

MIS2502:
Data Analytics
ETL - Excel Basics

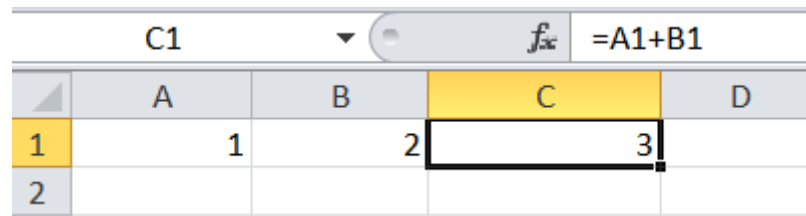
Alvin Zuyin Zheng

zheng@temple.edu

<http://community.mis.temple.edu/zuyinzheng/>

Excel References

- Many Excel formulas refer to cells or ranges of cells.
 - For example, the simple formula =A1+B1 refers to cells A1 and B1



The image shows a screenshot of an Excel spreadsheet. The active cell is C1, which contains the formula =A1+B1. The spreadsheet has columns A, B, C, and D, and rows 1 and 2. Cell A1 contains the value 1, and cell B1 contains the value 2. Cell C1 is highlighted in yellow and contains the value 3. The formula bar at the top shows the formula =A1+B1.

	A	B	C	D
1	1	2	3	
2				

- There are two types of cell references: **relative** and **absolute**.
- Relative and absolute references behave differently when copied and filled to other cells.

Relative Reference

- When copied across multiple cells, they change based on the relative position of rows and columns.
- By default, all cell references are **relative references**.
- Example:

	A	B	C	D
1	=E1			
2				
3				




	A	B	C	D
1	=E1	=F1	=G1	
2	=E2	=F2	=G2	
3	=E3	=F3	=G3	

Absolute Reference

- When copied across multiple cells, cell references remain constant
- Example:

	A	B	C	D
1	= $\$E\1			
2				
3				




	A	B	C	D
1	= $\$E\1	= $\$E\1	= $\$E\1	
2	= $\$E\1	= $\$E\1	= $\$E\1	
3	= $\$E\1	= $\$E\1	= $\$E\1	

Mixing Relative and Absolute References

- Using absolute referencing for the column and relative referencing for the row:


	A	B	C	D
1	=\$E1			
2				
3				



	A	B	C	D
1	=\$E1	=\$E1	=\$E1	
2	=\$E2	=\$E2	=\$E2	
3	=\$E3	=\$E3	=\$E3	

- Using relative referencing for the column and absolute referencing for the row:

	A	B	C	D
1	=E\$1			
2				
3				



	A	B	C	D
1	=E\$1	=F\$1	=G\$1	
2	=E\$1	=F\$1	=G\$1	
3	=E\$1	=F\$1	=G\$1	

Basic Functions

LEN()

- Return the number of characters in a text string

LEFT(), RIGHT(),
MID()

- Return a specific number of characters from a text string

CONCATENATE()

- Join two text strings together

IF()

- Check if a condition is met

LEN()

- =LEN(value)
 - returns the number of characters contained in a string value
- Example...
 - LEN(123) and LEN("DOG") both return 3.

Alvin	=LEN(B3)	5
Alvin Zheng	=LEN(B4)	11

LEN(text)

LEFT(), RIGHT() or MID()

=LEFT(value, n)

- Returns n characters from the **start** of a string value
- Example: LEFT("HELLO", 2) will return "HE".

=RIGHT(value, n)

- Returns n characters from the **end** of a string value
- Example: RIGHT("HELLO", 2) will return "LO".

=MID(value, start_index, n)

- Returns n characters from the **middle** of the text string given the index of first character to start at
- Example: MID("HELLO",2,3) will return "ELL"

CONCATENATE()

- =CONCATENATE(value1, value2...)
 - Combines two or more string values or data in cells
- Example...
 - CONCATENATE(A2, “, HELLO”) will append the string “, HELLO” to the end of whatever is in cell A2. Like this:

	A	B
1	Name	NewCell
2	Bob	Bob, HELLO
3	Jack	Jack, HELLO
4	Sue	Sue, HELLO
5	Janet	Janet, HELLO

IF()

- =IF(condition, value_if_true, value_if_false)
 - condition = conditional statement
(e.g., A2 > 3, C5 = "text")
 - value_if_true = value to return if condition is met
 - value_if_false = value to return if condition is not met
- Example...
 - IF(4 > 5, "red", "yellow") would give you a result of *yellow*

VLOOKUP()

- Match values and get associated fields
- Similar to SQL joins
- =VLOOKUP(lookup_value, table_array, column_index, range_lookup)
 - lookup_value = value that you're looking for
 - table_array = the table where you're going to do your search (e.g., A2:E5)
 - column_index = column number to return from matched record
 - range_lookup = TRUE for approximate matches and FALSE for exact matches

VLOOKUP()

- =VLOOKUP(lookup_value, table_array, column_index, range_lookup)
- Example

Hourly Pay

	A	B
1	Employee Name	Hourly Rate
2	Atkins, James	\$35.50
3	Benn, Carol	\$25.00
4	Benson, Paul	\$32.00
5	Cooper, David	\$28.50
6	Daley, Ann	\$41.00
7	Dawson, Helen	\$32.00
8	⋮	⋮

Sales Team Hours Worked with Vlookup Functions:

	A	B	C	D	E
1	Employee Name	Hours Worked	Hourly Rate	Payment Due	
2	Benson, Paul	37.5	=VLOOKUP(A2, 'Hourly Pay'!A:B, 2, FALSE)		
3	Cooper, David	40	=VLOOKUP(A3, 'Hourly Pay'!A:B, 2, FALSE)		
4	Dawson, Helen	39	=VLOOKUP(A4, 'Hourly Pay'!A:B, 2, FALSE)		
5	⋮	⋮	⋮		

Sales Team Hours Worked with Vlookup Results:

	A	B	C	D
1	Employee Name	Hours Worked	Hourly Rate	Payment Due
2	Benson, Paul	37.5	\$32.00	
3	Cooper, David	40	\$28.50	
4	Dawson, Helen	39	\$32.00	
5	⋮	⋮	⋮	