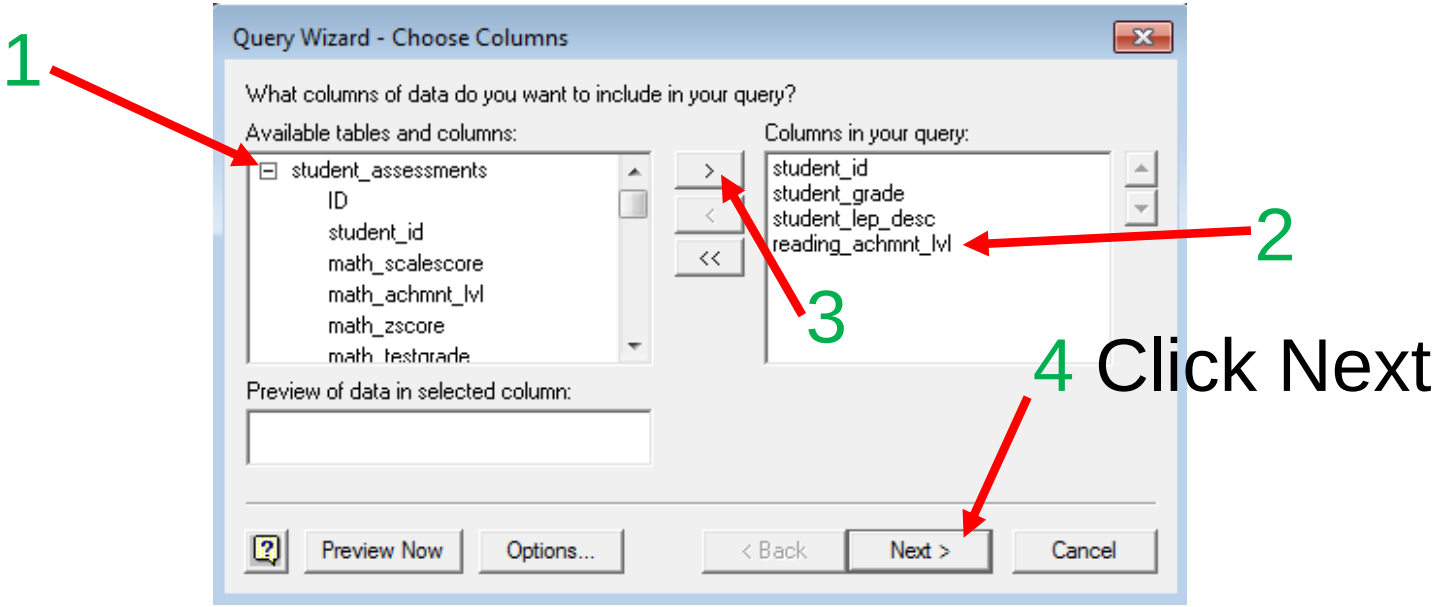
Select demographic variables

1. Click on variables in student_demographics table
2. Click '>' arrow to move variables to query window
   a) Student_id (student identification number)
   b) Student_grade (student grade level)
   c) Student_lep_desc (student LEP status description)
3. Click Next



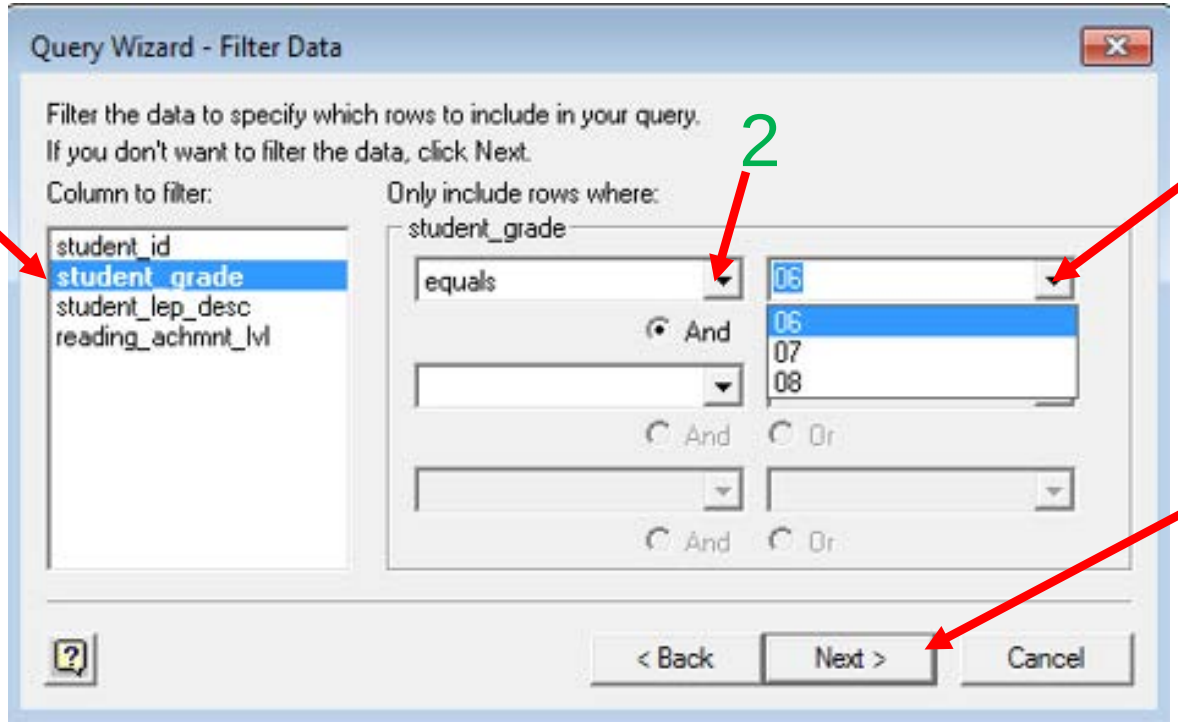
Select assessment variables

1. Press '+' to list variables in student_assessments table
2. Select variables (reading_achmnt_lvl) from student_assessments table
3. Click '>' arrow to move variable to query window



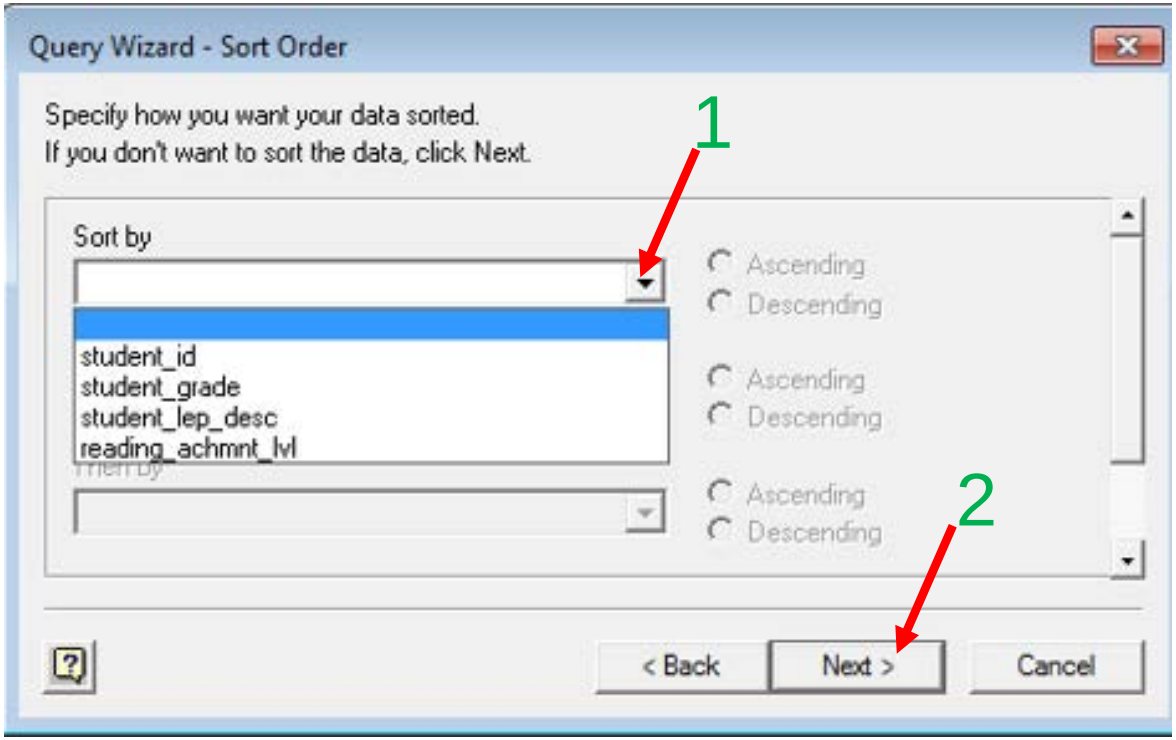
Filter data during data retrieval

- Returned data can be filtered during retrieval
  1. Select student_grade from Column to filter window
  2. Use drop-down to select criteria operation (‘=‘)
  3. Select ‘06’ to return only 6th grade student data



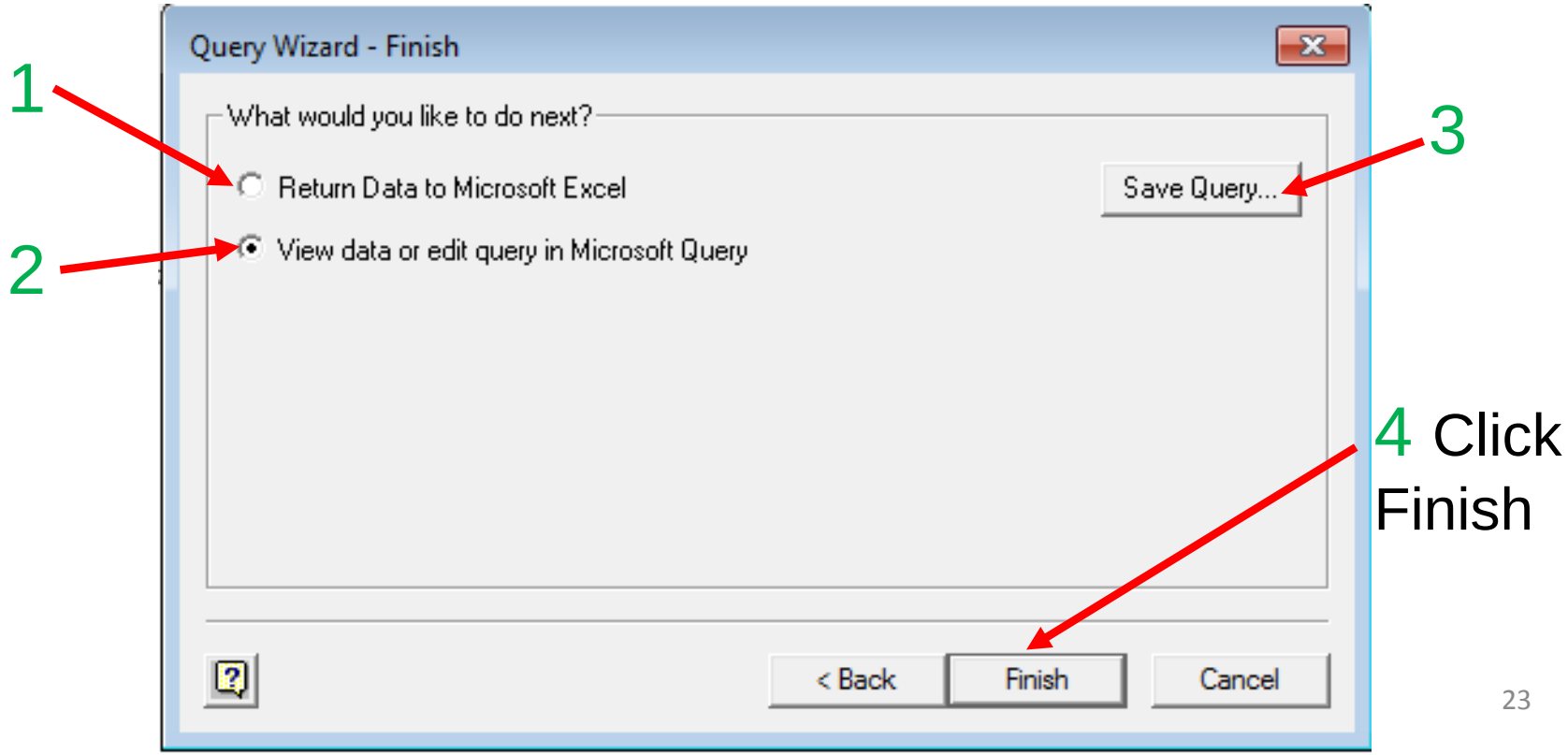
Sort data during data retrieval

- The query can sort returned data during retrieval
  1. Use drop-down to select variable to sort data by (in this example, we do not wish to sort our data)
  2. Click Next



Select how you wish to view the results

1. Results can be immediately returned to Excel, or
2. Query can be viewed/edited further and
3. Query can be saved

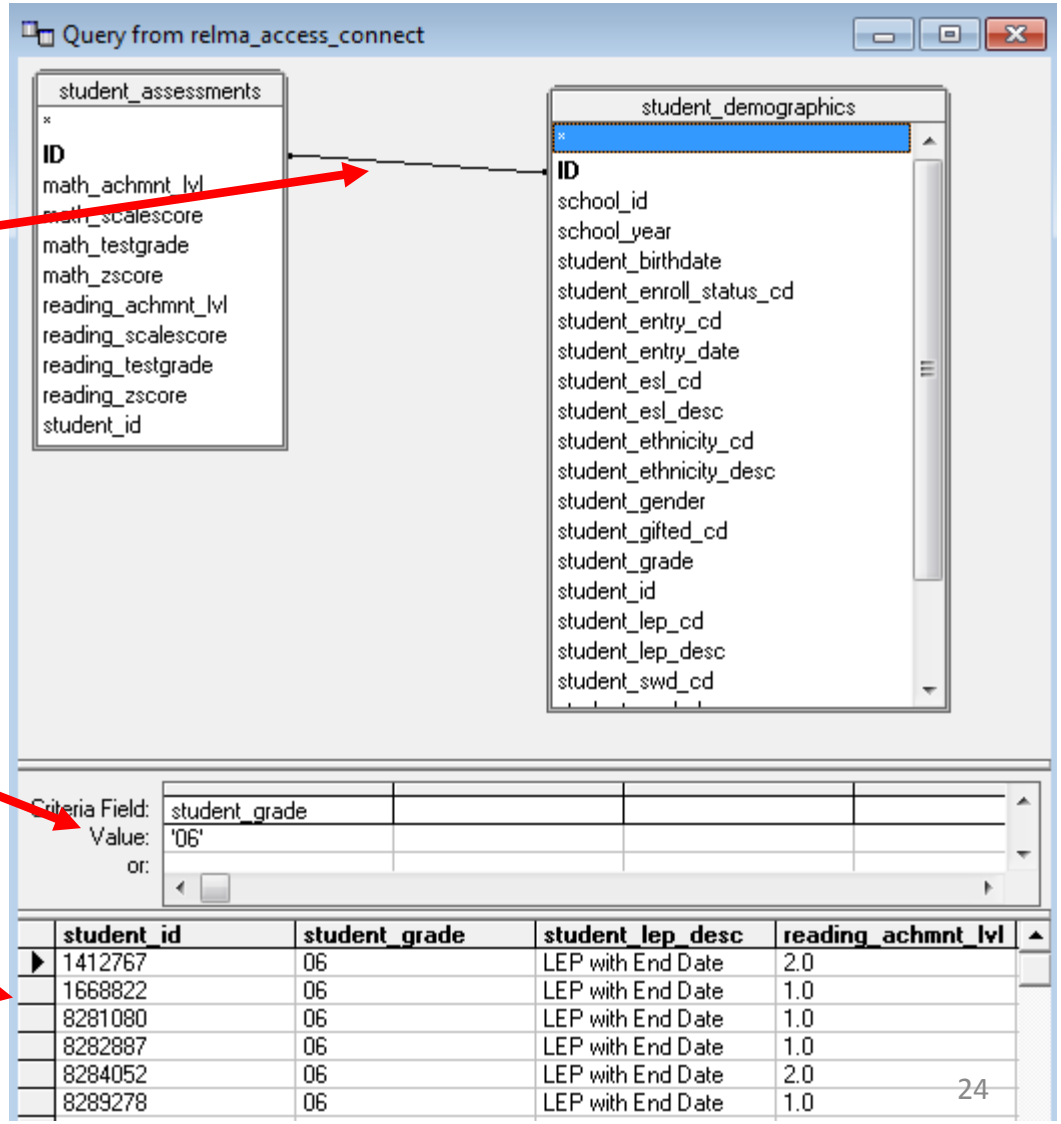


Results in Microsoft Query

Table joins depict relationships (Access creates 'ID' column in each table and joins tables by default)

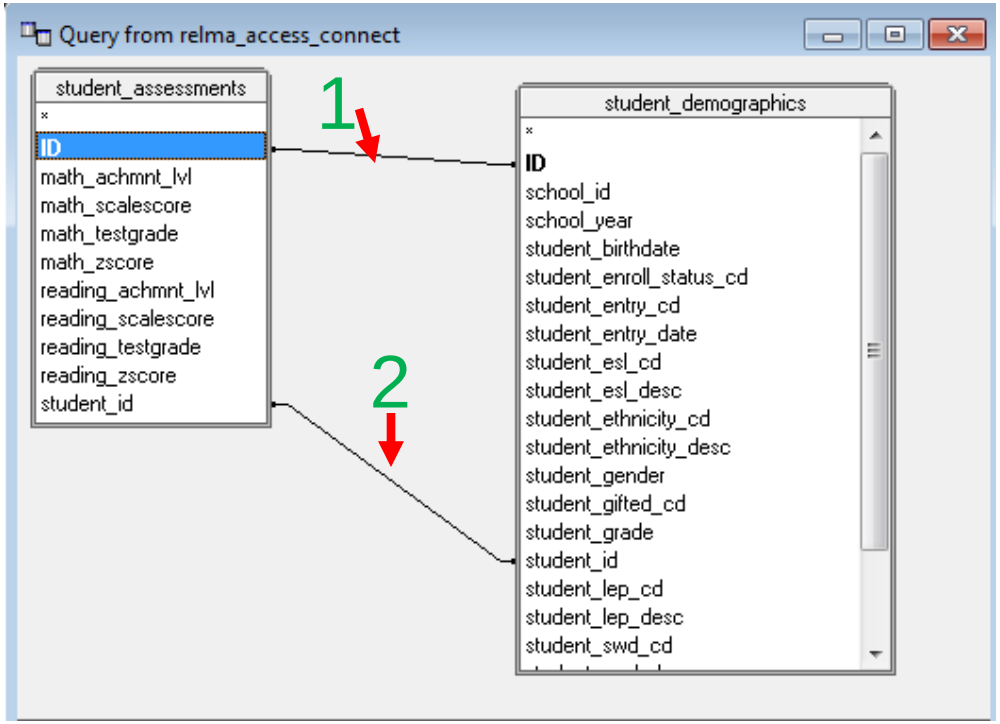
Filter criteria

Retrieved data preview



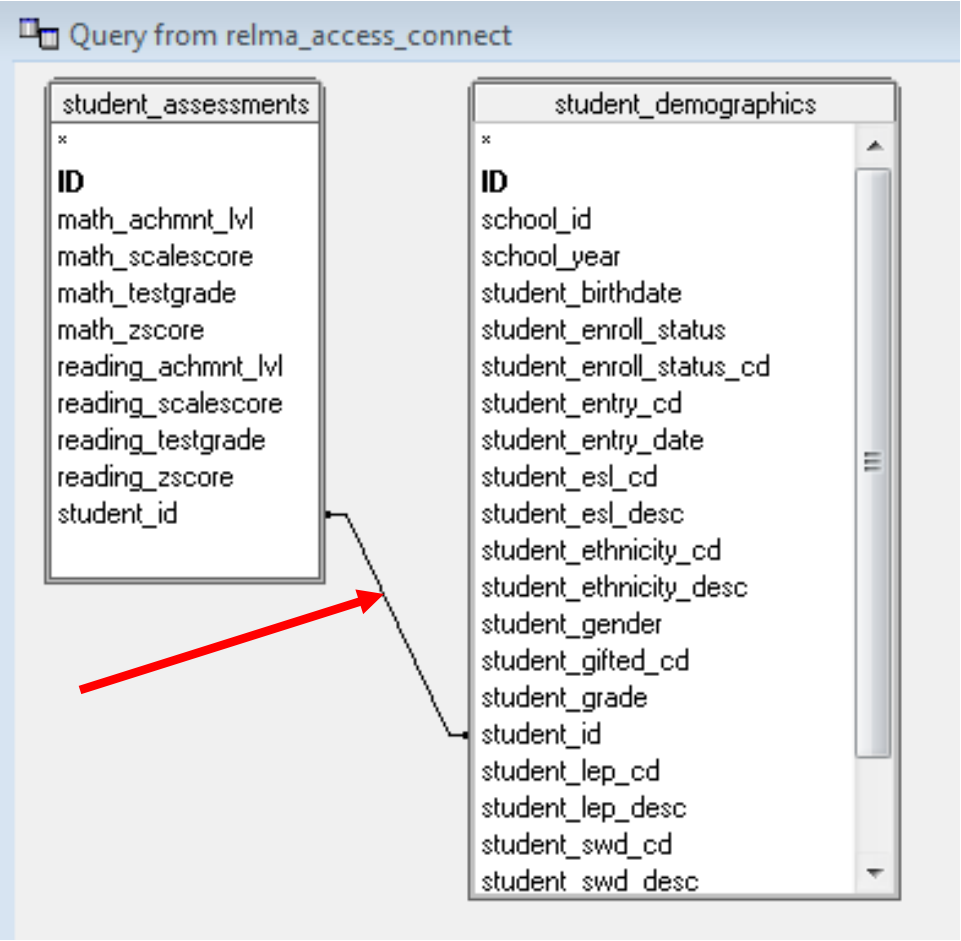
Adding/deleting relationships

1. Click on join to select, press ‘Delete’ to remove ID-to-ID relationship
2. Click and hold on variable in one table, drag mouse to corresponding field in other table (in this example, create relationship from student_id-to-student_id)



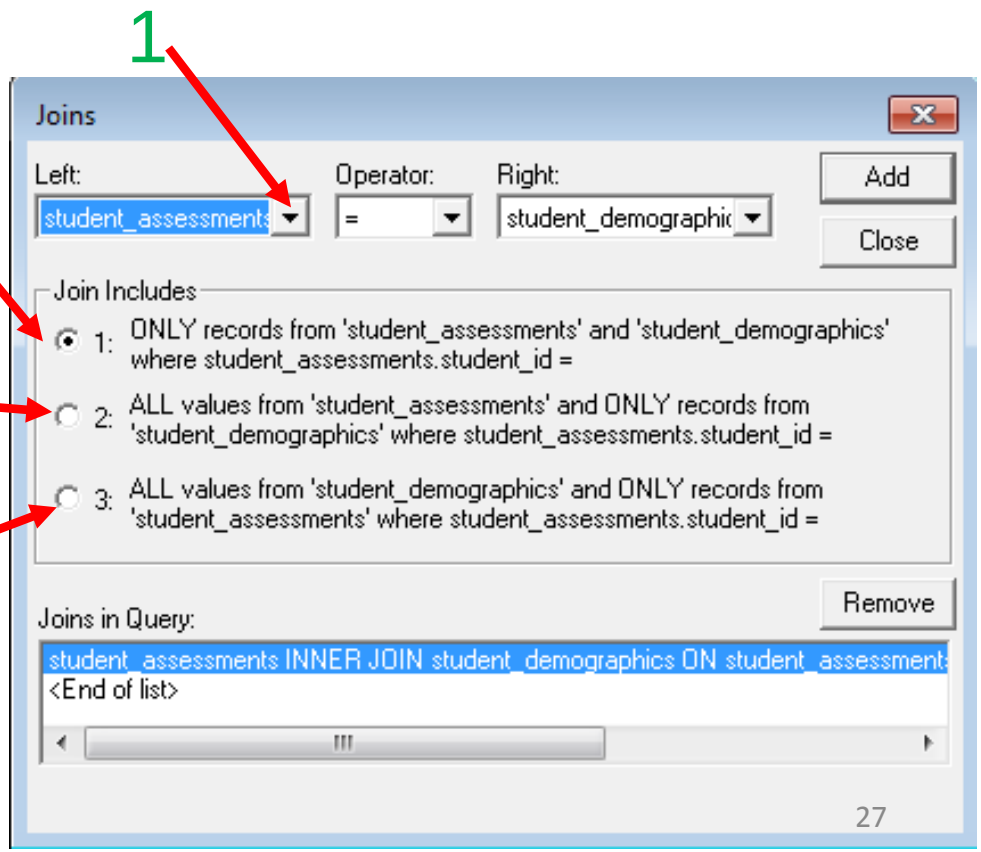
Changing the nature of relationships

- Double-click on student_id-to-student_id join



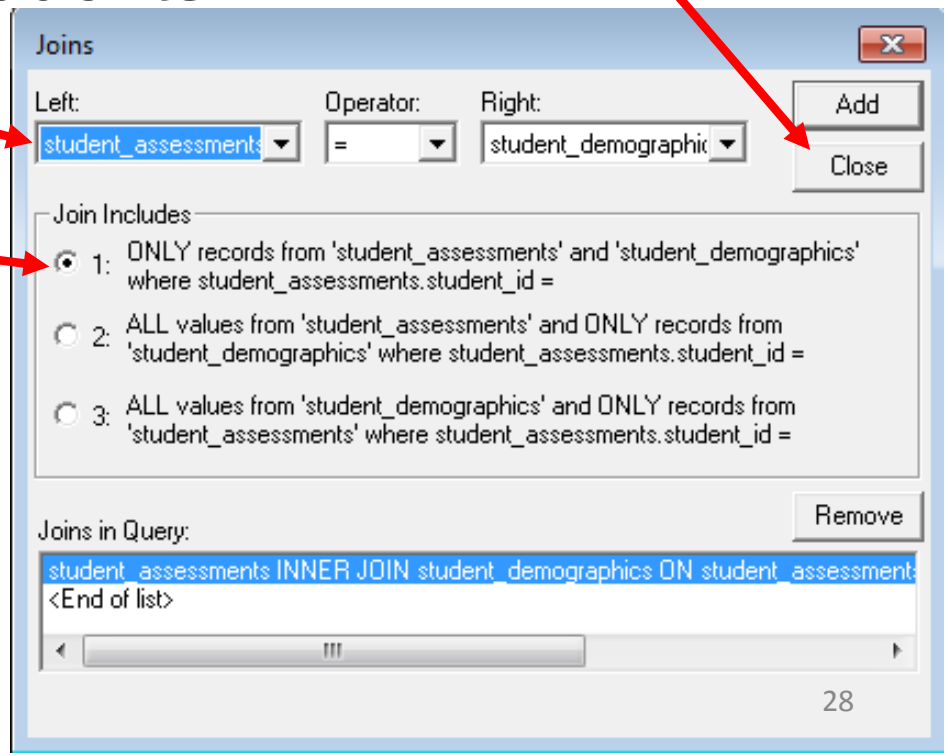
Changing the nature of relationships

1. Change variables to join on using drop-downs
2. Change type of join using radio button
   a) Join 1 returns *only* records with matching student_ids in both tables
   b) Join 2 returns *all* records from student_assessments and records with matches in student_demographics
   c) Join 3 returns *all* records from student_demographics and records with matches in student_assessments



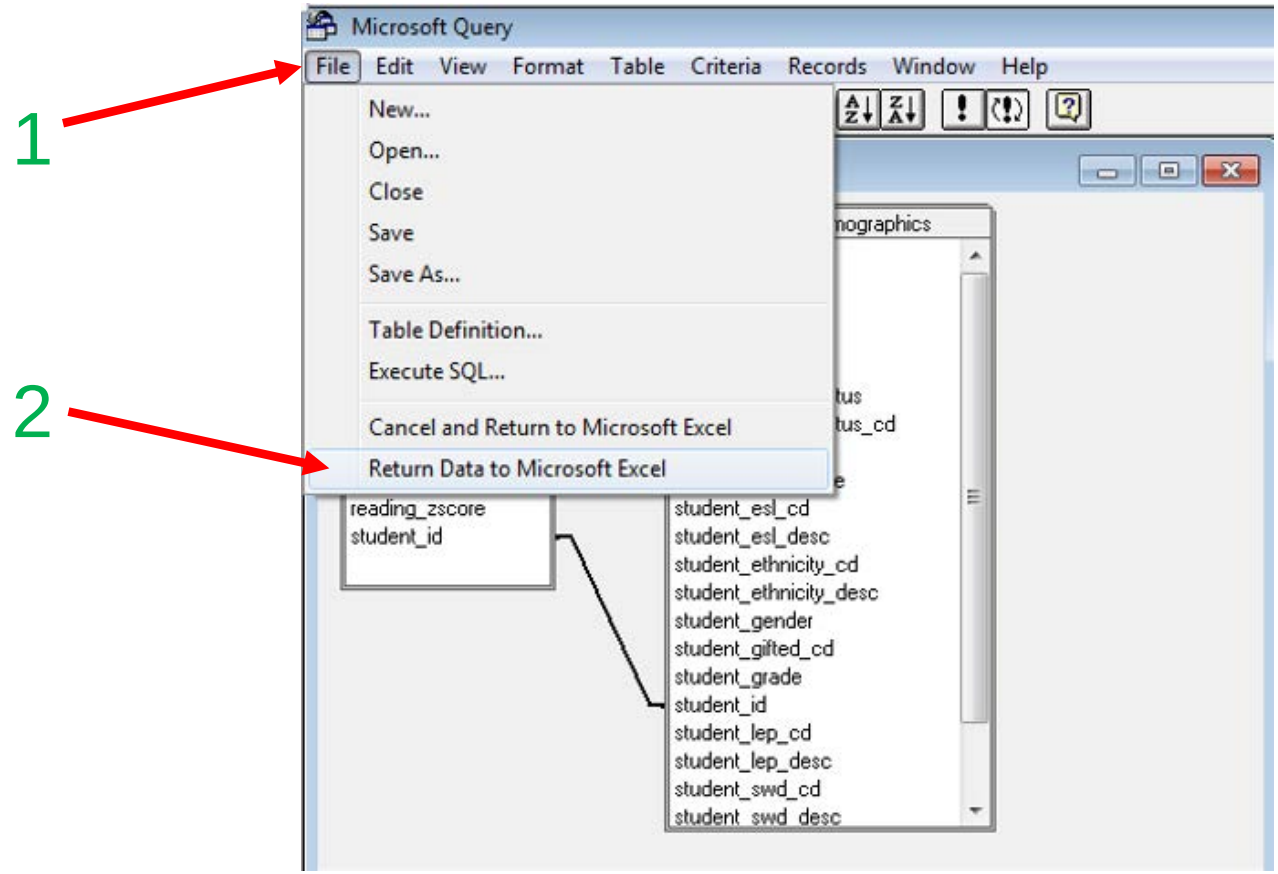
Changing relationships – this example

1. Joining on common student_id variables is appropriate for our example
2. Our join (1) is appropriate for comparing performance of LEP students
3. Click Close



Return data to Excel

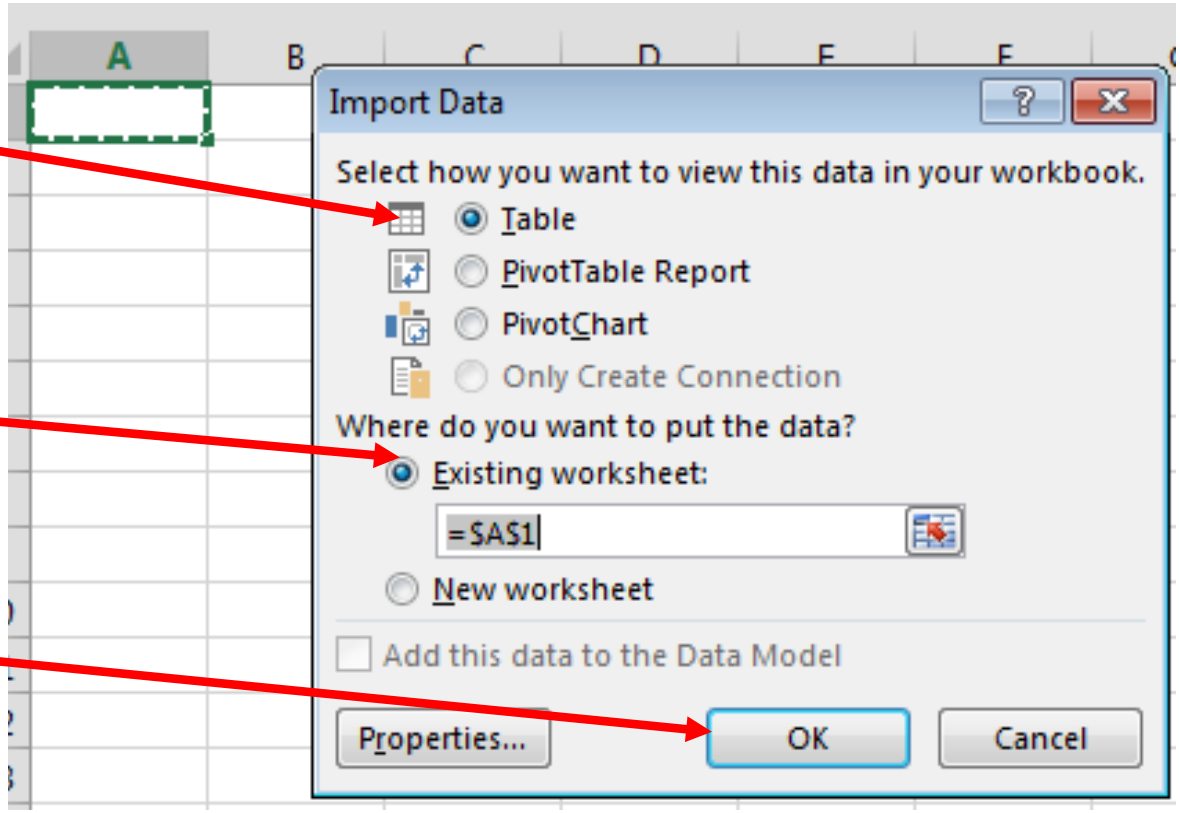
1. Click File
2. Select Return Data to Excel



Specify data type and destination

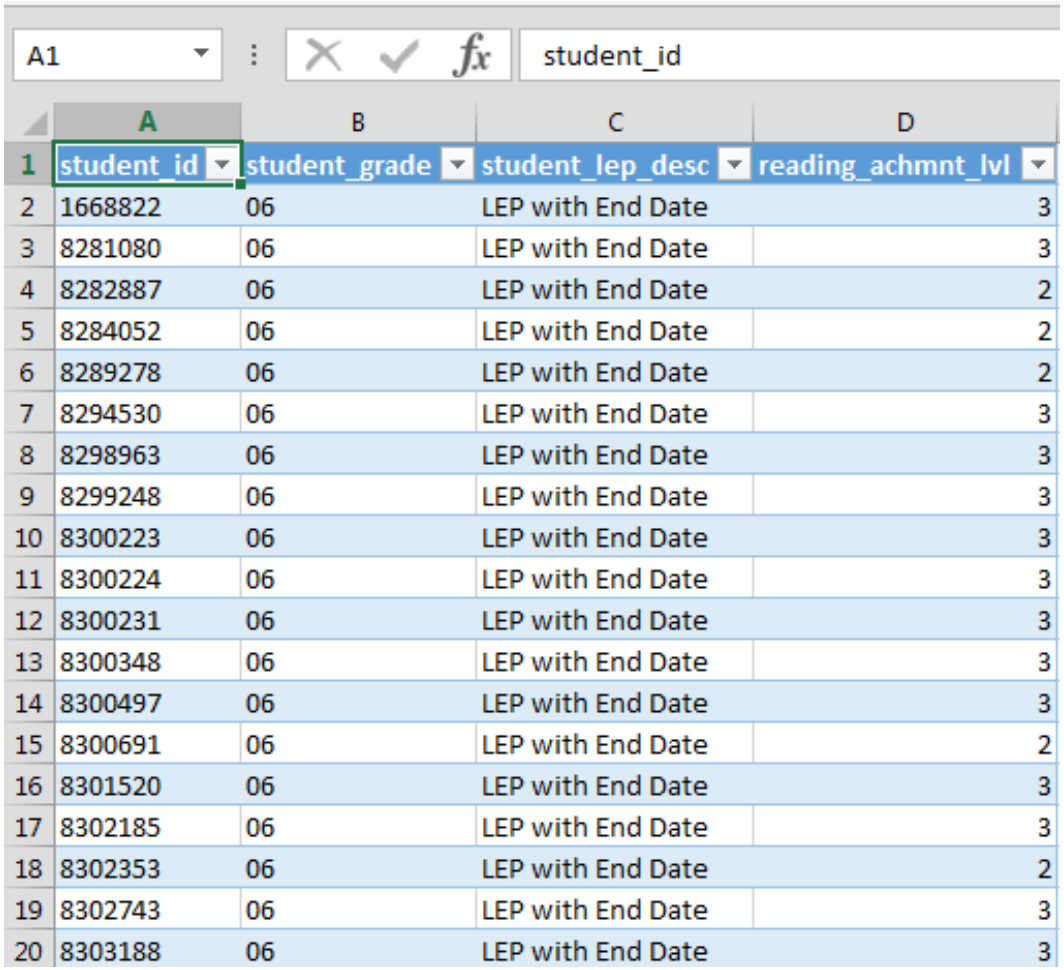
- Import Data window allows user to choose data type and destination

1. Leave 'view' as Table for raw data
2. Leave Existing Worksheet as destination
3. Click OK



Retrieved data in table form

- Resulting data returned to Excel, ready for analysis



A1 | student_id

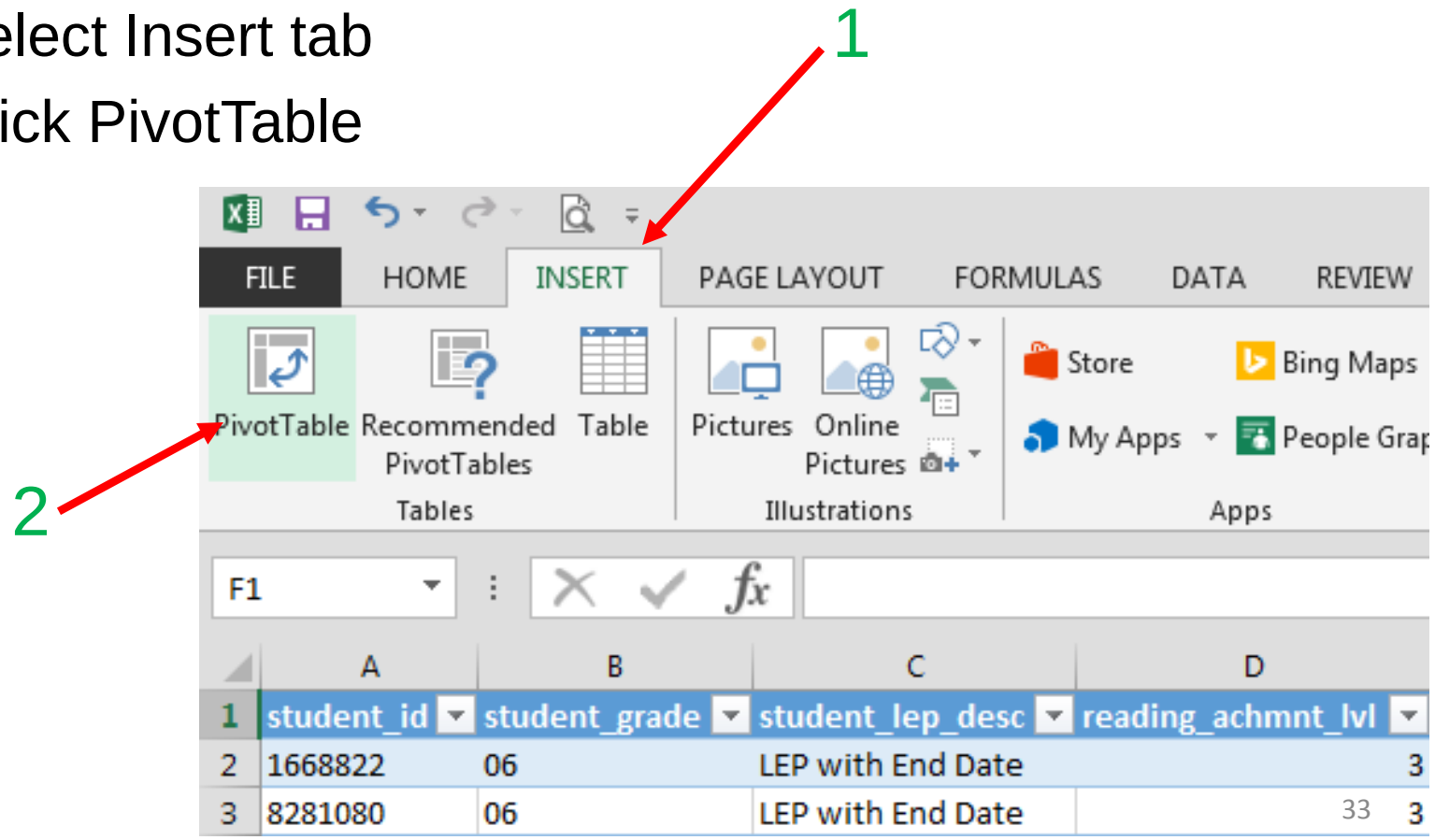
	A	B	C	D
1	student_id	student_grade	student_lep_desc	reading_achmnt_lvl
2	1668822	06	LEP with End Date	3
3	8281080	06	LEP with End Date	3
4	8282887	06	LEP with End Date	2
5	8284052	06	LEP with End Date	2
6	8289278	06	LEP with End Date	2
7	8294530	06	LEP with End Date	3
8	8298963	06	LEP with End Date	3
9	8299248	06	LEP with End Date	3
10	8300223	06	LEP with End Date	3
11	8300224	06	LEP with End Date	3
12	8300231	06	LEP with End Date	3
13	8300348	06	LEP with End Date	3
14	8300497	06	LEP with End Date	3
15	8300691	06	LEP with End Date	2
16	8301520	06	LEP with End Date	3
17	8302185	06	LEP with End Date	3
18	8302353	06	LEP with End Date	2
19	8302743	06	LEP with End Date	3
20	8303188	06	LEP with End Date	3

Summarize data using a connection

- If we want something that is more friendly for leadership, we may want to summarize the raw data
- Data summaries can be updated using database connections. In Excel, PivotTables provide this functionality.

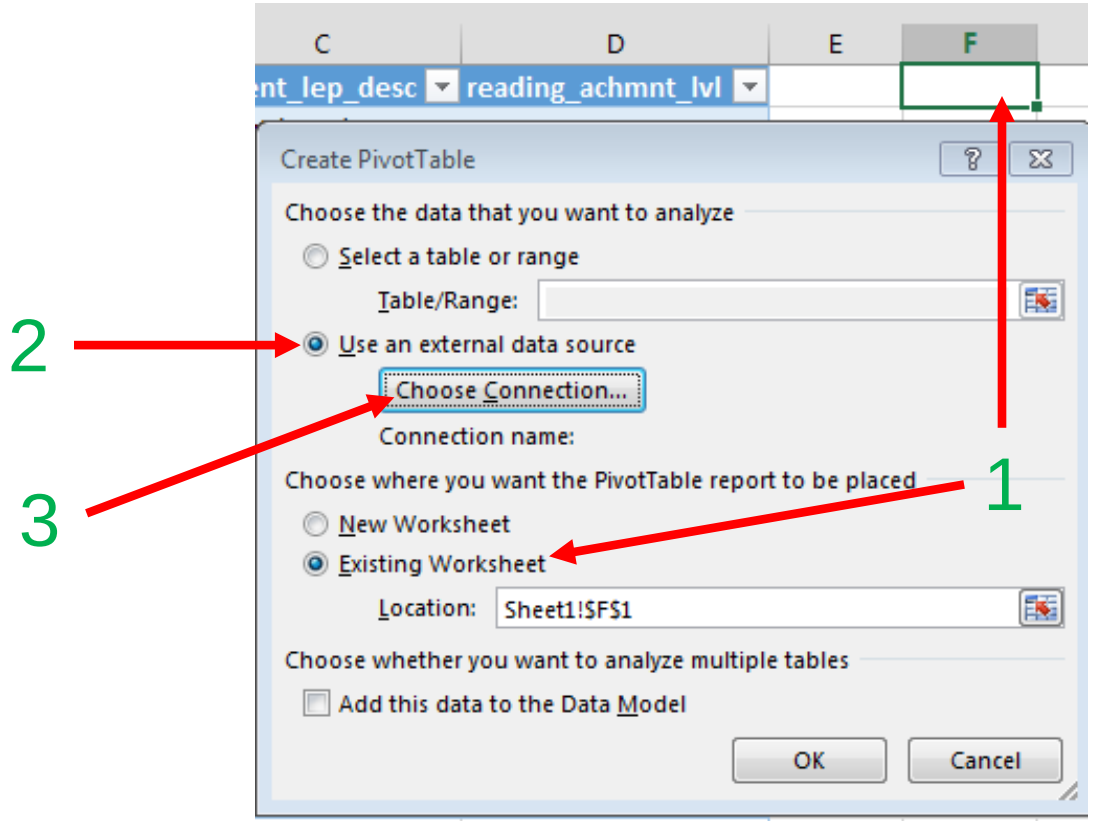
PivotTable – Using a connection

- Let's add a PivotTable to the returned sheet
  - We'll specify reading achievement levels as columns and student LEP status as rows
  1. Select Insert tab
  2. Click PivotTable



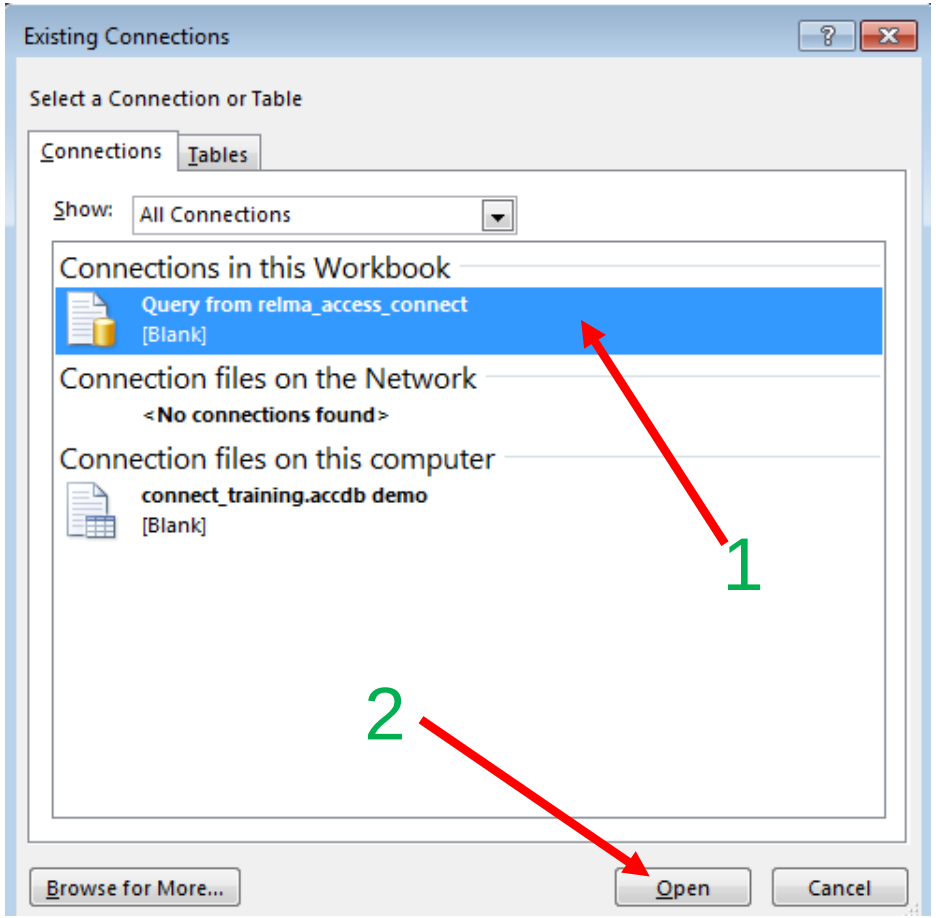
PivotTable – Identify source data & location

1. Select Existing Worksheet and choose cell F1
2. Select ‘Use an external data source
3. Click Choose Connection



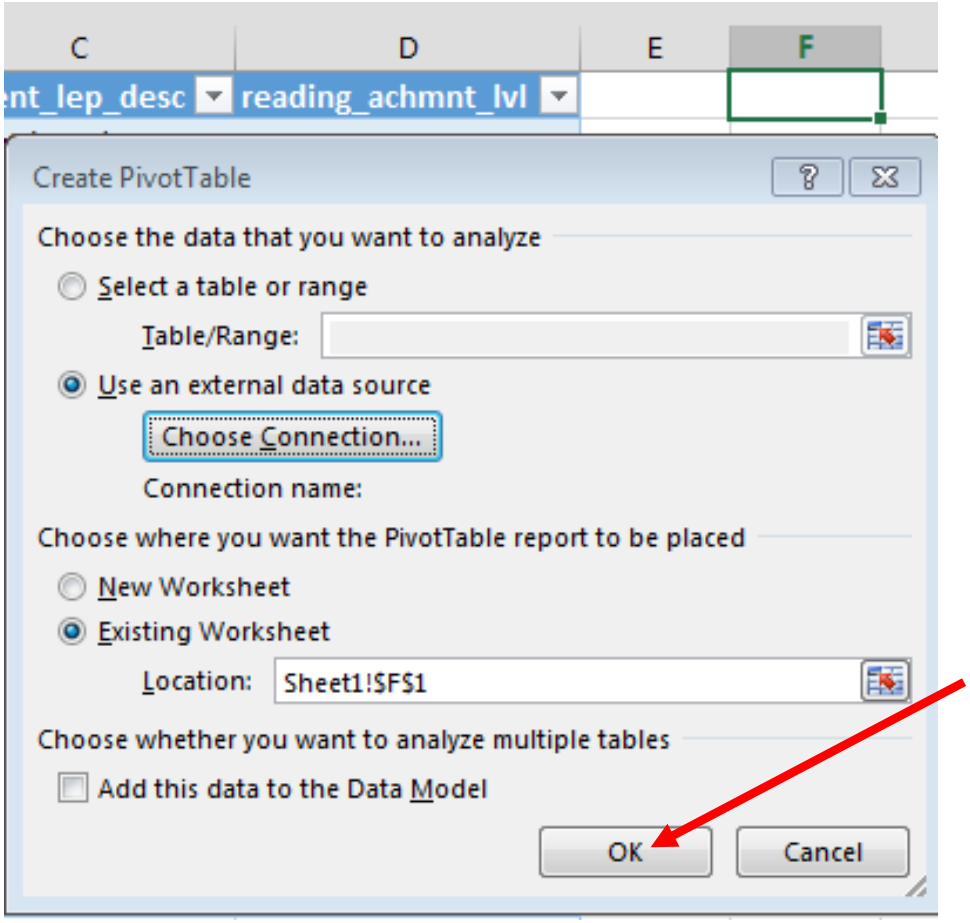
PivotTable – Select connection

1. Select Query from relma_access_connect (this is the query that retrieved our tabular data)
2. Click Open



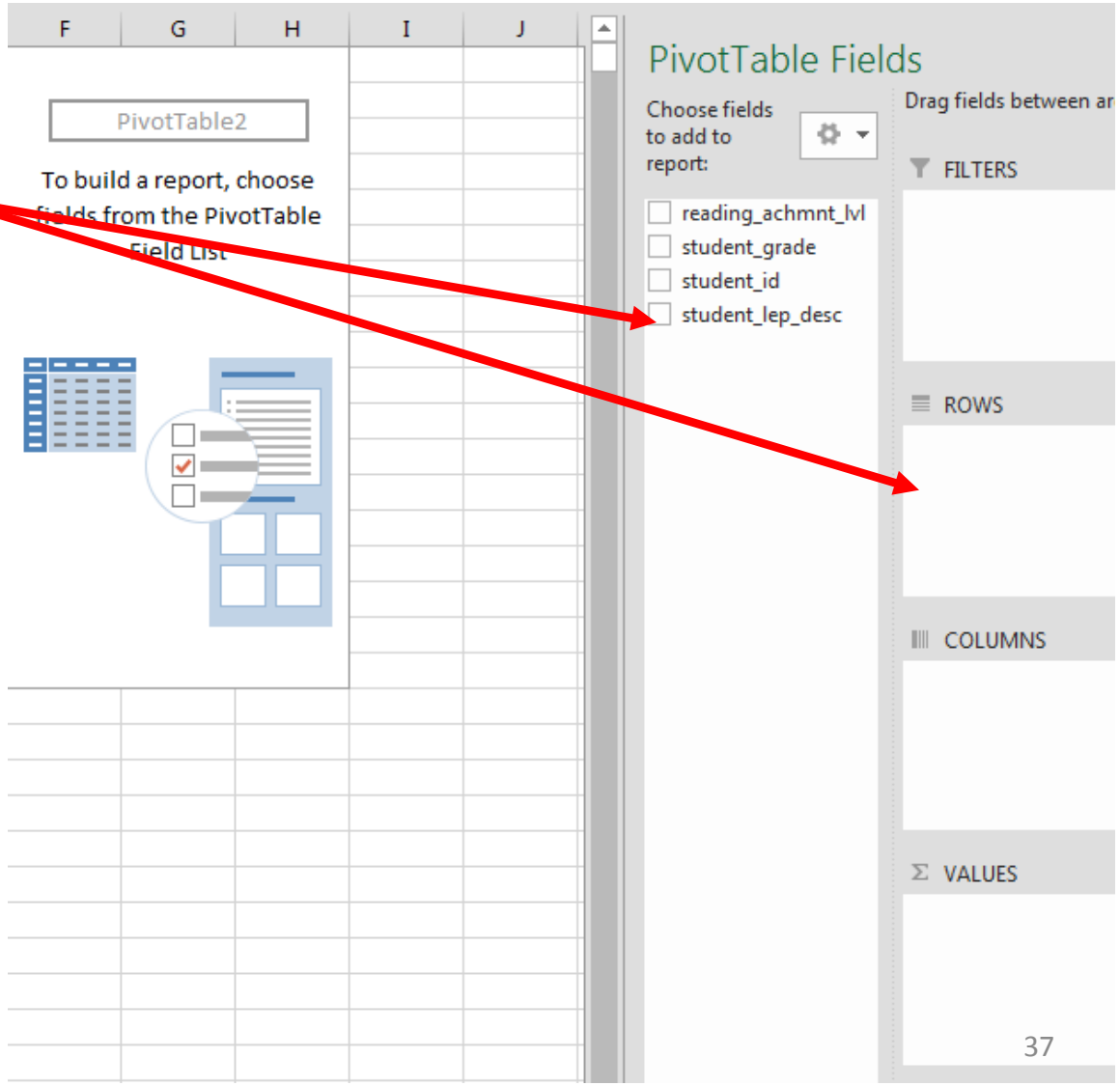
PivotTable – Identify source data & location

- Click OK



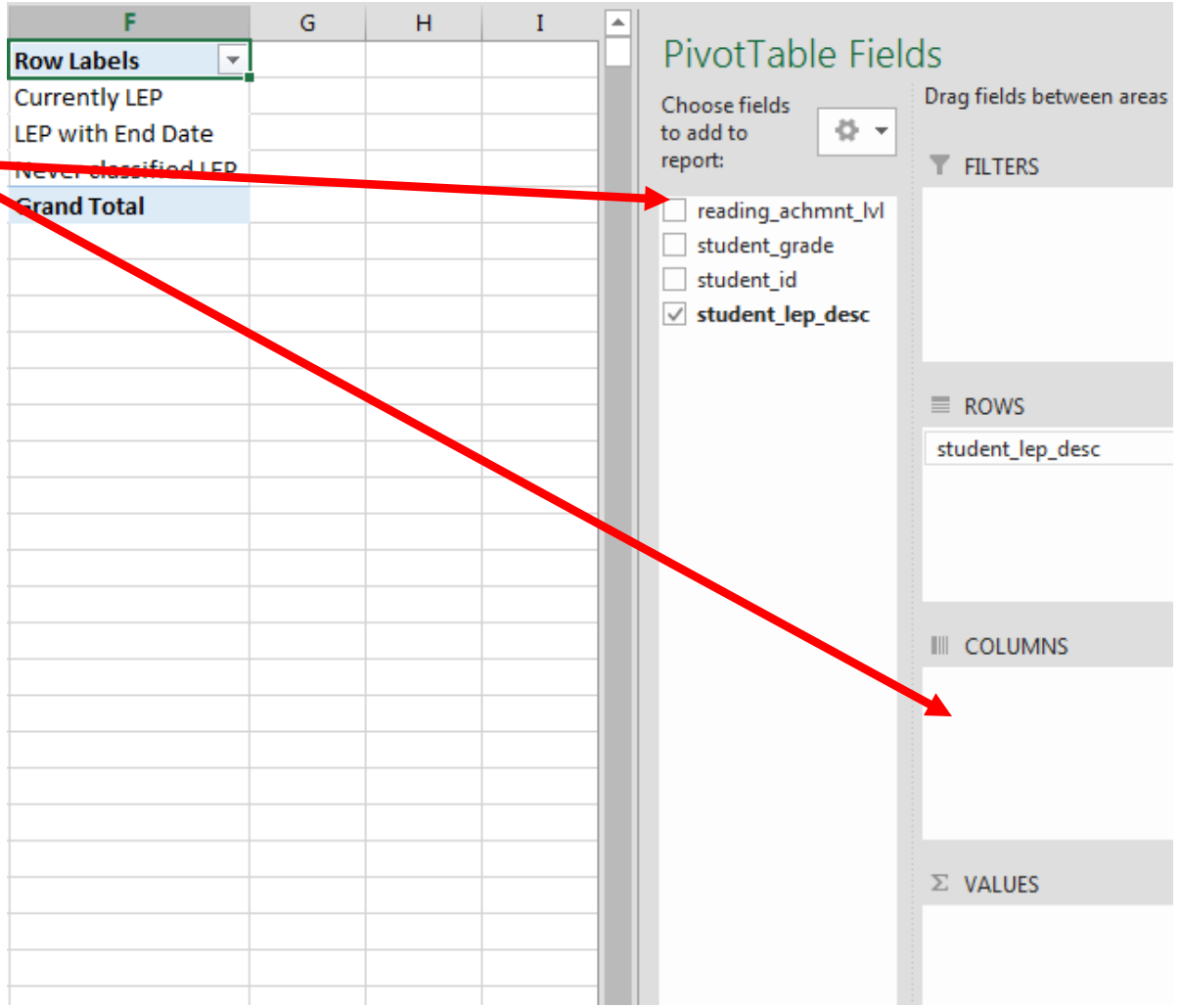
PivotTable – Specify rows of summary table

1 Click, hold and drag student_lep_desc to ROWS box



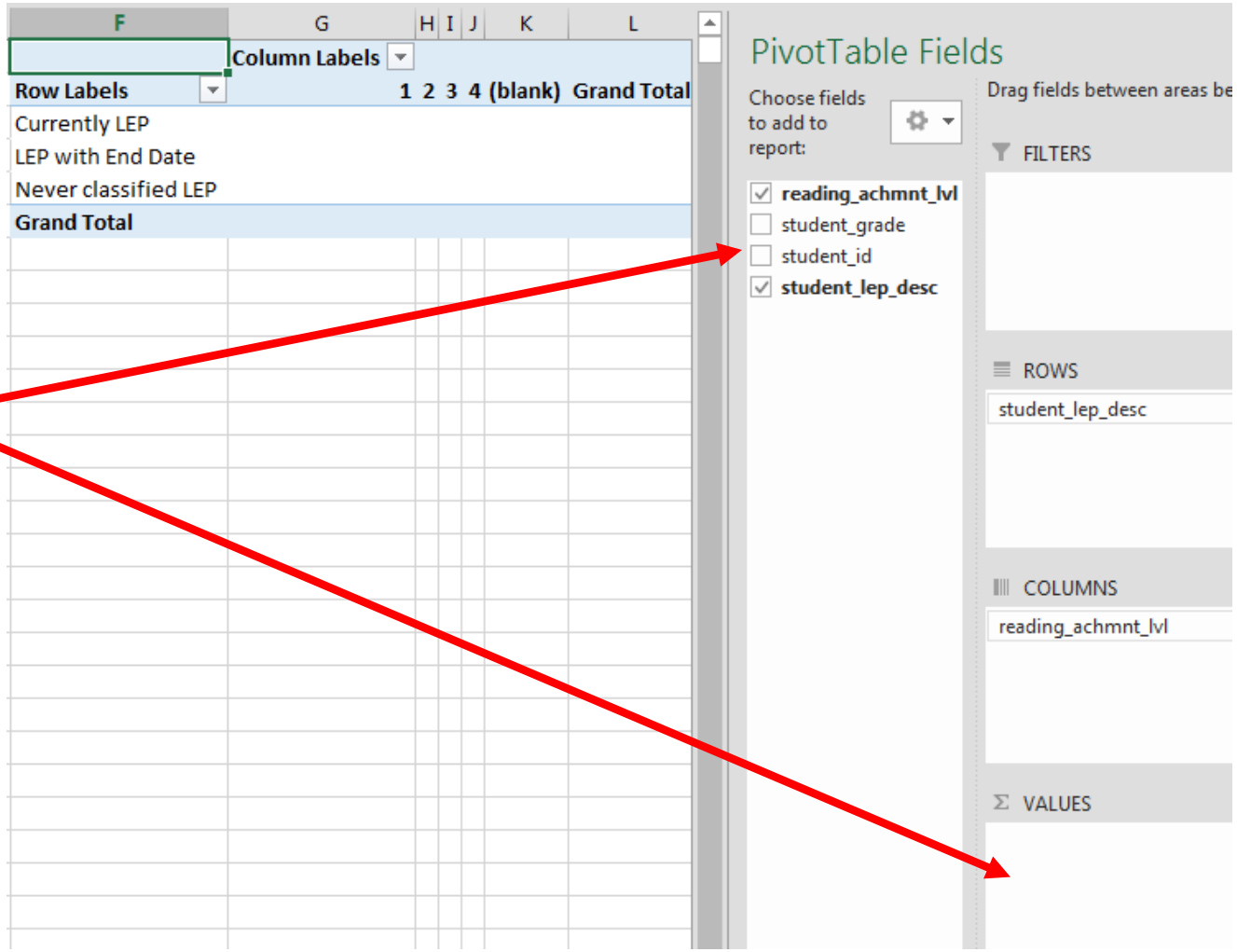
PivotTable - Specify columns of summary table

1 Click, hold and drag reading_achmnt_lvl to COLUMNS box

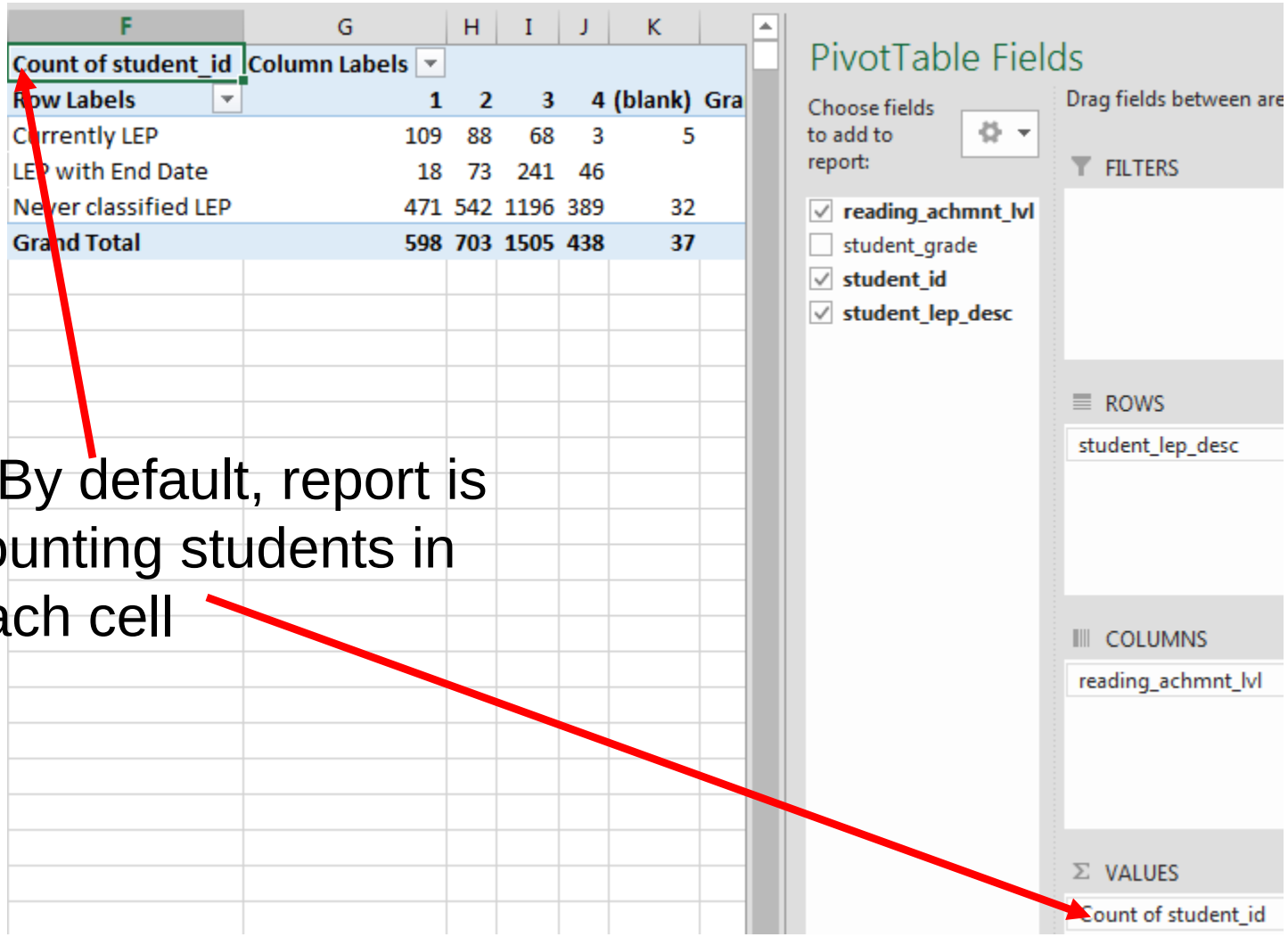


PivotTable - Specify variable to summarize

1 Click, hold and drag student_id to VALUES box

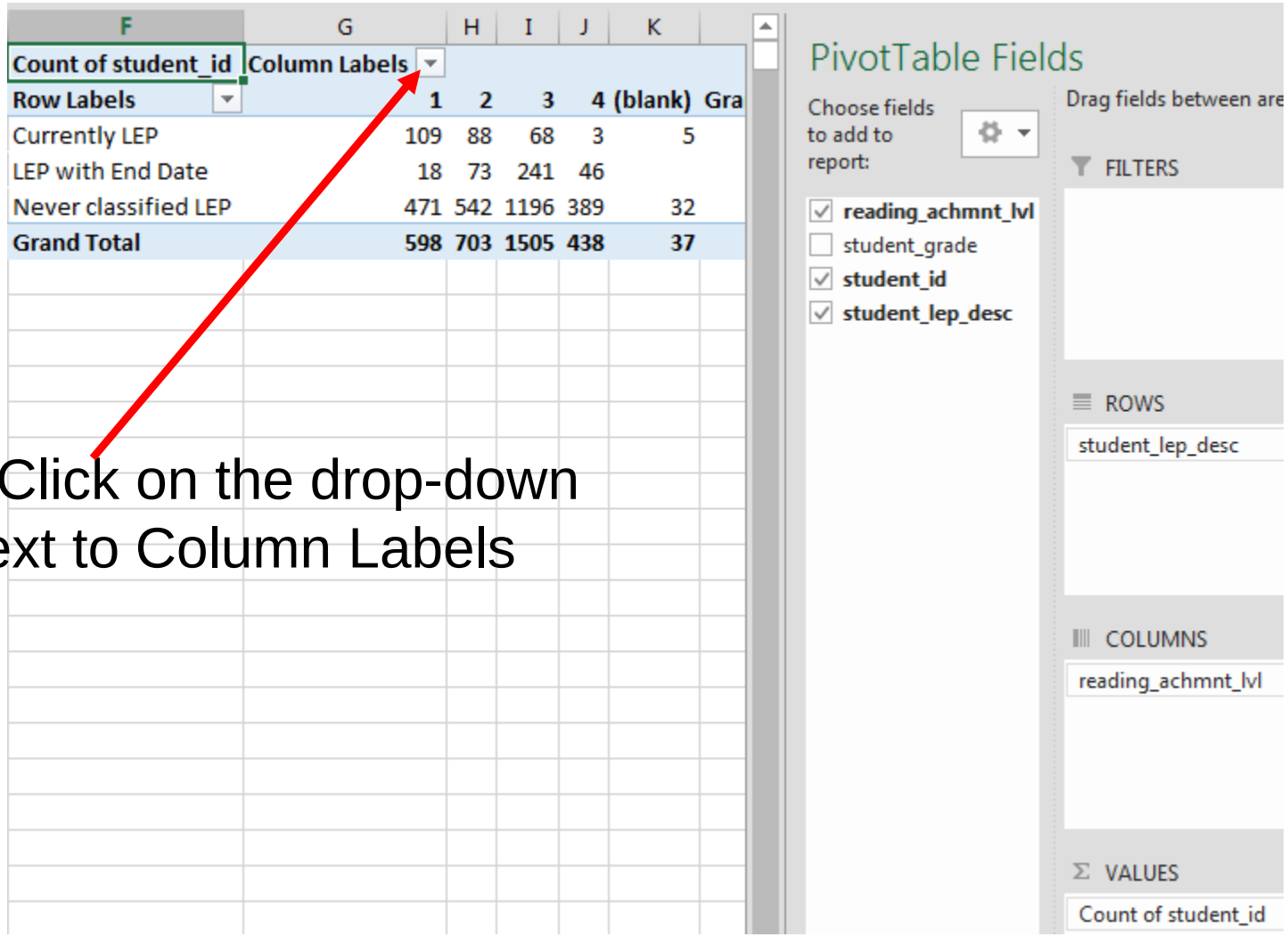


PivotTable – LEP status by achievement level counts



1 By default, report is counting students in each cell

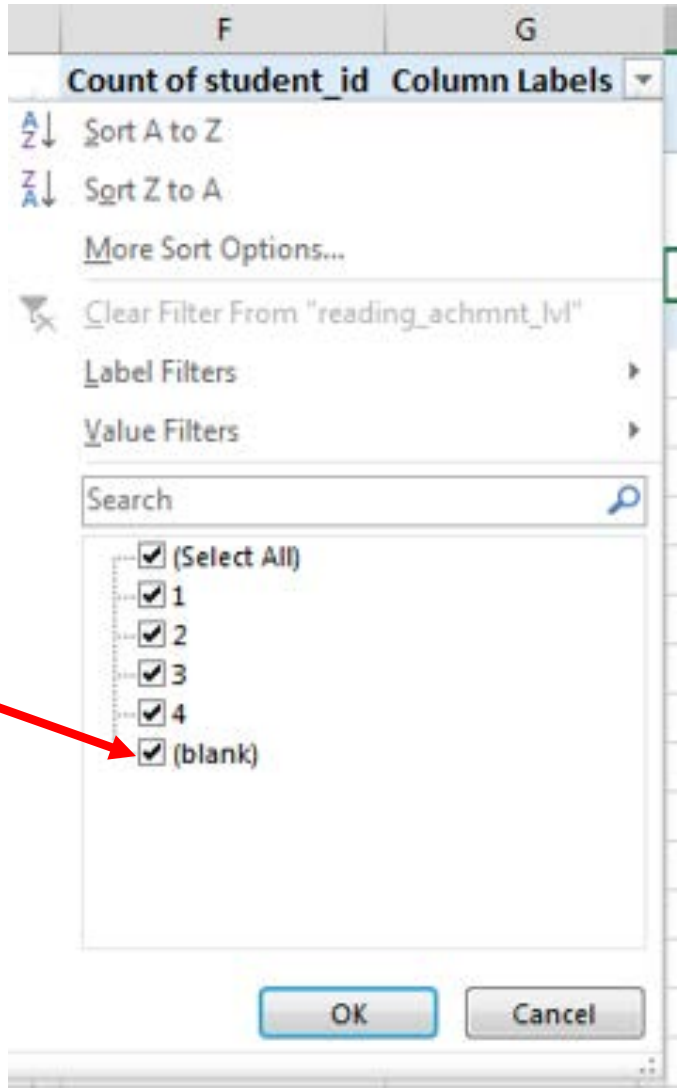
PivotTable – Lets polish the table a bit



1 Click on the drop-down next to Column Labels

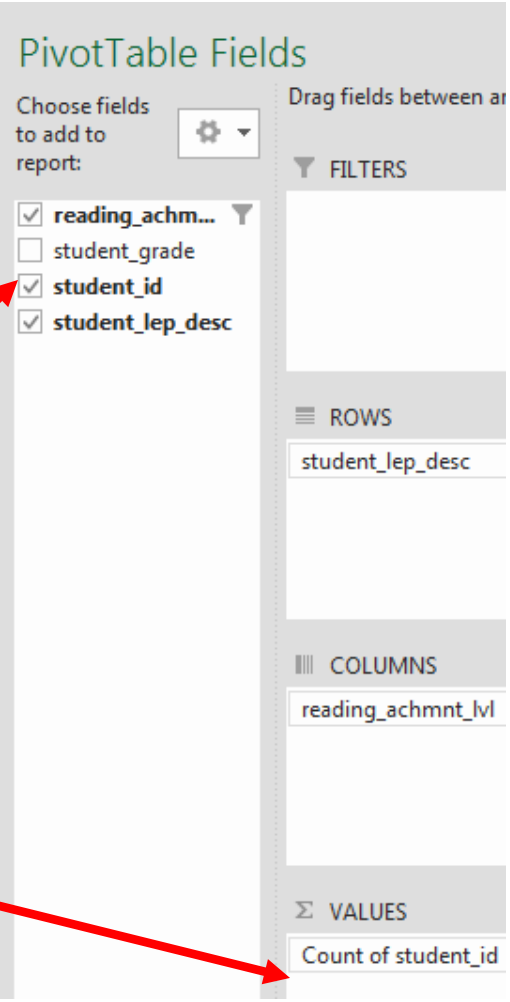
PivotTable – Using column filters

1 To remove the students without reading achievement levels, let's uncheck (blank) from our Column Labels filter



PivotTable – Lets polish a bit more

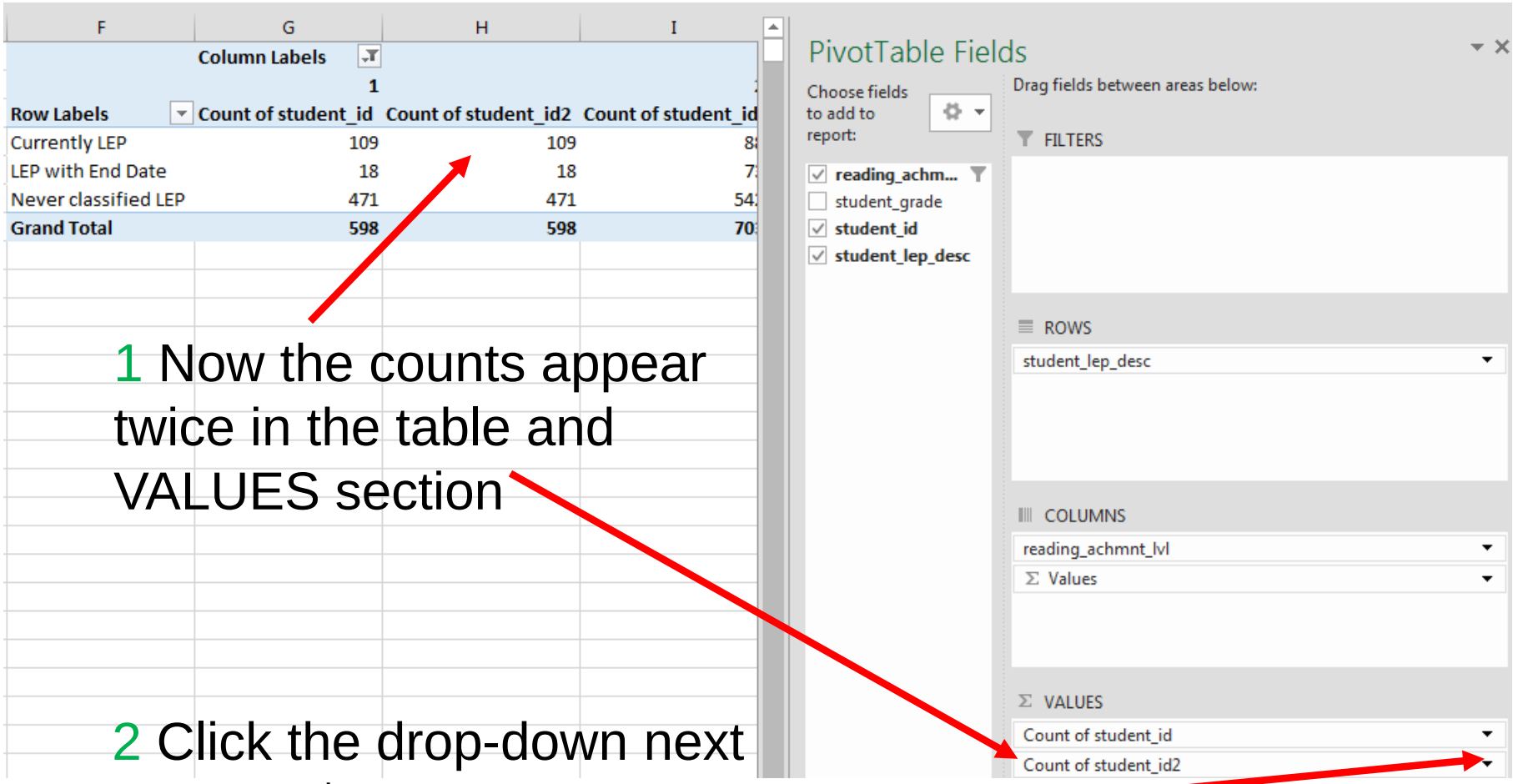
Count of student_id	Column Labels				
Row Labels	1	2	3	4	Grand Total
Currently LEP	109	88	68	3	268
LEP with End Date	18	73	241	46	378
Never classified LEP	471	542	1196	389	2598
Grand Total	598	703	1505	435	3244



1 Blank reading achievement levels have been removed, totals have been updated

2 Click, hold and drag student_id to VALUES box again

PivotTable – Adding more summary variables

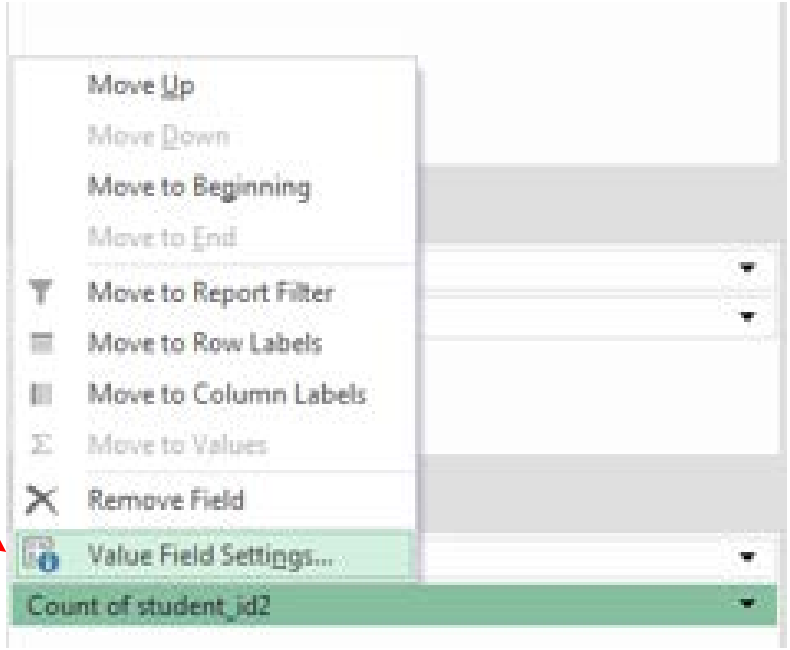


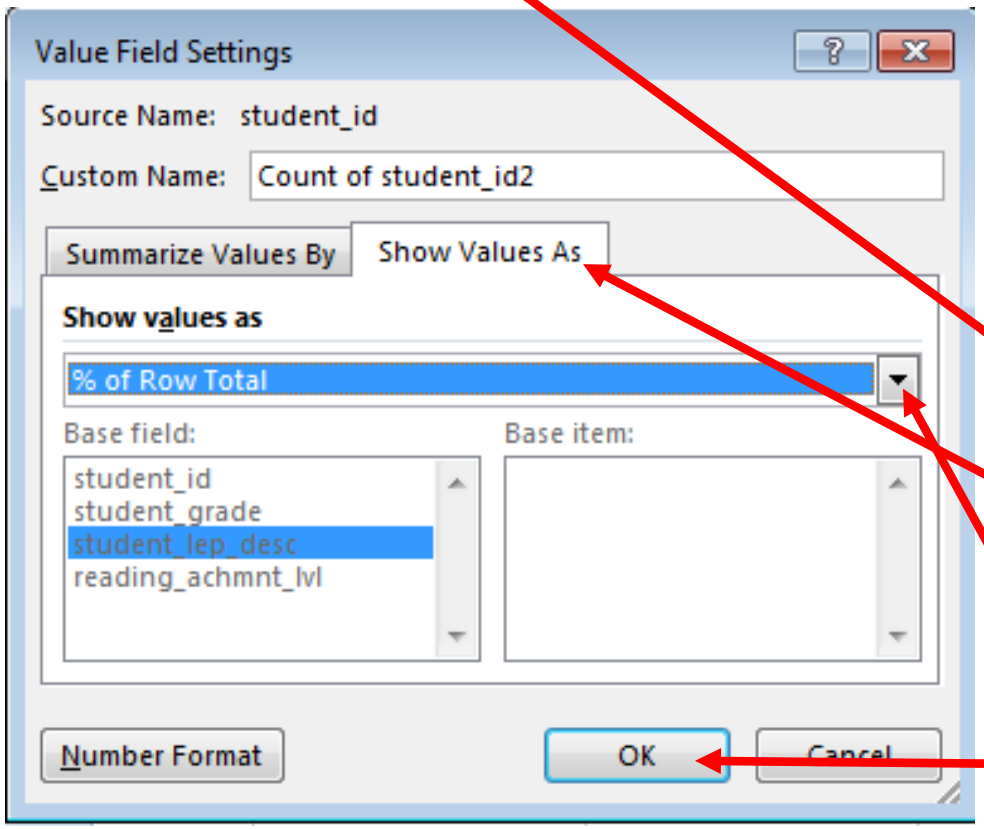
1 Now the counts appear twice in the table and VALUES section

2 Click the drop-down next to our 2nd student_id in VALUES box

PivotTable – Changing how summaries are displayed

1 Click on Value Field Settings…





2 Click Show Values As tab

3 Use drop-down to select % of Row Total

4 Click OK

PivotTable – Editing column labels

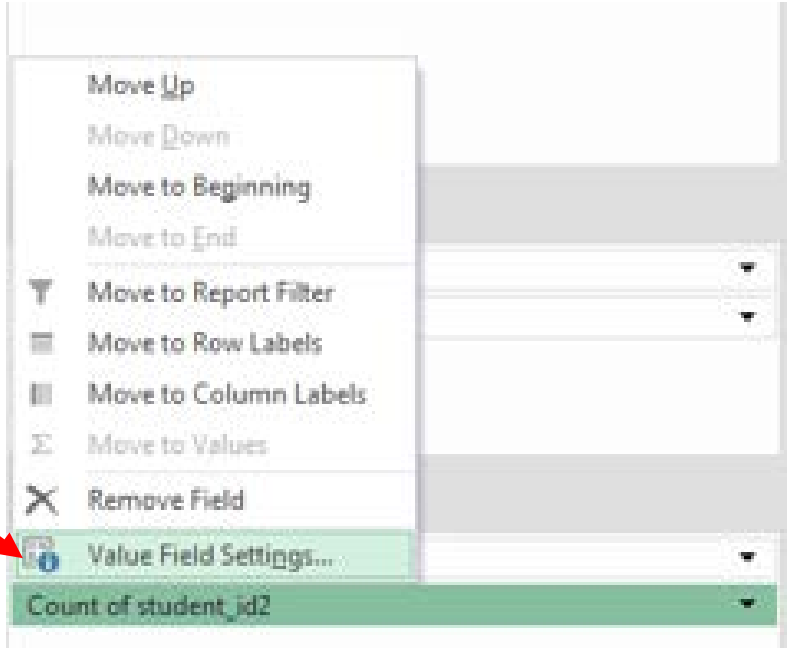
1. Now we have counts and percentages within LEP status
2. Right click on ‘Count of student_id’ in Column G

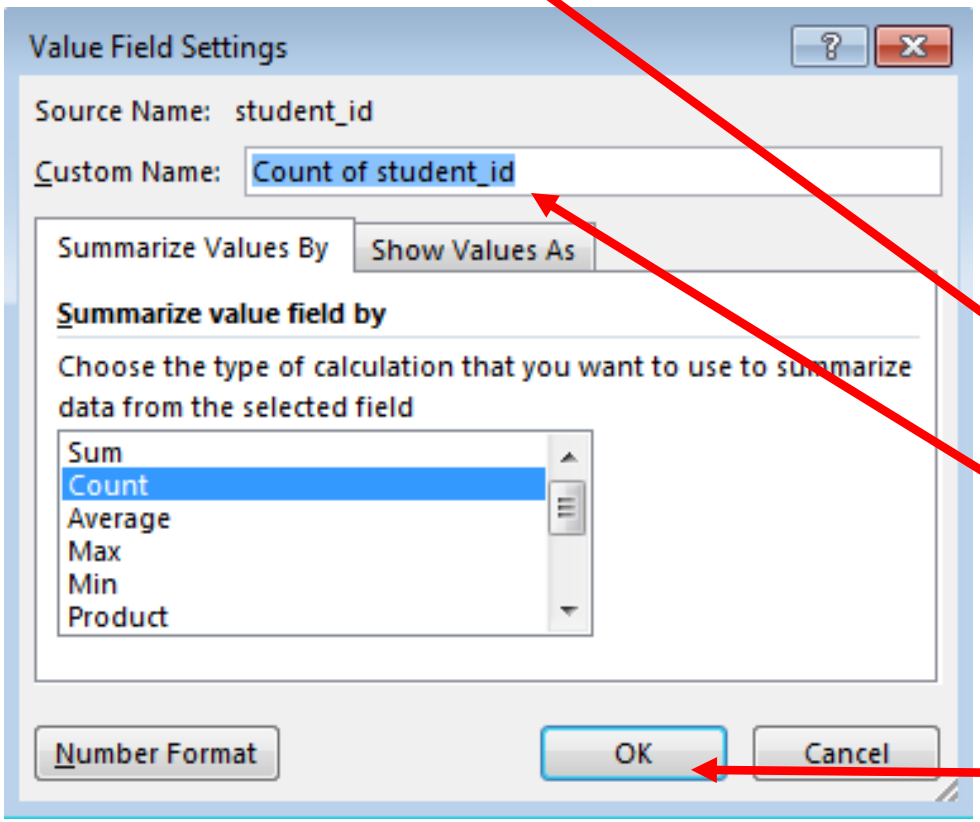
2

F	G	H	I	J
	Column Labels			
	1		2	
Row Labels	Count of student_id	Count of student_id2	Count of student_id	Count of student_id2
Currently LEP	109	40.67%	88	32.84%
LEP with End Date	18	4.76%	73	19.31%
Never classified LEP	471	18.13%	542	20.86%
Grand Total	**598**	**18.43%**	**703**	**21.67%**

PivotTable – Changing column labels

1 Click on Value Field Settings…





2 Change Count of student_id label to ‘n’

3 Click OK

PivotTable – Editing column labels

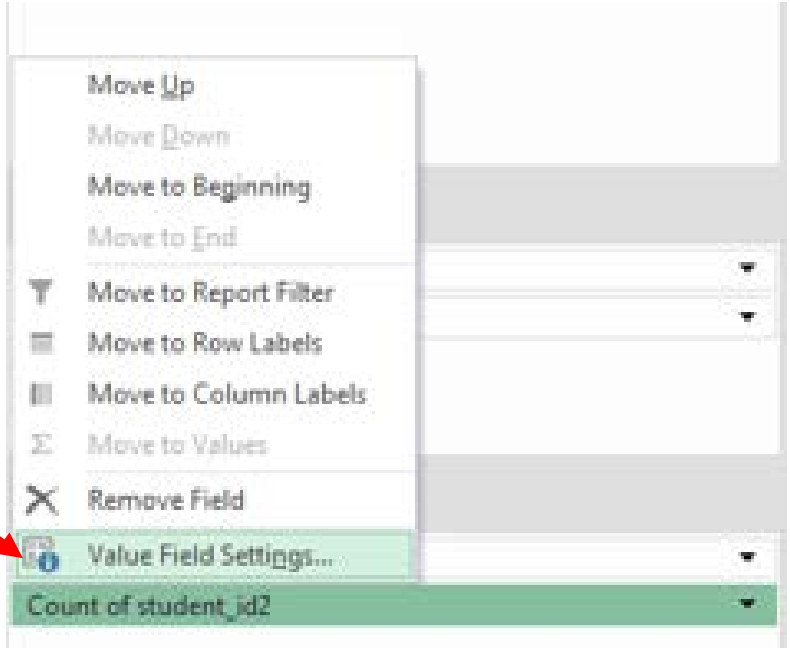
1. Now we have ‘n’ over our counts, let’s add ‘%’ over our percentages
2. Right click on ‘Count of student_id’ in Column H

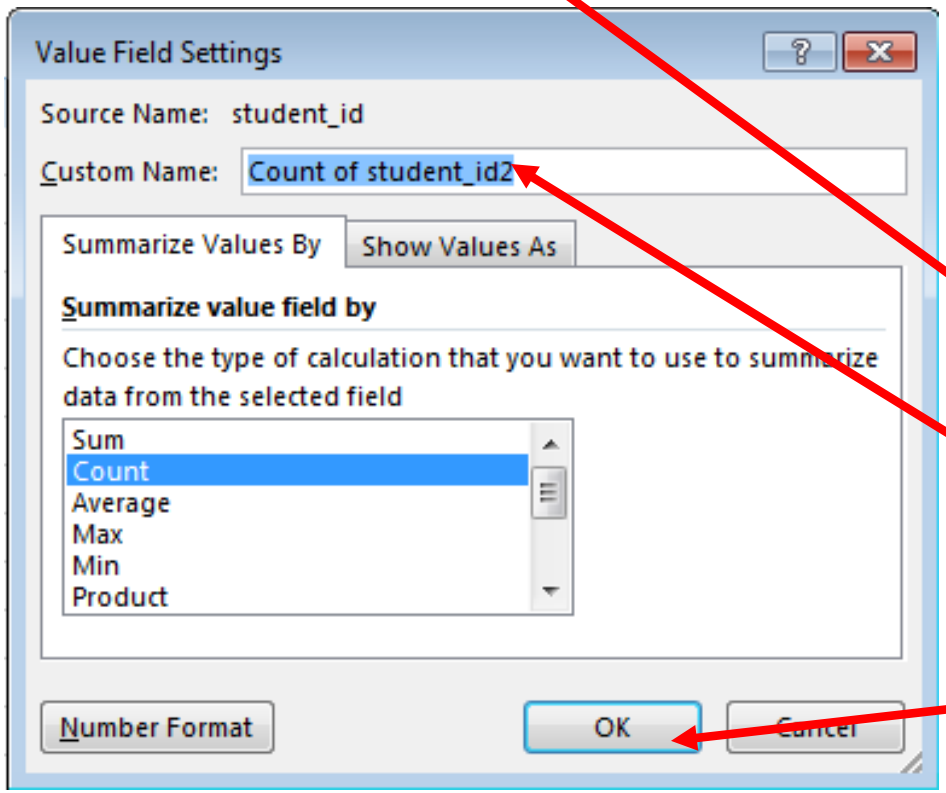
2

F	G	H	I	J
	Column Labels			
	1		2	
Row Labels	n	Count of student_id2	n	Count of student_id2
Currently LEP	109	40.67%	88	32.84%
LEP with End Date	18	4.76%	73	19.31%
Never classified LEP	471	18.13%	542	20.86%
Grand Total	**598**	**18.43%**	**703**	**21.67%**

PivotTable – Changing column labels

1 Click on Value Field Settings…





2 Change Count of student_id2 label to '%'

3 Click OK

PivotTable – Finished product

- Now we have ‘n’ over our counts, and ‘%’ over our percentages
- Adding the percentages quickly reveals for leadership the difference in performance among students of different LEP status
- Should the parameters of the request change, or new data become available, the summary table can quickly be refreshed using the connection

F	G	H	I	J	K	L	M	N	O	P
	Colu									
	1		2		3		4		Total n	Total %
Row Labels	n	%	n	%	n	%	n	%		
Currently LEP	109	40.67%	88	32.84%	68	25.37%	3	1.12%	268	100.00%
LEP with End Date	18	4.76%	73	19.31%	241	63.76%	46	12.17%	378	100.00%
Never classified LEP	471	18.13%	542	20.86%	1196	46.04%	389	14.97%	2598	100.00%
Grand Total	**598**	**18.43%**	**703**	**21.67%**	**1505**	**46.39%**	**438**	**13.50%**	**3244**	**100.00%**

Questions/Need help

Contact:

Jason Schoeneberger, Ph.D.

Senior Researcher and Task Lead

REL Mid-Atlantic at ICF International

jason.schoeneberger@icfi.com

704-307-9395

Please visit www.relmidatlantic.org for other data tools!