

SOC 4015/5050: Lecture-07 - Reshaping Data Replication

Christopher Prener, Ph.D.

Fall 2018

Part 1: Reshaping Long Data to Wide

participant	period	score
a	test1	.82
a	test2	.94
b	test1	.78
b	test2	.84

1. If we wanted to reshape these data, which verb is most appropriate? Why? **Since we want to reshape these data from long to wide, the verb that is most appropriate is “spread”.**
2. What is the key? **The key here is period, since it contains the two values that should be converted to column names.**
3. What is the value? **The value here is score, since it contains the data that should populate the new variables indicated by the key categories.**
4. Draw out a reshaped data table with new variable names and values filled in.

participant	test1	test2
a	.82	.94
b	.78	.84

Part 2: Reshaping Wide Data to Long

class	midterm	final
Social Studies, 4 th Period	88	91
Social Studies, 6 th Period	87	86

5. If we wanted to reshape these data, which verb is most appropriate? Why? *Since we want to reshape these data from wide to long, the verb that is most appropriate is “gather”.*
6. What could we name the key? *We could name the key period or exam given the context of the table.*
7. What could we name the value? *We could name the value score given the context of the table.*
8. Which variables (i.e. columns) will contribute to the values? *Both the midterm and final variables should contribute to the values.*
9. Draw out a reshaped data table with new variable names and values filled in.

class	period	score
Social Studies, 4 th Period	midterm	88
Social Studies, 4 th Period	final	91
Social Studies, 6 th Period	midterm	87
Social Studies, 6 th Period	final	86