

Para saber mais

Vimos que usamos o recurso de empilhar bases quando temos bases semelhantes e queremos juntar as observações de todas elas em uma única base. Mas o que acontece quando empilhamos bases que são *diferentes*?

Vamos usar como exemplo novamente duas listas de contatos, uma tem o número do telefone e a outra tem o endereço de email:

Nome	Telefone
Fulano	555-1234
Beltrano	321-7890
Sicrano	667-7882

Nome	Email
Fulano	fulano@xpto.org
Beltrano	siclano@acb.com
Sicrano	beltrano@xpto.org

Se ao invés de cruzarmos essas duas bases nós as empilhássemos usando um *data step*, teríamos como resultado a seguinte base:

Nome	Telefone	Email
Fulano	555-1234	
Beltrano	321-7890	
Sicrano	667-7882	
Fulano		fulano@xpto.org
Beltrano		siclano@acb.com
Sicrano		beltrano@xpto.org

Ou seja, no processo de empilhar bases com um *data step*, o SAS coloca na base final todas as variáveis das bases participantes, preenchendo com *missings* as observações de uma variável que uma base não tem.

Agora, se empilhássemos as bases usando o *SQL* com os comandos `union` ou `union all`, o resultado seria a base abaixo:

Nome	Telefone
Fulano	555-1234
Beltrano	321-7890
Sicrano	667-7882
Fulano	fulano@xpto.org
Beltrano	siclano@acb.com
Sicrano	beltrano@xpto.org

Ele simplesmente concatena a base, sem olhar se o nome das variáveis são iguais. Caso uma tabela tenha mais variáveis que a outra, essa variável adicional é tratada de uma forma semelhante que no *data step*: a variável é adicionada e preenchida com *missings* nas observações previamente inexistentes.