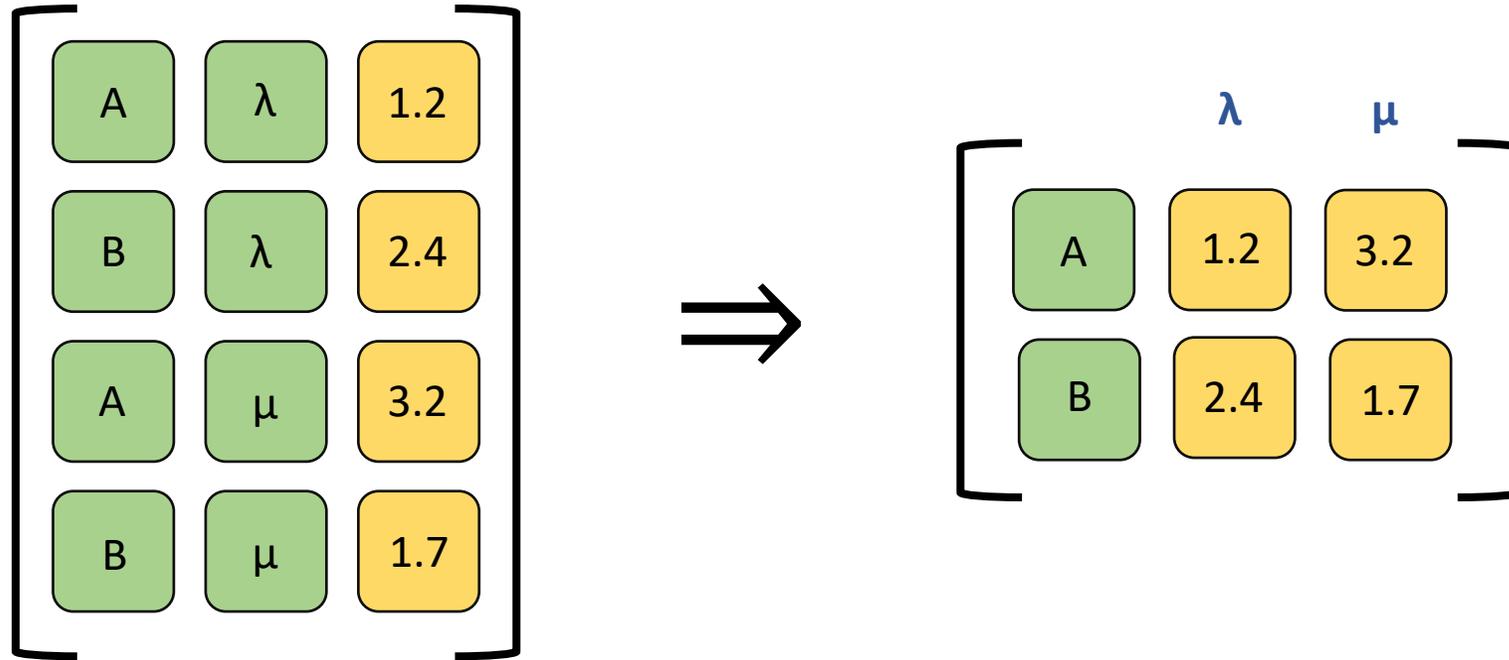
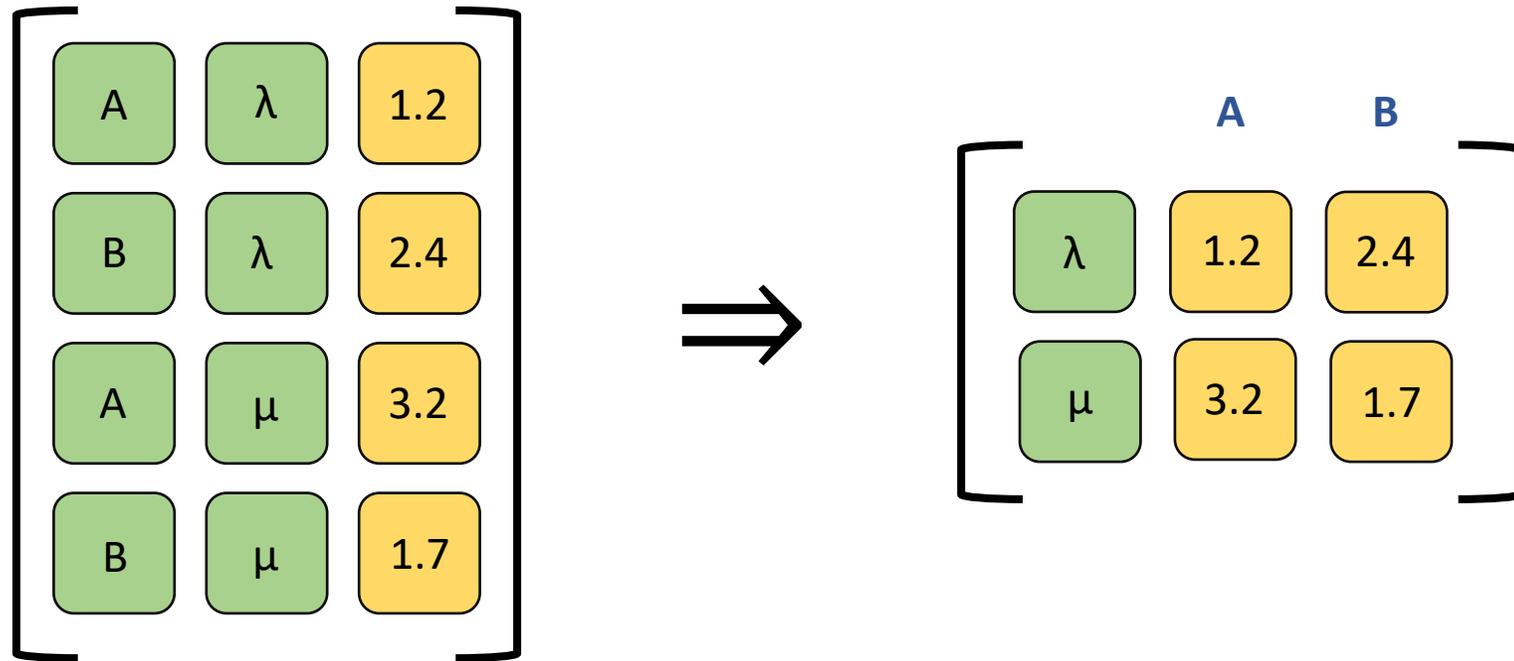


PIVOT WIDER



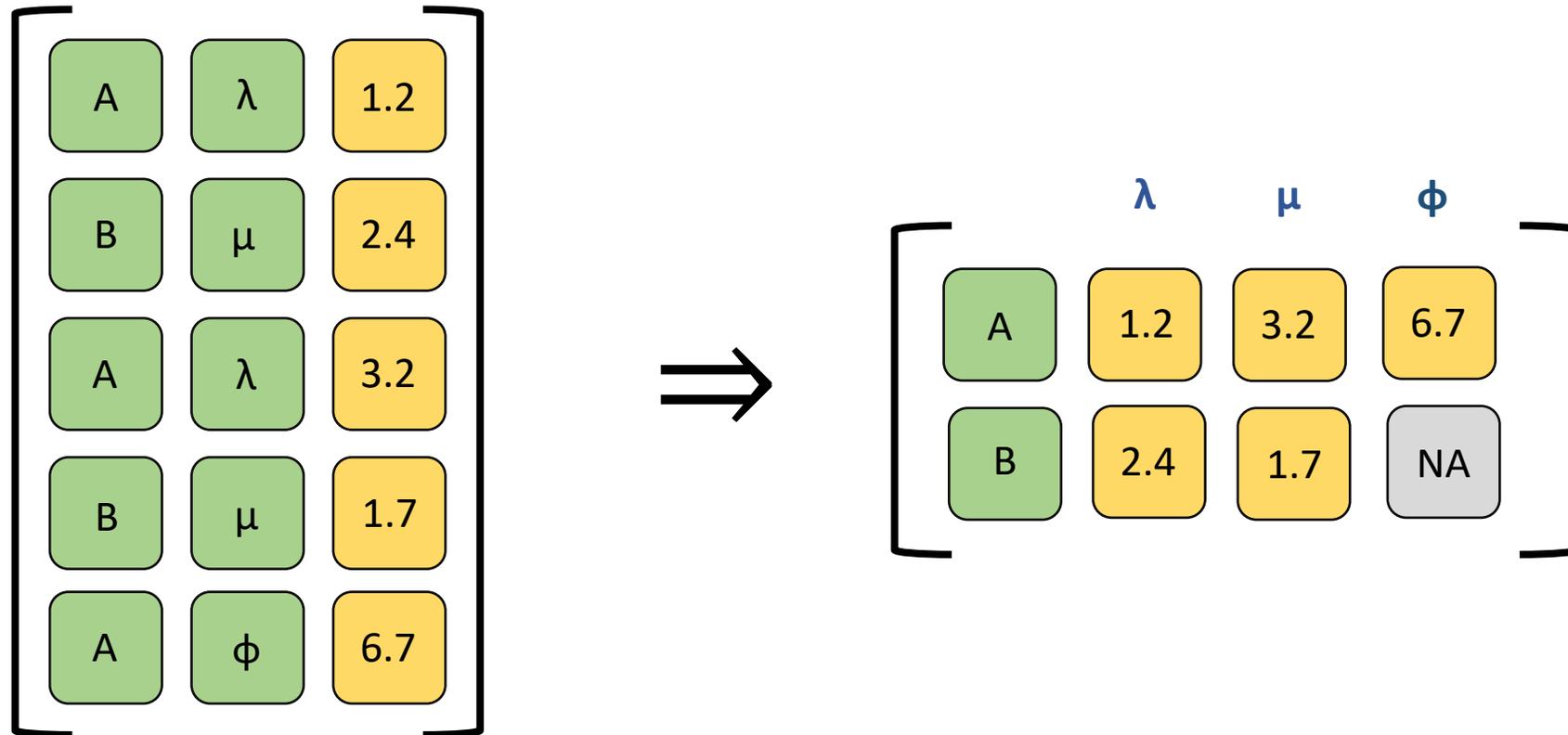
PIVOT WIDER

There are often different ways that we can pivot wider, depending on the type of analysis you want to perform.



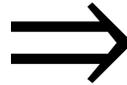
PIVOT WIDER

When pivoting wider, there is a chance that some values in the new table do not correspond to anything in the longer table. These are missing by default, but we can set a different default if needed.



PIVOT LONGER

	λ	μ	ϕ
A	1.2	3.2	6.7
B	2.4	1.7	NA



A	λ	1.2
B	λ	2.4
A	μ	3.2
B	μ	1.7
A	ϕ	6.7
B	ϕ	NA

BEST PRACTICES

It is highly recommended to try to store your data in the longer format rather than a wider format. This has several benefits:

- there are different ways to make the data in a wider format but usually only one way to make it longer
- there are fewer variables that you need to validate and document
- pivot longer is a real pain to work with if the column names in the original dataset represent numbers

The syntax for pivots are not terribly difficult, but using them really requires thinking about what you need.

THEORY

All of the previous one- and two-table verbs can be described by the theory of **relational algebra**.

Pivots operations are included in most database software, but are not part of this theoretical model. Pivots interchange rows and columns and that breaks the underlying assumptions of relational algebra.

This is not particularly important to know from the standpoint of doing data science, but does explain why we treat them separately. They are also included in a different R package (**tidyr** rather than **dplyr**).