

DATA STACKING

Dec 9, 2019



Data Stacking

Technique



stack

/stak/

verb

gerund or present participle: **stacking**

1. arrange (a number of things) in a pile, typically a neat one.
"the books had been stacked up in three piles"

Similar: heap (up) pile (up) make a heap/pile/stack of assemble

2. shuffle or arrange (a pack of cards) dishonestly so as to gain an unfair advantage.
"I know the cards are stacked"

Data stacking also commonly referred to as data rotation, is a data transformation technique whereby data for a single person is stacked in multiple rows. This technique is useful in instances where a comparative analysis is required. It's especially useful in projects where results need to be reviewed side by side for various ads, media, and/or concepts to be shown to responders.

In survey research, when a responder is shown multiple media and subsequent questions is asked about the exposure, data is generally in flat format, one responder per row.

Res pID	Ad1	Ad Feed back	Ad Recom mend	Ad Purchase Intent	Ad2	Ad Feed back	Ad Recom mend	Ad Purchase Intent	Ad3	Ad Feed back	Ad Recom mend	Ad Purchase Intent
1	ABC Ad	Love d it	Yes	Probably Would	GHI Ad	Not that great	Maybe	Probably Would	XYZ Ad	Love d it	Yes	Definitely Would

In above example, 3 ads were shown and after each ad, 3 questions were asked, in such a scenario, the data in its natural format would not allow for a comparative analysis.

In such cases, the data needs to be stacked. Stacking here would create three lines of data per responder with ads shown appearing in one column and all subsequent questions in their respective columns. Example below

RespID	Ad	Ad Feedback	Ad Recommend	Ad Purchase Intent
1	ABC Ad	Loved it	Yes	Probably Would
1	GHI Ad	Not that great	Maybe	Probably Would
1	XYZ Ad	Loved it	Yes	Definitely Would

Once the data has been reformatted (stacked) in the above format it can easily be plugged into any analytical/tabulations software for performing comparative analysis.

Use In tabulations

Stacked data can be very easily fed into any tabulations software for running a comparative analysis, also referred as grid tables. Example below

Table 1
Q1. How appealing is CONCEPT compared to other products currently available?
Base - Total Respondents

	Brands					
	Concept 1	Concept 2	Concept 3	Concept 4	Concept 5	Concept 6
	(A)	(B)	(C)	(D)	(E)	(F)
Total	103	103	103	102	101	103
Top 2 Box (Net)	70 68.0%	57 55.3%	67 65.0%	76 74.5%	65 64.4%	75 72.8%
Very appealing (5)	b 28 27.2%	26 25.2%	23 22.3%	B 25 24.5%	28 27.7%	B 41 39.8%
Appealing (4)	42 40.8%	31 30.1%	44 42.7%	51 50.0%	37 36.6%	aBCDe 34 33.0%
A little appealing (3)	22 21.4%	29 28.2%	b 25 24.3%	BeF 19 18.6%	25 24.8%	17 16.5%
Bottom 2 Box (Net)	11 10.7%	F 17 16.5%	11 10.7%	7 6.9%	11 10.9%	11 10.7%
Not very appealing (2)	9 8.7%	D 14 13.6%	6 5.8%	5 4.9%	6 5.9%	7 6.8%
Not at all appealing (1)	2 1.9%	cDe 3 2.9%	5 4.9%	2 2.0%	5 5.0%	4 3.9%
Sigma	103	103	103	102	101	103
Mean	3.83	3.61	3.72	B 3.90	3.76	Bc 3.98
Std.Dev.	0.99	1.10	1.03	0.90	1.08	1.09
Std.Err.	0.10	0.11	0.10	0.09	0.11	0.11

Comparison Groups: ABCDEF
T-Test for Means, Z-Test for Percentages
Uppercase letters indicate significance at the 95% level.
Lowercase letters indicate significance at the 90% level.
**#* denotes a cell for which statistical testing was suppressed because the filter's frequency was less than 30.

How it's done

Stacking is done most efficiently via SQL, ensuring the rotation is done with accuracy and speed. Once a syntax has been built for a project using a subset of data, it can be repeated without much effort in the project.

Benefits

- Stat testing of the ads/media/segments/concepts can be done automatically via tabulations software, no manual stat testing needed
- Reduces manual work in formatting tabulated output to allow for a side by side view thereby improving turnaround time
- Helps in reducing the reporting turnaround time

Disadvantages of stacked data files

Data files that have been stacked cannot be used to run analysis for questions that are outside the ad/media/concept evaluation section. As the data file will have multiple rows per responder, it would artificially inflate numbers for non-media evaluation questions. E.g. in a survey of 100 responders, where each person was shown three concepts, the stacked data file will have 300 lines of data, compared to 100 in unstacked/flat/standard file.