Creating Data Tables in Rstudio

```
# USING R CODE TO CREATE DATA TABLES (Typing content for Data Tables
# directly into Rstudio)

# (TWO METHODS)

# METHOD 1  VECTOR CONSTRUCTION DATA FRAME METHOD

# EXAMPLE


 Name <- c( "Jacob", "Elaine", "Alice", "Juan", "Ray", "Kate", "Leon")
 Name

## [1] "Jacob"  "Elaine" "Alice"  "Juan"   "Ray"    "Kate"   "Leon"

 Age <- c(26,31,42,31,28,25,30)
 Age

## [1] 26 31 42 31 28 25 30

 Department <- c("Accounting", "IT", "Sales", "IT", "Accounting", "Sales",
                 "Personnel")
 Department

## [1] "Accounting" "IT"         "Sales"      "IT"         "Accounting"
## [6] "Sales"      "Personnel"

 Salary <- c(70000, 75000, 72000, 68000, 67500, 68000, 63000)
 Salary

## [1] 70000 75000 72000 68000 67500 68000 63000

 data.frame(Name, Age, Department, Salary)

##     Name Age Department Salary
## 1  Jacob  26 Accounting  70000
## 2 Elaine  31         IT  75000
## 3  Alice  42      Sales  72000
## 4   Juan  31         IT  68000
## 5    Ray  28 Accounting  67500
## 6   Kate  25      Sales  68000
## 7   Leon  30  Personnel  63000

 # Now assign your data table a variable name

 data.frame(Name, Age, Department, Salary) ->  EmployeeInformation
 EmployeeInformation
```

```
##     Name Age Department Salary
## 1  Jacob  26 Accounting  70000
## 2 Elaine  31         IT  75000
## 3  Alice  42      Sales  72000
## 4   Juan  31         IT  68000
## 5    Ray  28 Accounting  67500
## 6   Kate  25      Sales  68000
## 7   Leon  30  Personnel  63000
```

```
# You can now use Base R commands to generate desired variable
```

```
 mean(EmployeeInformation$Age)
```

```
## [1] 30.42857
```

```
 median(EmployeeInformation$Salary)
```

```
## [1] 68000
```

```
# METHOD 2  TRIBBLE CONSTRUCTION (TIDYVERSE METHOD)
```

```
 library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.0.5
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.
3.1 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.5
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
## Warning: package 'tidyr' was built under R version 4.0.5
```

```
## Warning: package 'readr' was built under R version 4.0.5
```

```
## Warning: package 'forcats' was built under R version 4.0.5
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflict
s() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
 tribble(~Name,  ~Age,  ~Department, ~Salary,
         "Jacob",  26, "Accounting", 70000,
         "Elaine", 31, "IT",         75000,
         "Alice",  42, "Sales",      72000,
         "Juan",   31, "IT",         68000,
         "Ray",    28, "Accounting", 67500,
```

```
        "Kate",   25, "Sales",      68000,
        "Leon",   30, "Personnel",  63000 )
```

```
## # A tibble: 7 x 4
##   Name     Age Department Salary
##   <chr>  <dbl> <chr>       <dbl>
## 1 Jacob     26 Accounting  70000
## 2 Elaine    31 IT          75000
## 3 Alice     42 Sales       72000
## 4 Juan      31 IT          68000
## 5 Ray       28 Accounting  67500
## 6 Kate      25 Sales       68000
## 7 Leon      30 Personnel   63000
```

```
 # Now assigin the data table to a variable (Copy and paste; Do not retype
 # the table)


 tribble(~Name,  ~Age,  ~Department, ~Salary,
        "Jacob",  26, "Accounting", 70000,
        "Elaine", 31, "IT",         75000,
        "Alice",  42, "Sales",      72000,
        "Juan",   31, "IT",         68000,
        "Ray",    28, "Accounting", 67500,
        "Kate",   25, "Sales",      68000,
        "Leon",   30, "Personnel",  63000 ) ->  EmployeeInformation2
 EmployeeInformation2
```

```
## # A tibble: 7 x 4
##   Name     Age Department Salary
##   <chr>  <dbl> <chr>       <dbl>
## 1 Jacob     26 Accounting  70000
## 2 Elaine    31 IT          75000
## 3 Alice     42 Sales       72000
## 4 Juan      31 IT          68000
## 5 Ray       28 Accounting  67500
## 6 Kate      25 Sales       68000
## 7 Leon      30 Personnel   63000
```

```
 # We can now generate desired summaries

 mean(EmployeeInformation2$Age)
```

```
## [1] 30.42857
```

```
 max(EmployeeInformation2$Salary)
```

```
## [1] 75000
```

```
 sd(EmployeeInformation2$Salary)
```

```
## [1] 3790.653
```