

Slicing a data frame in R

Slice data rows in a data frame

Normally, we can easily slice a data frame based on given values in a column. Since data frame can be treated as a list, we can use expression "tbData[[colName]]" to represent a column in a data frame. That makes the program more generic. colName includes the column header information.

```
tbData[tbData[[colName]]==colValue, ]
```

The %in% operator in R is very useful. It can be used to test if there are common elements in two vectors, especially useful in character testing. It can be used to:

- 1) Test if shorter vectors are in longer vectors (ex, 6:10 %in% 1:36);
- 2) Test which elements of long vectors are in short vector (ex, 1:36 %in% 6:10)
- 3) Used in character vectors or factors (ex, c("d", "e") %in% c("a", "b", "c", "d"))

If you want to know the indexes of the specific elements inside a larger vector, use function "which" to achieve this goal. Find the indexes of which elements of (1:36 %in% 1:6) are: which(1:36 %in% 6:10).

By using this function, we can slice a data frame based on a given string or string vector.

```
myDf[myDf$contrast %in% compstr, ]
```

Slice data columns in a data frame

In R slicing data columns in a data frame is pretty simple. Either using the index information of columns or directly using column names.

Using index or position of columns to slice the data frame is more generic.

```
df[c(1, 4:ncol(df))]
```

Using columns to slice the data frame is useful when you know the column headers before you slice it.

```
df[c('c1', 'c5', 'c10')]
```