

# Transforming text in Power Query

DATA PREPARATION IN POWER BI





**Maarten Van den Broeck**




Content Developer at DataCamp

# What is clean text data?

- Free from typos
- Consistently formatted
  - Data points are consistently represented
  - Uniform capitalization
  - No leading or trailing whitespace
  - No punctuation(!) or control characters (\n, \r, etc.)
- Ideally, each column will only store one piece of information
  - Columns are split or merged to achieve this

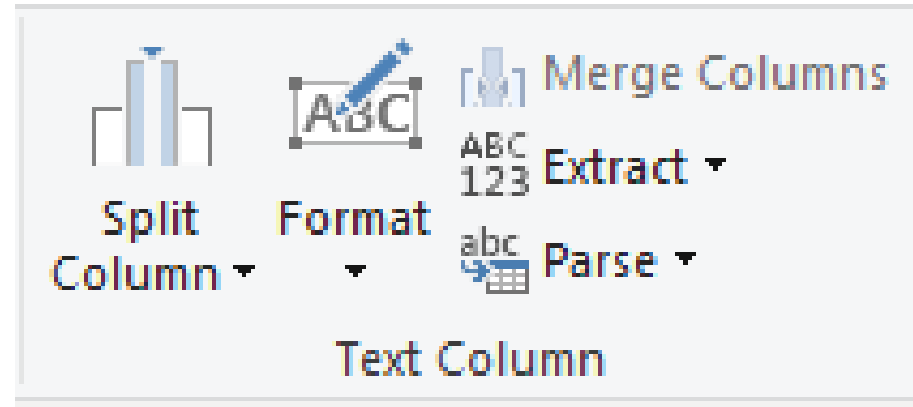
Yee**l**ow -> Yellow

	A <sup>B</sup> <sub>C</sub> Dirty	A <sup>B</sup> <sub>C</sub> Clean
	 4 distinct, 4 unique	 2 distinct, 1 unique
1	United States	United States of America
2	USA	United States of America
3	United States of America	United States of America
4	US	United States of America
5	united States of AMERICA	United States of America

	A <sup>B</sup> <sub>C</sub> Location (Dirty)	A <sup>B</sup> <sub>C</sub> Country	A <sup>B</sup> <sub>C</sub> City
	 5 distinct, 5 unique	 5 distinct, 5 unique	 5 distinct, 5 unique
1	Cairo, Egypt	Egypt	Cairo
2	Washington DC, United States of America	United States of America	Washington DC
3	Paris, France	France	Paris
4	Munich, Germany	Germany	Munich
5	Madrid, Spain	Spain	Madrid

# How to clean text data?

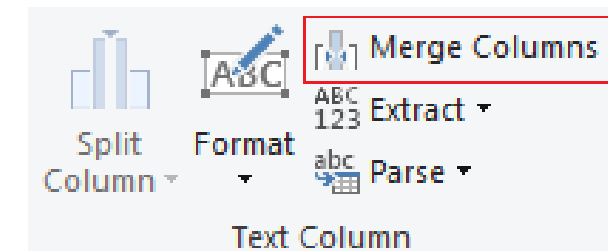
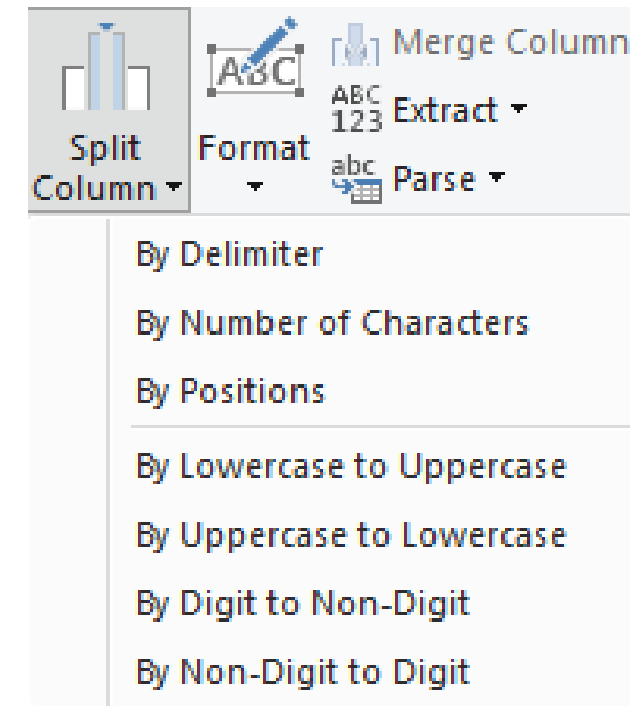
- Accessed from the Transform ribbon



- Use the Format option to change the capitalization and access Trim and Clean
- Trim & Clean should be applied to all text columns
  - Trim - removes trailing and leading whitespace
  - Clean - removes control characters (new line, carriage return, etc.)

# Splitting and combining columns

- Each column in your dataset should represent one piece of information
  - It is sometimes necessary to split or combine several columns
- When one column contains many pieces of information -> Split
  - Address -> Building, Street, City, Country
- When two or more columns can be combined to represent one piece of information -> Combine
  - First Name, Last Name -> Full Name



# Let's practice!

DATA PREPARATION IN POWER BI

# Text transformation

DATA PREPARATION IN POWER BI



**Maarten Van den Broeck**

Content Developer at DataCamp

# Let's practice!

DATA PREPARATION IN POWER BI