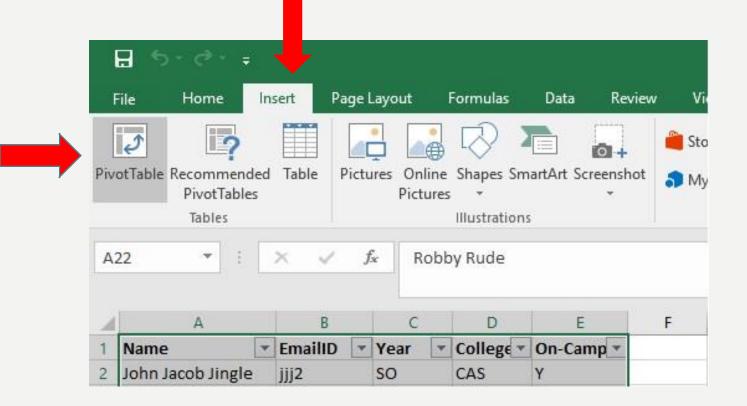
WORKING WITH DATA IN Excel from a student Affairs perspective

ZANE KRATZER DATA RESEARCH SCIENTIST LEHIGH UNIVERSITY OFFICE OF INSTITUTIONAL RESEARCH AND STRATEGIC ANALYTICS

- 'Ctrl-A' to select all relevant fields and rows in a spreadsheet
- Be sure to remove empty columns and empty header rows

1	A		В		С		D	E	F
1	Name	*	EmailID	¥	Year	¥	College •	On-Camp 🕶	
2	John Jacob Jingle		jjj2		so		CAS	Y	
3	Rammy Ramsbot				JR		CAS	Y	
4	Jani Lane		jql6 S		SO		CAS	Y	
5	Roddy Rude		rar6 F		FR		BUS	Y	
6	Sabintha Vali		slv7 i		FR I		BUS	Y	
7	Lani Lemuel		Ivl4 I		FR EN		EN	Y	
8	Baba O'Riley		bor5 S		SO EN		Y		
9	Harry Henderson		heh205		FR EN		EN	Y	
10	Rey Skywalker		ros222		so	SO CAS		Y	
11	Mega Tron		mat56		SO		CAS	Y	
12	Black Widow		bqw207		JR		BUS	N	
13	Pete Seeger		pas209		SO		BUS	Y	
14	Rory Rockerbock		ryr219		SO		EN	Y	
15	Fred Flinstone		fnf218		SR		IC	N	
16	Dorothy Dale		dtd217		FR		CAS	Y	
17	Suzy Q.		sqq5		SR		IC	N	
18	Betsy Black		blb218		JR		CAS	N	
19	Grimmy Reapa		gyr217		SO		EN	Y	
20	Jasper Jiver		jjj219		SO		IC	Y	
21	Rory Rockerbock		ryr216		SO		EN	Y	
22	Robby Rude		rar6		FR		BUS	Y	
23									

• Insert > Pivot Table



- Choose between 'New Worksheet' and 'Existing Worksheet'
- This tells Excel where to paste your new Pivot Table, on the current worksheet where your data resides or in a new bland worksheet.
- Also, you can check the table range here to make sure that Excel is grabbing the correct rows and columns.

Create PivotTable		? ×
Choose the data that	you want to analyze	
Select a table or	range	-
Table/Range:	Signin_Data!SAS1:SES22	1
O Use an external of	data source	
Choose Co	nnection	
Connection r	name:	-
🔘 Use this workbo	ok's Data Model	
Choose where you wa	ant the PivotTable report to be placed	
Choose where you wa		
• <u>N</u> ew Worksheet		
New Worksheet Existing Worksheet Location:	eet	
New Worksheet Existing Worksheet Location:	eet want to analyze multiple tables	
New Worksheet Existing Worksheet Location:	eet want to analyze multiple tables	

• Values: The field to be summarized (i.e. EmailID provides a Count of all EmailID values)

4	A	В
1		
2		
3	Count of EmailID	
4	21	
5		
6		
7		
8		

T FILTERS	
≡ ROWS	Σ VALUES
	Count of EmailID

• To change how the sum works, click on the dropdown arrow and select "Value Field Settings"

Drag fields between areas below: T FILTERS	Move Up Move Down Move to Beginning Move to End T Move to Report Filter
	Move to Row Labels Move to Column Labels Move to Values Remove Field
ROWS	Value Field Settings
	Count of EmailID

- Change the "Summarized value field by" to Sum, Count, Average, Max, Min, etc.
- Note that if your selected field is a text column then numeric options such as Sum and Average will not work

Custom Name: Count	of EmailID	
Summarize Values By	Show Values As	
Summarize value field	by	
	culation that you want to use to s	ummarize
data from the selected	field	
Sum Count	~	
Average		
Max		
Min		
Product		
Count Numbers StdDev		
StdDevp		
Var		
Varp	~	

 To show the summed values as a percentage or other form, click on the "Show Value As" tab and select an option such as "% of Column Total"

4	A	В	С
1			
2			
3	Row Labels	Count of EmailID	
4	FR	28.57%	
5	JR	14.29%	
6	so	47.62%	
7	SR	9.52%	
8	Grand Total	100.00%	
9			

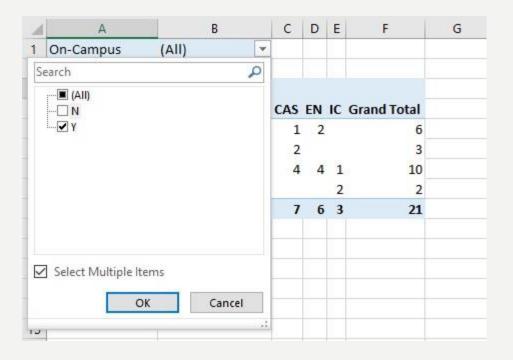
Source Name: EmailID		
Custom Name: Count of EmailID		
Summarize Values By Show Values A	As	
Show values as		
% of Column Total		~
No Calculation % of Grand Total % of Column Total		^
% of Row Total % Of % of Parent Row Total		Ų
On-Campus		

• Rows and Columns: Fields to be inserted into either the Row or Column of the pivot table

1	A	В	С	D	Е	F	G
1							
2							
3	Count of EmailII	Column Labels 💌					
4	Row Labels	BUS	CAS	EN	IC	Grand Total	
5	FR	3	1	2		6	
6	JR	1	2			3	
7	SO	1	4	4	1	10	
8	SR				2	2	
9	Grand Total	5	7	6	3	21	
10							
11							

▼ FILTERS		
	College	
Rows	Σ VALUES	
Year	Count of EmailID	8

• Filters: Fields to be used to filter out specific values from the data



T FILTERS		III COLUMNS	
On-Campus	-	College	Ŧ
ROWS		Σ VALUES	
Year	•	Count of EmailID	

- Counting Duplicate Records by Formula
 =IF(COUNTIF(\$B\$2:B2,B2)=1,COUNTIF(B:B,B2),"")
- Replace 'B' with whatever column your target field is in
- If your data does not begin on row 2, replace '2' with whatever row your data begins on
 A2
 Image: Source state of the stat

1	A	A B		C		D		E	F	G
1	DUPS	Name	•	EmailID	٣	Year	*	College 🔻	On-Camp 🔻	
2	1	John Jacob Jingle j		jjj2		SO		CAS	Y	
3		Rammy Ramsbot		rar5		JR		CAS	Y	
4		Jani Lane		jql6		SO		CAS	Y	
5		Roddy Rude		rar6		FR		BUS	Y	
6		Sabintha Vali		slv7		FR		BUS	Y	
7		Lani Lemuel		lvl4		FR		EN	Y	
8		Baba O'Riley		bor5		SO		EN	Y	
9		Harry Henderson		heh205		FR		EN	Y	

- Finding Duplicate Records with Conditional Formatting
- 'Ctrl-A' to select all relevant fields and rows in a spreadsheet
- Conditional Formatting > Highlight Cells Rules > Duplicate Values

General	*	N 🛃 🛃 🚺	Iormal	Bad	Good
\$ • %	00, 0. →		heck Cell	Explanatory	Input
Num	ber 5	Highlight Cells Rul		Greater Than	
		Top/Bottom Rules	s →	Less Than	
1	J	Data Bars	•	Between	Q
		Color Scales	•	Equal To	
		Icon Sets	•	Text that Contains.	
		New Rule Image: Clear Rules	,	A Date Occurring	
		Manage <u>R</u> ules		Duplicate Values	

• Check default settings on Duplicate Values dialog box on select

10	A	В	C	D	E	F	G	н	1	J	K
1	DUPS	Name	EmailID	Year	▼ College	 On-Camp 					
2	1	John Jacob Jingle	jjj2	so	CAS	Y					
3	1	Rammy Ramsbot	rar5	JR	CAS	Y					
4	1	Jani Lane	jql6	SO	CAS	Y					
5	2	Roddy Rude	rar6	FR	BUS	Y					
6	1	Sabintha Vali	slv7	FR	BUS	Y	-				
7	1	Lani Lemuel	lvl4	FR	EN	Y					
8	1	Baba O'Riley	bor5	SO	EN	Y					
9	1	Harry Henderson	heh205	FR	EN	Y Duplic	ate Values			?	
10	1	Rey Skywalker	ros222	SO	CAS	Y	Format cells that contain:				
11	1	Mega Tron	mat56	SO	CAS	Y					
12	1	Black Widow	bqw207	JR	BUS	N Duplic	ate 🗸 🗸 val	ues with Li	ght Red Fill v	vith Dark Red	d Tex
13		Pete Seeger	pas209	SO	BUS	Y			-		
	1.	Rory Rockerbock	ryr219	so	EN	Y			O	К	Can
14			E E	00	IC					-	1
15		Fred Flinstone	fnf218	SR		N					
15 16	1	Dorothy Dale	dtd217	FR	CAS	Y					
15 16 17	1	Dorothy Dale Suzy Q.	dtd217 sqq5	FR SR	CAS IC	Y N					
15 16 17 18	1 1 1	Dorothy Dale Suzy Q. Betsy Black	dtd217 sqq5 blb218	FR SR JR	CAS IC CAS	Y N N					
15 16 17 18 19	1 1 1 1	Dorothy Dale Suzy Q. Betsy Black Grimmy Reapa	dtd217 sqq5 blb218 gyr217	FR SR JR SO	CAS IC CAS EN	Y N N Y					
15 16 17 18 19 20	1 1 1 1 1	Dorothy Dale Suzy Q. Betsy Black Grimmy Reapa Jasper Jiver	dtd217 sqq5 blb218 gyr217 jjj219	FR SR JR SO SO	CAS IC CAS EN IC	Y N N Y Y					
15 16 17 18 19	1 1 1 1 1	Dorothy Dale Suzy Q. Betsy Black Grimmy Reapa	dtd217 sqq5 blb218 gyr217	FR SR JR SO	CAS IC CAS EN	Y N N Y					

To remove formatting, go to Conditional Formatting > Clear Rules
 > Clear Rules from Entire Sheet

General	•		Normal		Bad		Good	
\$ - %	• €.0 .00 .00 →.0		Check Ce	211	Explanat	ory	Input	
Number 🕞		Highlight Cells Ro	Styles					
		Top/Bottom Rule	es ▶					
1	J	Data Bars	•	N	0	P	Q	
		Color <u>S</u> cales	Þ					
		Icon Sets	×					
		🔝 New Rule						
		Clear Rules	*	Clear Rules from Selected Cells				
		Manage <u>R</u> ules	Clear Rules from Entire Sheet					
				Clear Rules from <u>T</u> his Table			ible	
				Clear Rules from This PivotTable			votTable	

MERGING MULTIPLE TABLES

Tools for merging tables, pulling data from one table into another:

• VLOOKUP

https://exceljet.net/excel-functions/excel-vlookup-function

• INDEX MATCH

https://exceljet.net/index-and-match

• Power Query

https://www.howtoexcel.org/power-query/the-complete-guide-to-power-query/

COMBINING YEARS/MULTIPLE EVENTS

- What is the unit of analysis?
- If the student record must be unique, then the data should be in a wide format, with each row representing a unique student and all subsequent fields or measurements for the student added as extra columns.
- If the unit of analysis is the point of measurement for each student (a specific term, date, event), then the data should be in a long format, with each row representing a unique data entry for each student. Students will have duplicate rows if data is recorded for them more than once (i.e. for every event or semester attended).

OIRSA CONTACT INFO

For general OIRSA contact and information: Website: oirsa.lehigh.edu Email: oir@lehigh.edu Phone: 610-758-5890

To request our services: oirsa.lehigh.edu/make-request

For specific questions regarding this presentation: Zane Kratzer, Data Research Scientist Email: <u>zlk205@lehigh.edu</u> Phone number: 610-758-5918