

Statistics with R – Beginner Level

Section 3

Creating Frequency Tables and Cross Tables

Lesson 17 - Frequency Tables in Base R

```
demo <- read.csv("demographics.csv")

View(demo)

#####
### how to create frequency tables
### in base R
#####

### we will build a table for the variable educ (education
level)
### this table will contain the following:
### absolute frequencies (counts), cumulative absolute
frequencies,
### relative frequencies and cumulatitve relative
frequencies

### create the initial table (with the counts only)

mytable <- table(demo$educ, exclude = NULL)  ### the
missing values will be excluded
```

```

print(mytable)

### compute the cumulative counts (using the cumsum
function)

cumul <- cumsum(mytable)

print(cumul)

### compute the relative frequencies

relative <- prop.table(mytable)

print(relative)

### compute the cumulative relative frequencies

n <- nrow(demo)  ### number of rows (cases) of the data
frame demo

cumulfreq <- cumul/n

print(cumulfreq)

### create the final table with the cbind function

mytable2 <- cbind(Freq=mytable, Cumul=cumul,
Relative=relative, CumFreq=cumulfreq)

print(mytable2)

###
### the commands above can be used with numeric variables
as well
###

```

Lesson 18 - Frequency Tables with plyr

```
demo <- read.csv("demographics.csv")
```

```
View(demo)
```

```
#####
```

```
### how to create frequency tables
```

```
### with the package plyr (function count)
```

```
#####
```

```
### we will build a table for the variable educ (education level)
```

```
### this table will contain the following:
```

```
### absolute frequencies (counts), cumulative counts,
```

```
### relative frequencies and cumulative relative frequencies
```

```
### load the package
```

```
require(plyr)
```

```
### build the initial table, with the absolute frequencies
```

```
mytable <- count(demo, 'educ')
```

```
print(mytable)
```

```
# compute the percentages (relative frequencies)
```

```
perc <- mytable$freq/nrow(demo)
```

```
print(perc)
```

```
### compute the cumulative counts
```

```
cumul <- cumsum(mytable$freq)
```

```
print(cumul)
```

```
### compute the cumulative percentages
```

```
cumperc <- cumul/nrow(demo)
```

```
print(cumperc)
```

```
# add the cumulative counts and the percentages to the
initial table
```

```
mytable <- cbind(mytable, cumul, perc, cumperc)
```

```
print(mytable)
```

Lesson 19 - Building Cross Tables using xtabs

```
demo <- read.csv("demographics.csv")
```

```
View(demo)
```

```
#####
```

```
### how to create cross-tables
```

```
### in base R (with the xtabs function)
```

```
#####
```

```
### we will build a cross table with the variables gender
and carcat (car category)
```

```
ct <- xtabs(~gender+carcat, data=demo)
```

```
fTable(ct)
```

Lesson 20 - Building Cross Tables with CrossTable

```
demo <- read.csv("demographics.csv")
```

```
View(demo)
```

```
#####
```

```
### how to create cross-tables
```

```
### with the CrossTable function (package gmodels)
```

```
#####
```

```
### we will build a cross table with the variables gender
and carcat (car category)
```

```
### load the package
```

```
require(gmodels)

CrossTable(demo$gender, demo$carcat, prop.chisq = FALSE)

### we don't want the chi square contributions

### some other options of the CrossTable function

CrossTable(demo$gender, demo$carcat, digits=3,
expected=TRUE, prop.r=TRUE, prop.c=TRUE,
           prop.t=TRUE, prop.chisq=TRUE, chisq = FALSE,
fisher=FALSE, mcnemar=FALSE,
           missing.include=FALSE)

### digits: number of decimals

### expected: shows the expected frequencies

### prop.r: shows the row proportions

### prop.c: shows the column proportions

### prop.chisq: shows the chi square contributions

### chisq: computes the chi square test for association

### fisher: computes the Fisher exact test

### mcnemar: computes the McNemar test (for the 2x2 tables
only)

### missing.include: if TRUE removes the unused factor
levels

### this table will show the observed counts only
```

```
CrossTable(demo$gender, demo$carcat, digits=3,  
expected=FALSE, prop.r=FALSE, prop.c=FALSE,  
           prop.t=FALSE, prop.chisq=FALSE, chisq = FALSE,  
fisher=FALSE, mcnemar=FALSE,  
missing.include=FALSE)
```

Learn more complex analysis techniques in R (click for a big discount!)

[Take the intermediate course](#)

Become an expert in statistical analysis with R (click for a big discount!)

[Take the advanced course](#)