



Microsoft: DAT209x Programming in R for Data Science



Bookmarks

- ▶ 0. Start Here
- ▶ 1. Introduction
- ▶ 2. Functions and Data Structures
- ▶ 3. Loops and Flow Control
- ▶ 4. Working with Vectors and Matrices
- ▶ 5. Reading in Data
- ▶ 6. Writing Data to Text Files
- ▶ 7. Reading Data from SQL Databases

9. Manipulating Data > Lab > Lab

Bookmark

Create a dataset in the following way:

```
my.data<-data.frame(Treatment=c(rep("A",4),rep("B",4)),  
                    Stone=rep(rep(c("Small","Large"),c(2,2)),2),  
                    Success=rep(c(1,0),4),  
                    Count=c(81,6,192,71,234,36,55,25))
```

Question 1

(2/2 points)

How many rows does the data frame have?

☐ 4☐ 6☒ 8 ✓

► 8. Working with Data

▼ 9. Manipulating Data

Lecture

Knowledge Checks

Quiz



Lab

Lab



☐ 10

How many columns does the data frame have?

☒ 4 ✓

☐ 6

☐ 8

☐ 10

EXPLANATION

You can observe the content of the data frame by typing `my.data`

The dataset contains information on successful and unsuccessful surgery on kidney stones after two methods of treatment, A and B. Use the `xtabs()` and `prop.table()` functions to work out which method that is the best.

Question 2

(1/1 point)

You want to create a three way contingency table from `my.data` to analyze the data.

Which command could you use to do so?

- ☐ `my.table<-xtabs(Treatment~Stone+Success+Count,data=my.data)`
- ☐ `my.table<-xtabs(Stone~Treatment+Success+Count,data=my.data)`
- ☐ `my.table<-xtabs(Success~Treatment+Stone+Count,data=my.data)`
- ☒ `my.table<-xtabs(Count~Treatment+Success+Stone,data=my.data)` ✓

EXPLANATION

Question 3

(2/2 points)

You want to compare the success of treatment A vs treatment B using `margin.table()` and `prop.table()`.

Which command could you use to do so?

- ☐ `margin.table(prop.table(my.table, 1:2),1)`
- ☐ `margin.table(prop.table(my.table, 1:2),2)`
- ☒ `prop.table(margin.table(my.table, 1:2),1)` ✓
- ☐ `prop.table(margin.table(my.table, 1:2),2)`

From the table you created, which treatment seems to have better success rate?

- ☐ It cannot be concluded from the table you created
- ☐ A
- ☒ B ✓
- ☐ They are the same

EXPLANATION

Question 4

(2/2 points)

However, it is required that treatment A and B are applied to comparable populations, before it can be concluded from the above that treatment B is better than treatment A. Let us see if this is the case. Let's start by looking into the **small** stone type. You want to compare the success rate of each treatment for the small stone type.

Which command could you use to do so?

- ☐ `prop.table(my.table[,1],1)`
- ☐ `prop.table(my.table[,1],2)`
- ☒ `prop.table(my.table[,2],1)` ✓
- ☐ `prop.table(my.table[,2],2)`

Which treatment seems to have better success rate for the **small** stone type?

☐ It cannot be concluded from the table you created

☒ A ✓

☐ B

☐ They are the same

EXPLANATION

What we can see here is that Treatment A is better than Treatment B among small stones.

Question 5

(2/2 points)

Let's continue our analysis by looking into the **large** stone type. You want to compare the success rate of each treatment for the large stone type.

Which command could you use to do so?

☒ `prop.table(my.table[,1],1)` ✓

☐ `prop.table(my.table[,1],2)`

☐ `prop.table(my.table[,2],1)`

☐ `prop.table(my.table[,2],2)`

Which treatment seems to have better success rate for the **large** stone type?

☐ It cannot be concluded from the table you created

☒ A ✓

☐ B

☐ They are the same

EXPLANATION

Treatment A is also better than Treatment B among large stones.

Question 6

(2/2 points)

Let us look at the success rate for the stones among each treatment. Let's start by comparing the success rate of each stone types for Treatment **A**.

Which command could you use to do so?

- ☐ `prop.table(my.table[1,,],1)`
- ☒ `prop.table(my.table[1,,],2)` ✓
- ☐ `prop.table(my.table[2,,],1)`
- ☐ `prop.table(my.table[2,,],2)`

Which stone type seems to have better success rate for the Treatment **A**?

- ☐ It cannot be concluded from the table you created

☒ Small ✓

☐ Large

☐ They are the same

EXPLANATION

Among treatment A, small stones have a better success rate.

Question 7

(2/2 points)

Let's continue by comparing the success rate of each stone types for Treatment **B**.

Which command could you use to do so?

☐ `prop.table(my.table[1,,1])`

☐ `prop.table(my.table[1,,2])`

☐ `prop.table(my.table[2,,1])`

☒ `prop.table(my.table[2,,2])` ✓

Which stone type seems to have better success rate for the Treatment **B**?

☐ It cannot be concluded from the table you created

☒ Small ✓

☐ Large

☐ They are the same

EXPLANATION

Also among treatment B, small stones have a better success rate.

Question 8

(1/1 point)

Are the stones randomly allocated to the treatments? Let us look at the marginal table for Stone and Treatment.

Which two commands could you use to do so?

☐ `prop.table(margin.table(my.table, 1:3),1)`

☒ `prop.table(margin.table(my.table, c(3,1)),1)` ✓

☐ `prop.table(margin.table(my.table, 3:1),2)`

☒ `prop.table(margin.table(my.table, c(1,3)),2)` ✓



EXPLANATION

What this table tells us is that large stones are largely allocated to treatment A, while small stones are largely allocated to treatment B. Since larger stones are more difficult to treat, this is what causes treatment B to appear to have a higher success-rate.



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

