### data.table

**data.table** is an R package that provides a high-performance version of base R's `data.frame` with syntax and feature enhancements for ease of use, convenience and programming speed.

#### Load the package:

```r
> library(data.table)
```

### Creating A data.table

```r
> set.seed(123) create a data table and call it DF
> DF <- data.table(x=1:5).y=LETTERS[1:5]

### Subset rows using i

```r
> DF[3:5] subset rows from 3 to 5
> DF[-4] subset from 1 to 3
> DF[which(x>2)] subset rows that have value greater than 2
> DF[V1!=2] subset rows that have value 1 or 2 in the second column
```

### Manipulating on Columns in j

```r
> DF[, x2:=x*2] duplicate x and add it as a new column
> DF[, .SDcols=LETTERS[1:5]] keep the columns that are specified
```

### Chaining

```r
> DT <- .(DF[, .SD[row=1]]).y=LETTERS[1:5] calculate sum of y, grouped by i
> DT[, .SD] subset rows in which the sum is > 0
> DT[, V1:= as.numeric(V1)] cast column V1 to numeric type
> DT[, V1:= V1 V2] generate a new column V1

### Doing j by Group

```r
> DT[, .(V1=sum(V1),V2=sum(V2)), V1=1] sum of v grouped by v1
> DT[, .(mean(V1),median(V1))) V1=1] average and median of v1 grouped by v1
> DT[, .SD[3:5,1]] keep the first three rows
```

### Adding/Updating Columns By Reference in j Using :=

```r
> DT[, .(V1=V1+10,,V2=V2+10), V1=1] add 10 to all values in column V1
> DT[, .(V1=x V2=x), V1=1] divide by 2
```

### Advanced Data Table Operations

```r
> DT[, .SD[1:2], V1=1] subset the first two columns grouped by V1
> DT[, .SD[3:5,1], V1=1] subset the third to fifth columns grouped by V1
```

### Indexing and Keys

```r
> setkey(DF,V2) an easy way set on V2, output is returned invisibly
> DT[.i] subset all rows where the key column (set to V2) is 2 or the value A
```

### set() - Family

**set()**

```r
> DF[1:3,.SDcols=LETTERS[1:5]] keep the columns that are specified

**setnames()**

```r
> setnames(DF,.cols=LETTERS[1:5]) give a new name for each column
```

**setorder()**

```r
> setorder(DF,.key=V1) order by V1
```

The Cheat Sheet is available for download at [DataCamp](https://www.datacamp.com).