Shaping tables

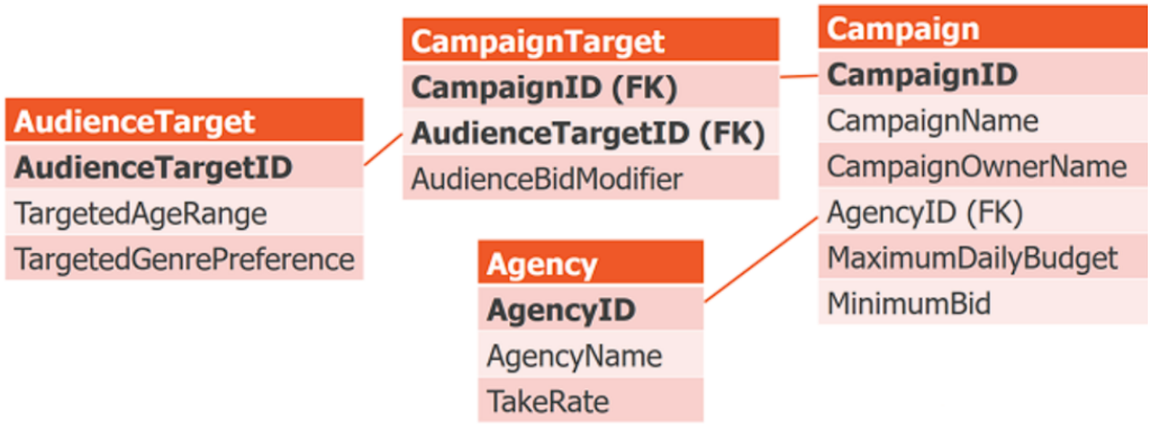
DATA MODELING IN POWER BI

Maarten Van den Broeck
Content Developer at DataCamp

datacamp

Database normalization

- A set of logical rules and processes to follow for data modeling
- Organizing a database
- Goals of normalization
  - Remove redundant data
  - Achieve a design which is a good representation of the real world
- Tables are connected through relationships in Power BI

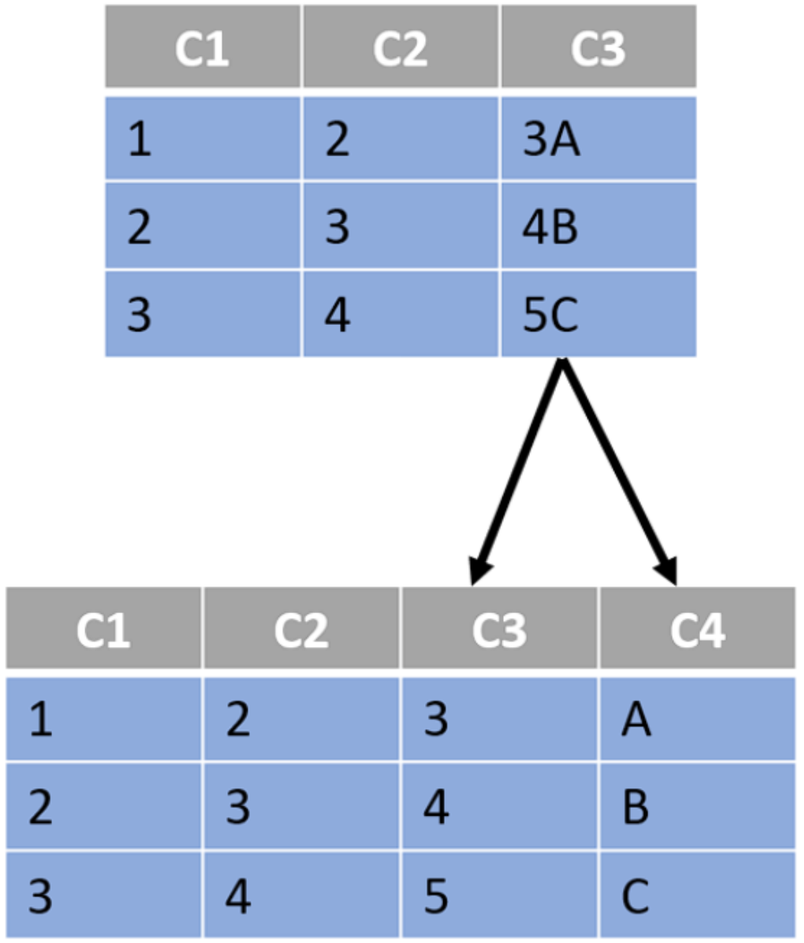


Data shaping in Power Query

- Power Query includes several data shaping operations to get closer to a normalized data model.
- Key techniques:
  1. Column splitting
  2. Column extraction
  3. Query merging
  4. Query appending
- There are additional techniques as well!

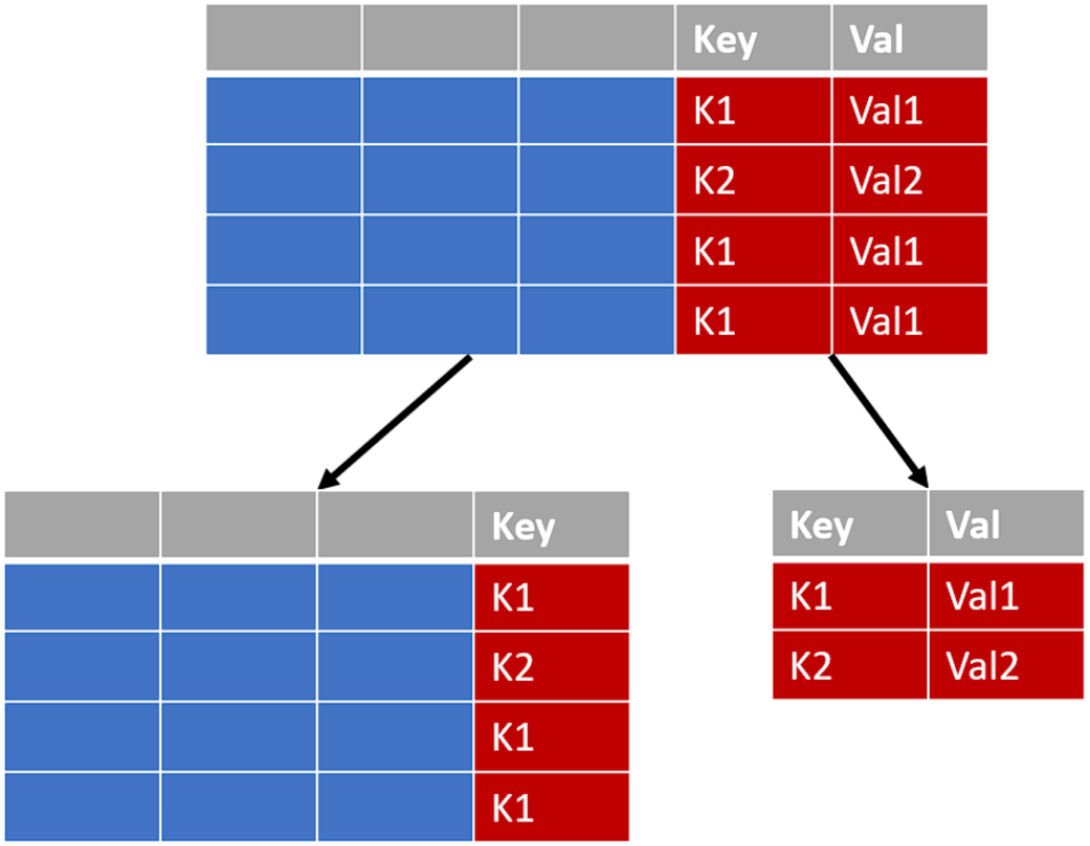
1. Column splitting

- Break out one column into multiple columns
- Split criteria
  - Delimiter
  - Number of characters
  - Position in string
  - Lower vs. upper casing
  - Digit vs. non-digit



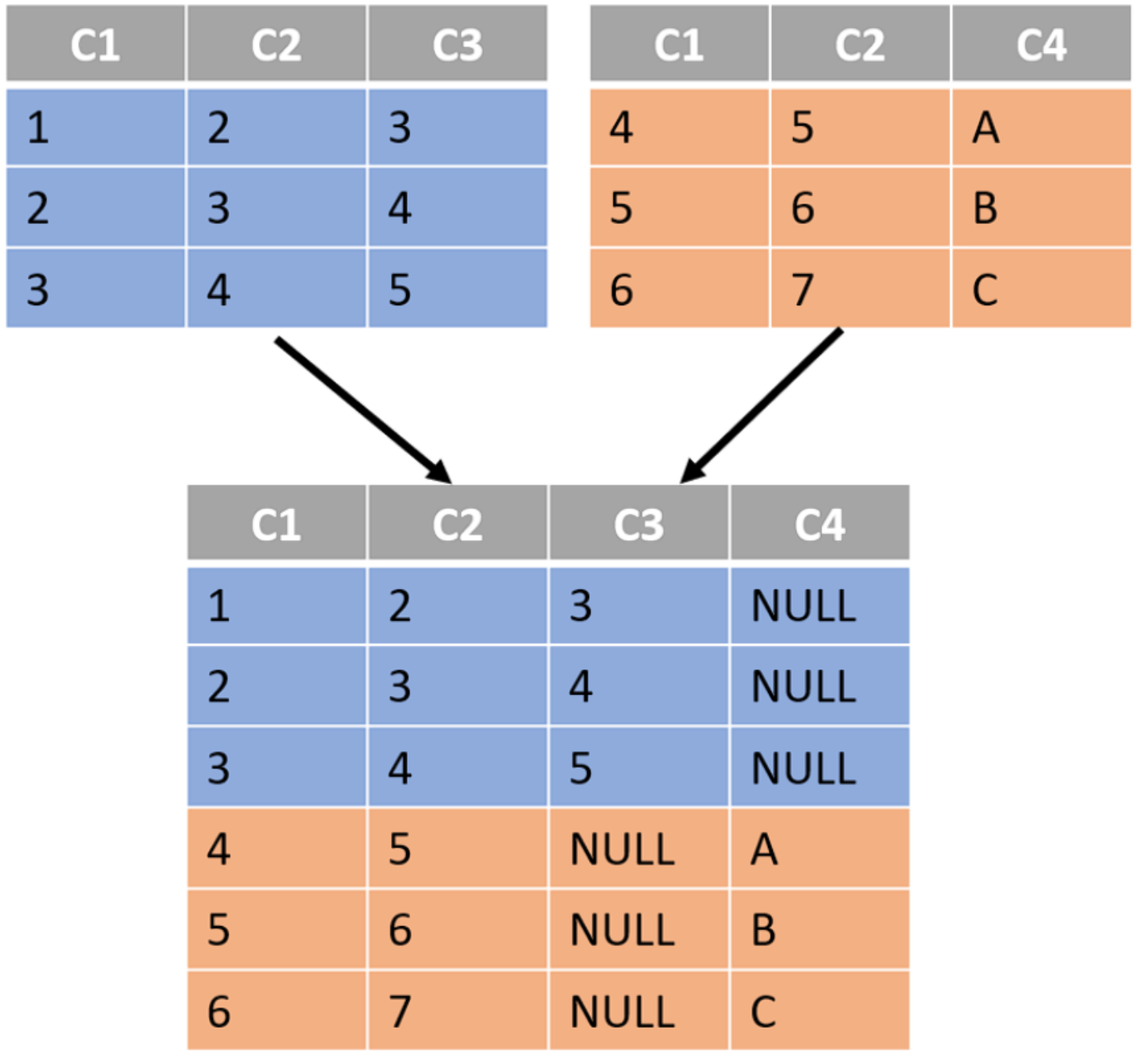
2. Column extraction

- Take columns from one table and break them out into another table
- Keep a **key** on the original table to know which values fit together
- Result: keep distinct rows, shrinking total data model size and reducing redundancy



3. Query appending

- Combine contents of two or more tables into a single table
- Match rows based on column names, adding *NULL* for missing columns
- Equivalent to a UNION ALL statement in SQL



C1	C2	C3
1	2	3
2	3	4
3	4	5

C1	C2	C4
4	5	A
5	6	B
6	7	C

C1	C2	C3	C4
1	2	3	NULL
2	3	4	NULL
3	4	5	NULL
4	5	NULL	A
5	6	NULL	B
6	7	NULL	C

4. Query merging

- Join together two existing tables based on values from one or more columns[1]
- Types of joins:
  - Inner join
  - Left outer join
  - Right outer join
  - Full outer join

ColA	ColB	ColC	Key
			K1
			K2
			K1
			K4

Key	Val
K1	Val1
K2	Val2
K3	Val3

Inner Join

ColA...ColC	Key	Val
	K1	Val1
	K2	Val2
	K1	Val1

(Left) Outer Join

ColA...ColC	Key	Val
	K1	Val1
	K2	Val2
	K1	Val1
	K4	*null*

Full Outer Join

ColA...ColC	Key	Val
	K1	Val1
	K2	Val2
	K1	Val1
	K4	*null*
null	K3	Val3

[1] This won't be covered in the exercises; check other DataCamp courses on joining tables for more detail

Let's practice!

DATA MODELING IN POWER BI

Merging and appending queries

DATA MODELING IN POWER BI

Maarten Van den Broeck
Content Developer at DataCamp

datacamp

Let's practice!

DATA MODELING IN POWER BI