How can reverse coding be performed in R, and do you have an example?

Authored by

stats writer

June 27, 2024

Reverse coding in R refers to the process of changing the coding of a variable from a positive to a negative scale, or vice versa. This is often done in research studies to avoid response bias or to create a more balanced scale. In order to perform reverse coding in R, one can use the "reverse.code" function from the "psych" package. This function takes in a variable and reverses its coding, generating a new variable with the reversed scores. For example, if a variable has scores ranging from 1 to 5, the "reverse.code" function will change them to 5 to 1. This allows researchers to easily perform reverse coding in a systematic and efficient way.

Perform Reverse Coding in R (With Example)

When creating surveys, researchers sometimes rephrase "positive" questions in a "negative" way to make sure that individuals are giving consistent responses.

We say that these types of questions are reverse-coded.

When using a survey to assign a composite score to individuals, it's important to make sure the reverse-coded questions are reverse-scored as well.

The following example shows how to reverse the scores on reverse-coded questions in R.

Example: Reverse Coding in R

Suppose researchers administer a survey with 5 questions to 10 individuals in which the possible responses to each questions are:

Strongly AgreeAgreeNeither Agree Nor DisagreeDisagreeStrongly Disagree

The following data frame contains the results of the survey in which "Strongly Agree" is assigned a value of 5, "Agree" is assigned a value of 4, and so on:

```
#create data frame that contains survey results
df <- data.frame(Q1=c(5, 4, 4, 5, 4, 3, 2, 1, 2, 1),
Q2=c(1, 2, 2, 1, 2, 3, 4, 5, 4, 5),
Q3=c(4, 4, 4, 5, 4, 3, 2, 4, 3, 1),
Q4=c(3, 4, 2, 2, 1, 2, 5, 4, 3, 2),
Q5=c(2, 2, 3, 2, 3, 1, 4, 5, 3, 4))

#view data frame
df
```

```
Q1 Q2 Q3 Q4 Q5
1 5 1 4 3 2
2 4 2 4 4 2
3 4 2 4 2 3
4 5 1 5 2 2
5 4 2 4 1 3
6 3 3 3 2 1
7 2 4 2 5 4
```

8 1 5 4 4 5

9 2 4 3 3 3

10 1 5 1 2 4

Suppose questions 2 and 5 are reverse coded, so we must reverse their scores.

That is:

1 should become 5.2 should become 4.3 should become 3.4 should become 2.5 should become 1.

The easiest way to do this is to take the max possible score (5) and add 1 to get 6. Then subtract the original scores from 6 to get the reverse scored value.

For example:

5 becomes: 6 - 5 = 1.4 becomes: 6 - 4 = 2.3 becomes: 6 - 3 = 3.2 becomes: 6 - 2 = 4.1 becomes: 6 - 1 = 5.

We can use the following code to do this in R:

```
#define columns to reverse code
reverse_cols = c("Q2", "Q5")
```

#reverse code Q2 and Q5 columns

df = 6 - df

#view updated data frame

df

Q1 Q2 Q3 Q4 Q5

1 5 5 4 3 4

2 4 4 4 4 4

3 4 4 4 2 3

4 5 5 5 2 4

5 4 4 4 1 3

6 3 3 3 2 5

7 2 2 2 5 2

8 1 1 4 4 1

9 2 2 3 3 3

10 1 1 1 2 2