# Filter & Slice Function

*Pema Lama*

*April 5, 2017*

In this session, we will explore the filter function of dplyr package.

How do I load R-packages using library() function?

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

How do I set the current working directory useing setwd() function?

```
setwd("C:/Users/Owner/Desktop/dplyr")
```

How do I load a data set using read.csv() function?

```
titanic <- read.csv('titanic.csv', stringsAsFactors = TRUE, header = TRUE)
```

How do I see the structure of the data set using str() function?

```
str(titanic)
```

```
## 'data.frame':    1310 obs. of  14 variables:
##  $ pclass   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ survived : int  1 1 0 0 0 1 1 0 1 0 ...
##  $ name     : Factor w/ 1308 levels "","Abbing, Mr. Anthony",..: 23 25 26 27 28 32 47 48 52 56 ...
##  $ sex      : Factor w/ 3 levels "","female","male": 2 3 2 3 2 3 2 3 2 3 ...
##  $ age      : num  29 0.917 2 30 25 ...
##  $ sibsp    : int  0 1 1 1 1 0 1 0 2 0 ...
##  $ parch    : int  0 2 2 2 2 0 0 0 0 0 ...
##  $ ticket   : Factor w/ 930 levels "","110152","110413",..: 189 51 51 51 51 126 94 17 78 827 ...
##  $ fare     : num  211 152 152 152 152 ...
##  $ cabin    : Factor w/ 187 levels "","A10","A11",..: 45 81 81 81 81 151 147 17 63 1 ...
##  $ embarked : Factor w/ 4 levels "","C","Q","S": 4 4 4 4 4 4 4 4 4 2 ...
##  $ boat     : Factor w/ 28 levels "","1","10","11",..: 13 4 1 1 1 14 3 1 28 1 ...
##  $ body     : int  NA NA NA 135 NA NA NA NA NA 22 ...
##  $ home.dest: Factor w/ 370 levels "","?Havana, Cuba",..: 310 232 232 232 232 238 163 25 23 230 ...
```

How do I slice data set according to row values with filter() function?

```
titanic_sub <- filter(titanic, sex == 'female')
head(titanic_sub)
```

```
##   pclass survived                            name    sex
## 1      1        1    Allen, Miss. Elisabeth Walton female
## 2      1        0    Allison, Miss. Helen Loraine female
```

```
## 3      1        0    Allison, Mrs. Hudson J C (Bessie Waldo Daniels) female
## 4      1        1                    Andrews, Miss. Kornelia Theodosia female
## 5      1        1       Appleton, Mrs. Edward Dale (Charlotte Lamson) female
## 6      1        1 Astor, Mrs. John Jacob (Madeleine Talmadge Force) female
##    age sibsp parch   ticket      fare   cabin embarked boat body
## 1  29     0     0    24160 211.3375      B5        S    2   NA
## 2   2     1     2   113781 151.5500 C22 C26        S        NA
## 3  25     1     2   113781 151.5500 C22 C26        S        NA
## 4  63     1     0    13502  77.9583      D7        S   10   NA
## 5  53     2     0    11769  51.4792     C101        S    D   NA
## 6  18     1     0 PC 17757 227.5250 C62 C64        C    4   NA
##                         home.dest
## 1                    St Louis, MO
## 2 Montreal, PQ / Chesterville, ON
## 3 Montreal, PQ / Chesterville, ON
## 4                      Hudson, NY
## 5             Bayside, Queens, NY
## 6                    New York, NY
```

How do I slice a data set by row values of multiple conditions with filter() function?

```
titanic_sub <- filter(titanic, sex == 'female', survived == 1)
head(titanic_sub)
```

```
##    pclass survived                                       name    sex
## 1      1        1                Allen, Miss. Elisabeth Walton female
## 2      1        1              Andrews, Miss. Kornelia Theodosia female
## 3      1        1       Appleton, Mrs. Edward Dale (Charlotte Lamson) female
## 4      1        1 Astor, Mrs. John Jacob (Madeleine Talmadge Force) female
## 5      1        1                 Aubart, Mme. Leontine Pauline female
## 6      1        1                  Barber, Miss. Ellen "Nellie" female
##    age sibsp parch   ticket      fare   cabin embarked boat body
## 1  29     0     0    24160 211.3375      B5        S    2   NA
## 2  63     1     0    13502  77.9583      D7        S   10   NA
## 3  53     2     0    11769  51.4792     C101        S    D   NA
## 4  18     1     0 PC 17757 227.5250 C62 C64        C    4   NA
## 5  24     0     0 PC 17477  69.3000     B35        C    9   NA
## 6  26     0     0    19877  78.8500                S    6   NA
##            home.dest
## 1       St Louis, MO
## 2         Hudson, NY
## 3 Bayside, Queens, NY
## 4       New York, NY
## 5      Paris, France
## 6
```

We can also use (&) to filter out the intersection of two variables.

```
titanic_sub <- filter(titanic, sex == 'female' & survived == 1)
head(titanic_sub)
```

```
##    pclass survived                                       name    sex
## 1      1        1                Allen, Miss. Elisabeth Walton female
## 2      1        1              Andrews, Miss. Kornelia Theodosia female
## 3      1        1       Appleton, Mrs. Edward Dale (Charlotte Lamson) female
## 4      1        1 Astor, Mrs. John Jacob (Madeleine Talmadge Force) female
```

```
## 5       1         1                           Aubart, Mme. Leontine Pauline female
## 6       1         1                           Barber, Miss. Ellen "Nellie" female
##    age sibsp parch    ticket      fare   cabin embarked boat body
## 1  29     0     0     24160 211.3375      B5        S    2   NA
## 2  63     1     0     13502  77.9583      D7        S   10   NA
## 3  53     2     0     11769  51.4792    C101        S    D   NA
## 4  18     1     0 PC 17757 227.5250 C62 C64        C    4   NA
## 5  24     0     0 PC 17477  69.3000     B35        C    9   NA
## 6  26     0     0     19877  78.8500              S    6   NA
##               home.dest
## 1        St Louis, MO
## 2          Hudson, NY
## 3 Bayside, Queens, NY
## 4        New York, NY
## 5        Paris, France
## 6
```

We can slice out multiple conditions in filter() function.

```r
titanic_sub <- filter(titanic, sex == 'female', survived == 1, pclass == 1, sibsp == 0)
head(titanic_sub)
```

```
##    pclass survived                                              name    sex
## 1       1         1                 Allen, Miss. Elisabeth Walton female
## 2       1         1                Aubart, Mme. Leontine Pauline female
## 3       1         1                 Barber, Miss. Ellen "Nellie" female
## 4       1         1 Baxter, Mrs. James (Helene DeLaudeniere Chaput) female
## 5       1         1                        Bazzani, Miss. Albina female
## 6       1         1                        Bidois, Miss. Rosalie female
##    age sibsp parch    ticket      fare   cabin embarked boat body
## 1  29     0     0     24160 211.3375      B5        S    2   NA
## 2  24     0     0 PC 17477  69.3000     B35        C    9   NA
## 3  26     0     0     19877  78.8500              S    6   NA
## 4  50     0     1 PC 17558 247.5208 B58 B60        C    6   NA
## 5  32     0     0     11813  76.2917     D15        C    8   NA
## 6  42     0     0 PC 17757 227.5250              C    4   NA
##        home.dest
## 1  St Louis, MO
## 2 Paris, France
## 3
## 4  Montreal, PQ
## 5
## 6
```

We can separate all the conditions by (&).

```r
titanic_sub <- filter(titanic, sex != 'male' & survived != 0 & pclass == 1 & sibsp == 0)
head(titanic_sub)
```

```
##    pclass survived                                              name    sex
## 1       1         1                 Allen, Miss. Elisabeth Walton female
## 2       1         1                Aubart, Mme. Leontine Pauline female
## 3       1         1                 Barber, Miss. Ellen "Nellie" female
## 4       1         1 Baxter, Mrs. James (Helene DeLaudeniere Chaput) female
## 5       1         1                        Bazzani, Miss. Albina female
## 6       1         1                        Bidois, Miss. Rosalie female
```

```
##   age sibsp parch   ticket     fare   cabin embarked boat body
## 1  29     0     0    24160 211.3375      B5        S    2   NA
## 2  24     0     0 PC 17477  69.3000     B35        C    9   NA
## 3  26     0     0    19877  78.8500               S    6   NA
## 4  50     0     1 PC 17558 247.5208 B58 B60        C    6   NA
## 5  32     0     0    11813  76.2917     D15        C    8   NA
## 6  42     0     0 PC 17757 227.5250               C    4   NA
##        home.dest
## 1  St Louis, MO
## 2 Paris, France
## 3
## 4  Montreal, PQ
## 5
## 6
```

Slicing data set with row index in slice() function.

```
titanic_sub <- slice(titanic, 1:10)
titanic_sub
```

```
##    pclass survived                                         name    sex
## 1       1        1                 Allen, Miss. Elisabeth Walton female
## 2       1        1                Allison, Master. Hudson Trevor   male
## 3       1        0                  Allison, Miss. Helen Loraine female
## 4       1        0        Allison, Mr. Hudson Joshua Creighton   male
## 5       1        0 Allison, Mrs. Hudson J C (Bessie Waldo Daniels) female
## 6       1        1                          Anderson, Mr. Harry   male
## 7       1        1             Andrews, Miss. Kornelia Theodosia female
## 8       1        0                          Andrews, Mr. Thomas Jr   male
## 9       1        1    Appleton, Mrs. Edward Dale (Charlotte Lamson) female
## 10      1        0                          Artagaveytia, Mr. Ramon   male
##         age sibsp parch   ticket     fare   cabin embarked boat body
## 1   29.0000     0     0    24160 211.3375      B5        S    2   NA
## 2    0.9167     1     2   113781 151.5500 C22 C26        S   11   NA
## 3    2.0000     1     2   113781 151.5500 C22 C26        S        NA
## 4   30.0000     1     2   113781 151.5500 C22 C26        S       135
## 5   25.0000     1     2   113781 151.5500 C22 C26        S        NA
## 6   48.0000     0     0    19952  26.5500     E12        S    3   NA
## 7   63.0000     1     0    13502  77.9583      D7        S   10   NA
## 8   39.0000     0     0   112050   0.0000     A36        S        NA
## 9   53.0000     2     0    11769  51.4792    C101        S    D   NA
## 10  71.0000     0     0 PC 17609  49.5042               C        22
##                       home.dest
## 1                  St Louis, MO
## 2  Montreal, PQ / Chesterville, ON
## 3  Montreal, PQ / Chesterville, ON
## 4  Montreal, PQ / Chesterville, ON
## 5  Montreal, PQ / Chesterville, ON
## 6                  New York, NY
## 7                    Hudson, NY
## 8                   Belfast, NI
## 9            Bayside, Queens, NY
## 10          Montevideo, Uruguay
```