

Data types for data science

INTERMEDIATE SPREADSHEETS



Richie Cotton

Data Evangelist at DataCamp

Calculating the mean, badly

	A	B	C	D
1	49.53669111			
2	AVERAGE(A1:A5)			%+2
3	SUM(A1:A5)			Enter
4	Press Tab then ↑/↓ to select suggestions			X
5	15.57589729			
6	=			

What went wrong?

	A
1	49.53669111
2	7.42453201
3	63.47273292
4	28.37423 I 27
5	15.57589729
6	34.00246333

Please don't stare!



The 4 common data types

- Numbers (123.456)
- Text ("Hello!")
- Dates (1900-02-29)
- Logical (TRUE)

ISNUMBER() checks if a cell contains a number

	A	B	C
1	<i>Value</i>	<i>Command</i>	<i>Result</i>
2	9.869604	=ISNUMBER(A2)	TRUE
3	Data science is fun!	=ISNUMBER(A3)	FALSE

Other "IS" functions for checking types

- `ISTEXT()` checks for text.
- `ISDATE()` checks for dates and times.
- `ISLOGICAL()` checks for `TRUE` and `FALSE`.

Checking cell properties and rarer types

- `ISURL()` checks if text is web URL.
- `ISFORMULA()` checks if the cell contains a formula.

Copying formulas between cells

fx		=ISNUMBER(A1)					
	A	B	C	D	E	F	G
1	Mersenne	3	seven	31	127	8 191	131,071
2	FALSE						

Drag the bottom-right corner of the cell to copy. Alternatively,

- to copy right, use CTRL + R (CMD + R on macOS).
- to copy down, use CTRL + D (CMD + D on macOS).

Let's go!

INTERMEDIATE SPREADSHEETS

Convert or die!

INTERMEDIATE SPREADSHEETS



Richie Cotton

Data Evangelist at DataCamp

Formatting numbers

	A	B	C
1	<i>Value</i>	<i>Command</i>	<i>Result</i>
2	1.234	=TO_DOLLARS(A2)	\$1.23
3		=TO_PERCENT(A2)	123%

Converting text to numbers

	A	B	C
1	<i>Value</i>	<i>Command</i>	<i>Result</i>
2	'1.23	=ISTEXT(A2)	TRUE
3		=N(A2)	1.23
4		=ISNUMBER(C3)	TRUE

Converting logicals to numbers

	A	B	C
1	<i>Value</i>	<i>Command</i>	<i>Result</i>
2	TRUE	=N(A2)	1
3	FALSE	=N(A3)	0

Converting units of numbers

	A	B	C
1	<i>Value</i>	<i>Command</i>	<i>Result</i>
2	1000	=CONVERT(A2, "ft^2", "m^2")	92.90304

Summary

- `T0*()` functions format numbers.
- `N()` changes text to numbers.
- `IF(logical, 1, 0)` changes logicals to numbers.
- `CONVERT()` changes the units of a number.

Let's fix some data!

INTERMEDIATE SPREADSHEETS