Transforming text in Power Query

DATA PREPARATION IN POWER BI



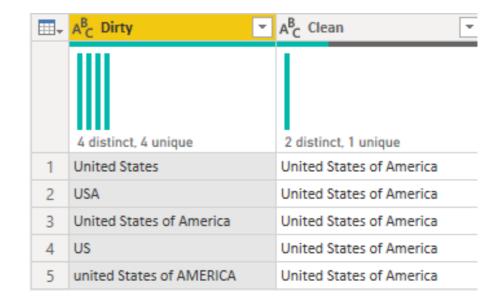
Maarten Van den Broeck Content Developer at DataCamp

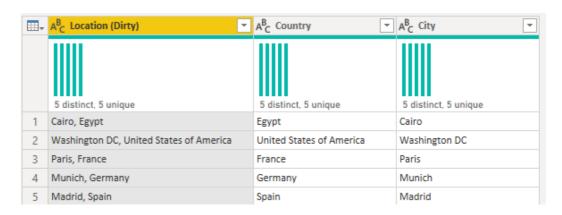


What is clean text data?

- 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

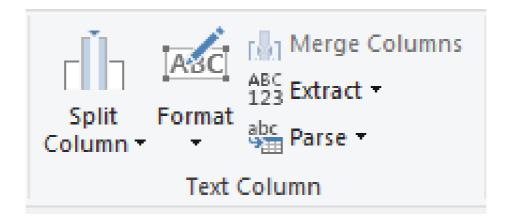
Yeelow -> Yellow





How to clean text data?

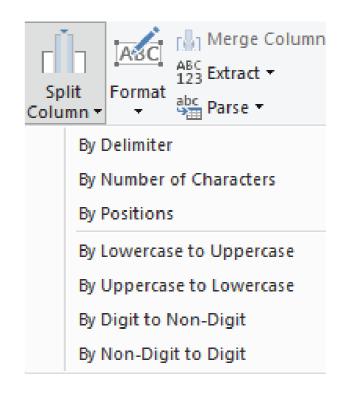
Accessed from the Transform ribbon

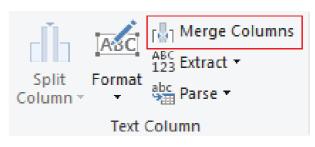


- Use the Format option to change the capitalization and access Trim and Clean
- Trim & Clean should be applied to all text columns
 - Trim removes trailing and leading whitespace
 - Clean removes control characters (new line, carriage return, etc.)

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

