

Coding examples

FL2012.B55.MKT.473.01 - Marketing Research
Washington University in St. Louis

Professor César Zamudio

Below are some examples of how to code questions based on different levels of measurement:

NOMINAL SCALE - Dichotomous

Q. Are you currently enrolled as an undergraduate student at Washington University in St. Louis?

- Yes
 No

Coding options

		Indices		Dummy variable
Text coding	Prespecified values	UG_YES	UG_NO	UG_YES
Yes	1	1	0	1
No	2	0	1	0

NOMINAL SCALE - Multichotomous (with "Other" option to be specified)

Q. Which academic school are you in?

- Olin Business School
 Arts & Sciences
 Engineering
 Other (please specify): _____

Coding options

Text coding	Prespecified values	Indices				Dummy variable set			
		OLIN	ARTSSCI	ENG	OTHER	OLIN	ARTSSCI	ENG	OTHER
OLIN	1	1	0	0	0	1	0	0	0
ARTSSCI	2	0	1	0	0	0	1	0	0
ENG	3	0	0	1	0	0	0	1	0
OTHER	4	0	0	0	1	0	0	0	0

If "Other" is chosen, an additional column for the consumers' response is required. Below is an example where 5 subjects answered both questions above:

Subject ID	YES	OLIN	ARTSSCI	ENG	OTHER	OTHER_TXT
1	1	1	0	0	0	.
2	1	0	1	0	0	.
3	0
4	1	0	0	0	1	Sam Fox
5	1	0	0	0	1	Nursery

NOMINAL SCALE - Multiple choice (with "Other" option to be specified)

This coding procedure is similar to the above, except that the categories are no longer mutually exclusive, so indices are used.

12. How would you prefer to find about enrolling in our MBA program? (check all that apply)

- Career Center
- Academic advisor
- Friends and relatives
- E-Mail
- Flyers
- Other (please specify) _____

Subject ID	CCENTER	ADVISOR	FRIEND_REL	E-MAIL	FLYER	OTHER	OTHER_TXT
1	1	0	0	0	0	0	
2	0	0	0	1	1	1	Twitter
3	0	0	0	1	1	0	
4	0	0	0	1	0	0	
5	1	1	1	1	1	1	Facebook

Here Subject 1 selected only Career Center; Subject 2 selected E-Mail, Flyer, and Other; Subject 3 selected E-Mail and Flyer; Subject 4 selected E-Mail only; and Subject 5 selected all methods.

ORDINAL SCALE - Exhaustive task

Please rank the following headphone brands in order of preference, where 1 is most preferred and 5 is least preferred.

- _____ Beats by Dr. Dre
- _____ Bose
- _____ Koss
- _____ Monster Cable
- _____ Sennheiser

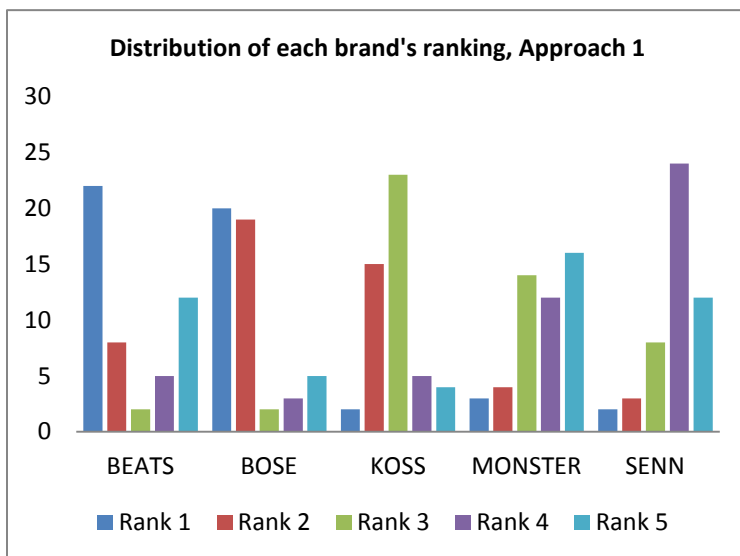
The key here is that **you cannot** code ordinal responses using a single column. You need multiple columns! The first approach is to have **objects in columns, ranks in data**. Here's an example with 5 subjects and their responses. In this case 1=Brand is ranked #1, 2= Brand is ranked #2, and so forth.

Subject ID	BEATS	BOSE	KOSS	MONSTER	SENN
1	1	2	3	4	5
2	2	3	4	5	1
3	3	4	5	1	2
4	4	5	1	2	3
5	5	1	2	3	4

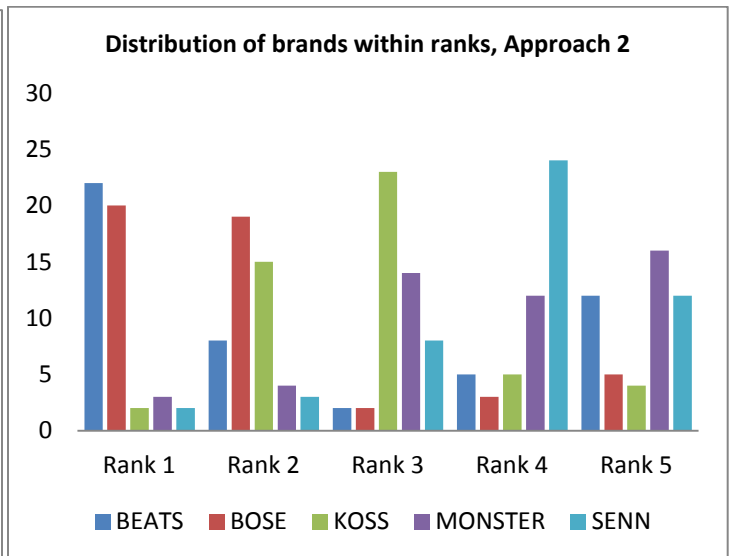
An alternative is to have **ranks in columns, objects in data**. This approach requires to (1) specify a coding scheme for the objects evaluated (e.g. Beats=1, Bose =2, Koss=3, Monster=4, Sennheiser=5), (2) create a column for each possible ranking, and (3) fill up the columns according to how each brand was ranked. The values in the table below correspond to the same responses in the table above using the alternative approach.

Subject ID	RANK1	RANK2	RANK3	RANK4	RANK5
1	1	2	3	4	5
2	5	1	2	3	4
3	4	5	1	2	3
4	3	4	5	1	2
5	2	3	4	5	1

Choosing the first vs. the second approach depends on the information the researcher is interested in. The first approach is useful to determine **the distribution of each brand's ranking**; the second approach is useful to determine **the distribution of each ranking option among brands**. Examples:



n=49




n=49



You **do not need to use both approaches to create both graphs**. You only need to conduct the following steps: (1) find the frequency distribution of each column using Approach 1; (2) transpose the matrix; (3) find the frequency distribution of the resulting matrix, as follows:

Freqs.	BEATS	BOSE	KOSS	MONSTER	SENN
Rank 1	22	20	2	3	2
Rank 2	8	19	15	4	3
Rank 3	2	2	23	14	8
Rank 4	5	3	5	12	24
Rank 5	12	5	4	16	12
n	49	49	49	49	49



Freqs.	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
BEATS	22	8	2	5	12
BOSE	20	19	2	3	5
KOSS	2	15	23	5	4
MONSTER	3	4	14	12	16
SENN	2	3	8	24	12
n	49	49	49	49	49

SPSS or Excel's PivotTables can be used to this end.

ORDINAL SCALE - Non-Exhaustive task

Sometimes the ranking task the subject is asked to perform is not exhaustive, as in the following example:

Q. Below is a list of business graduate program characteristics. Please choose the three most important characteristics for you, and assign them a number, where 1 means "Most important among my 3 choices" and 3 means "Least important among my 3 choices."

Time investment required Academic challenge
 Students' ethnic diversity Program prestige
 Potential for advancing career

Here the coding is similar to the previous example, except that some objects will not be ranked all the time:

Subject ID	TIME	ACAD_CH	ETHNIC	PRESTIGE	CAREER
1	1	2	3	.	.
2	1	.	.	3	2
3	3	.	1	2	.
4	.	1	2	.	3
5	.	.	1	2	3

Frequency analysis is similar as in the example above.

INTERVAL SCALE

Attitudinal scales are the easiest to code. Only input the data directly from the survey. In this example we'll have three items of each scale. Likelihood and other scales are coded and entered in the same way.

Subject ID	LIKERT			SEM. DIFFERENTIAL			STAPEL			RATING TASK		
	ITEM1	ITEM2	ITEM3	ITEM1	ITEM2	ITEM3	ITEM1	ITEM2	ITEM3	ITEM1	ITEM2	ITEM3
1	3	4	5	7	4	3	5	-5	-3	9	7	9
2	5	2	3	6	1	2	-4	-5	2	10	1	10
3	4	3	2	4	5	6	2	3	4	7	2	8
4	3	4	5	7	7	7	4	-5	2	10	3	4
5	5	3	2	1	7	1	-2	-2	3	6	6	7

RATIO SCALE

Ratio variables should always be entered directly from the survey. One can later produce categories or brackets using quartile splits, observing the distribution of the data, etc.