

How can I create a contingency table in R?

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I create a contingency table in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138764>

A contingency table is a useful tool for organizing and analyzing categorical data. In order to create a contingency table in R, the user must first import the data into the software using the appropriate function, such as "read.csv" or "read.table". Once the data is loaded, the user can use the "table" function to generate a contingency table, which displays the frequency of each category combination. Additional functions, such as "prop.table" and "margin.table", can be used to calculate proportions and marginal totals, respectively. The resulting contingency table can then be further analyzed and visualized using various statistical techniques and graphical tools in R.

Create a Contingency Table in R

A contingency table (sometimes called "crosstabs") is a type of table that summarizes the relationship between two categorical variables.

Fortunately it's easy to create a contingency table for variables in R by using the pivot table function. This tutorial shows an example of how to do so.

Example: Contingency Table in R

Suppose we have the following dataset that shows information for 20 different product orders, including the type of product purchased along with the country that the product was purchased in:

#create data

```
df <- data.frame(order_num = 1:20,  
product=rep(c('TV', 'Radio', 'Computer'), times=c(9, 6,  
5)),
```

```
country=rep(c('A', 'B', 'C', 'D'), times=5))
```

```
#view data
```

```
df
```

```
order_num product country
```

```
1 1 TV A
```

```
2 2 TV B
```

```
3 3 TV C
```

```
4 4 TV D
```

```
5 5 TV A
```

```
6 6 TV B
```

```
7 7 TV C
```

```
8 8 TV D
```

```
9 9 TV A
```

```
10 10 Radio B
```

```
11 11 Radio C
```

```
12 12 Radio D
```

```
13 13 Radio A
```

```
14 14 Radio B
```

```
15 15 Radio C
```

```
16 16 Computer D
```

```
17 17 Computer A
```

```
18 18 Computer B
```

19 19 Computer C

20 20 Computer D

To create a contingency table, we can simply use the `table()` function and provide the variables `product` and `country` as the arguments:

```
#create contingency table
```

```
table <- table(df$product, df$country)
```

```
#view contingency table
```

```
table
```

```
A B C D
```

```
Computer 1 1 1 2
```

```
Radio 1 2 2 1
```

```
TV 3 2 2 2
```

We can also use the `addmargins()` function to add margins to the table:

```
#add margins to contingency table
```

```
table_w_margins <- addmargins(table)
```

```
#view contingency table
```

table_w_margins

A	B	C	D	Sum	
Computer	1	1	1	2	5
Radio	1	2	2	1	6
TV	3	2	2	2	9
Sum	5	5	5	5	20

Here is how to interpret the table:

The value in the bottom right corner shows the total number of products ordered: 20. The values along the right side show the row sums: A total of 5 computers were ordered, 6 radios were ordered, and 9 TV's were ordered. The values along the bottom of the table show the column sums: A total of 5 products were ordered from country A, 5 from country B, 5 from country C, and 5 from country D. The values inside the table show the number of specific products ordered from each country: 1 computer from country A, 1 radio from country A, 3 TV's from country A, etc.