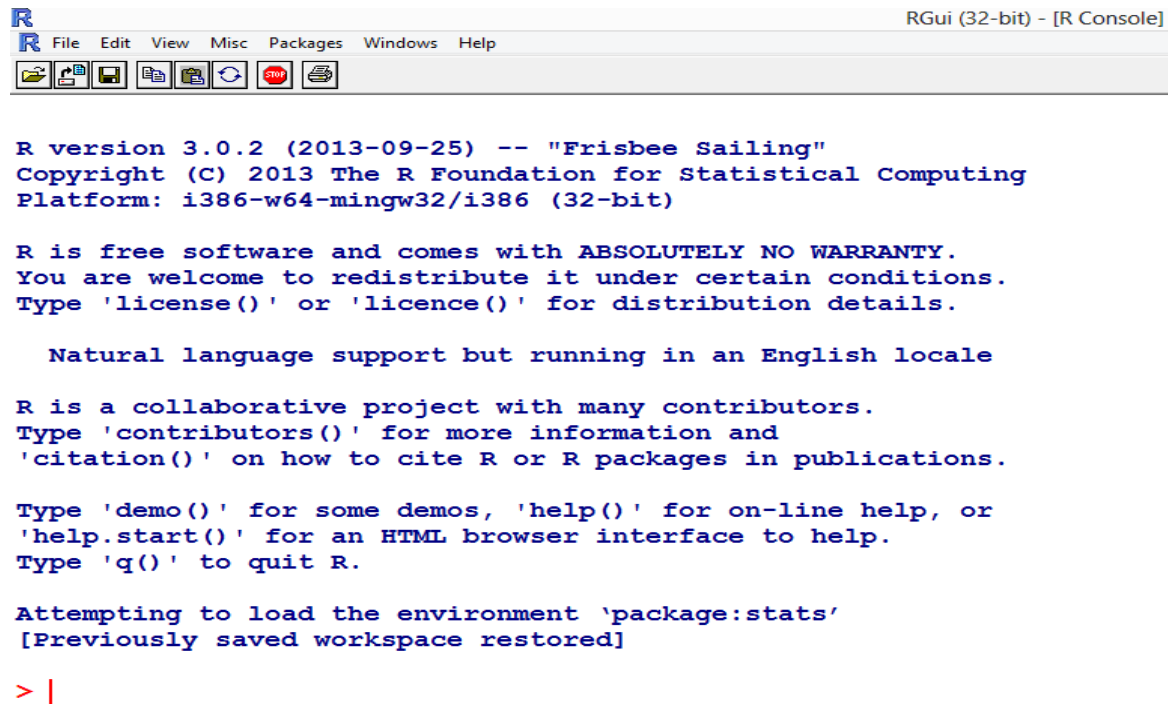




# Getting Started With R

# Data Science: Getting Started with R



```
RGui (32-bit) - [R Console]

R version 3.0.2 (2013-09-25) -- "Frisbee Sailing"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Attempting to load the environment 'package:stats'
[Previously saved workspace restored]

> |
```

R is an **open-source** software program, developed by volunteers as a service to the community of scientists, researchers, and data analysts who use it. R is **free** to download and use. **Lots of advice is available** online to help users learn R, which is good because it is a powerful and complex program, in reality a full-featured programming language dedicated to data.

# Some Basic Information on R

- Command line oriented
- Not especially good at giving feedback or error messages
- One Needs to know the data

# An Introduction to Data: Strings

**Character string**



**“this is a piece of text”**

011101000110100001101001011100110010000001101001011100110010  
000001100001001000000111000001101001011001010110001101100101  
001000000110111101100110001000000111010001100101011110000111  
0100



**R code:**

**myText <- “this is a piece of text”**

# An Introduction to Data: Type / Mode

**List of integers**

**43, 42, 12, 8, 5**

*Each integer represents the age of a family member.*

**Integer list is all the same “type/mode.”**

R refers to a list as a “vector.”

**R code for this vector looks like**

```
c (43, 42, 12, 8, 5)
```

```
myFamilyAges <- c (43, 42, 12, 8, 5)
```

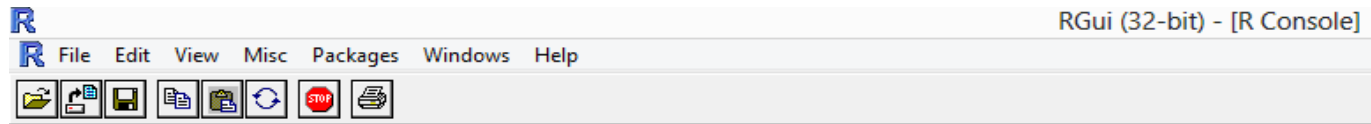
**Data set**



# Getting Started in R

- Start up R
- Create/key in at the command prompt '>'  
    > MyText<- "this is a piece of text"  
    Retrieve MyText  
    Retrieve mytext
- Create/key in at the command prompt '>'  
    > myFamilyAges <-c(**43, 42, 12, 8, 5**)  
    Retrieve myFamilyAges
- End R .... q() at the command prompt '>'
- Respond 'yes' to save your work space

# The Initial Console View



```
R version 3.0.2 (2013-09-25) -- "Frisbee Sailing"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-bit)
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```

```
  Natural language support but running in an English locale
```

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```

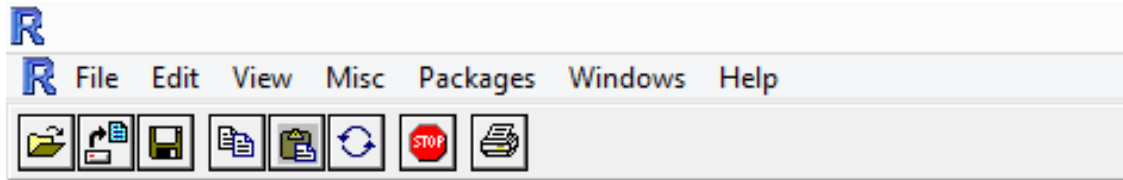
```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

```
Attempting to load the environment 'package:stats'
[Previously saved workspace restored]
```

```
> |
```

# Clear the Console

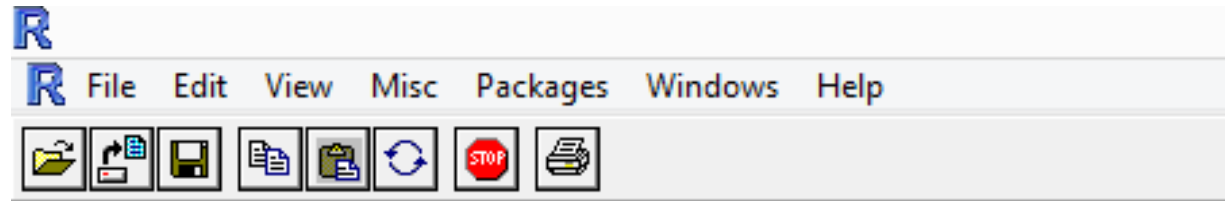
Edit → Clear Console or Control L



> |

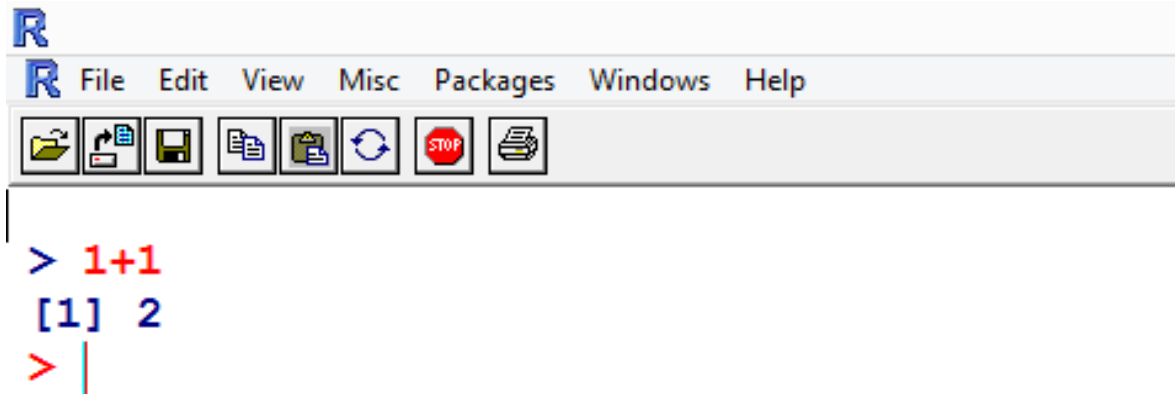


# Adding Two Numbers



```
> 1+1|
```

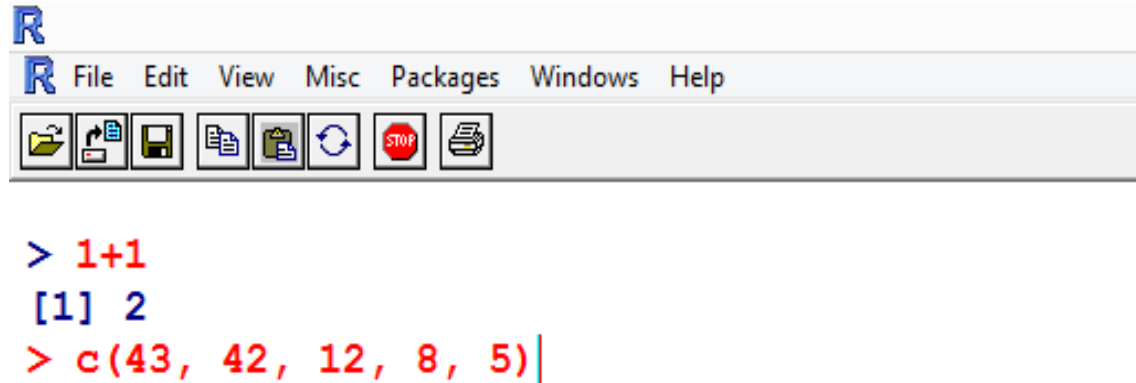
# Adding Two Numbers



The image shows a screenshot of the R console interface. At the top, there is a menu bar with the following options: File, Edit, View, Misc, Packages, Windows, and Help. Below the menu bar is a toolbar with icons for file operations (open, save, print, etc.) and a red stop button. The console area displays the following text:

```
> 1+1  
[1] 2  
> |
```

# Creating a Vector

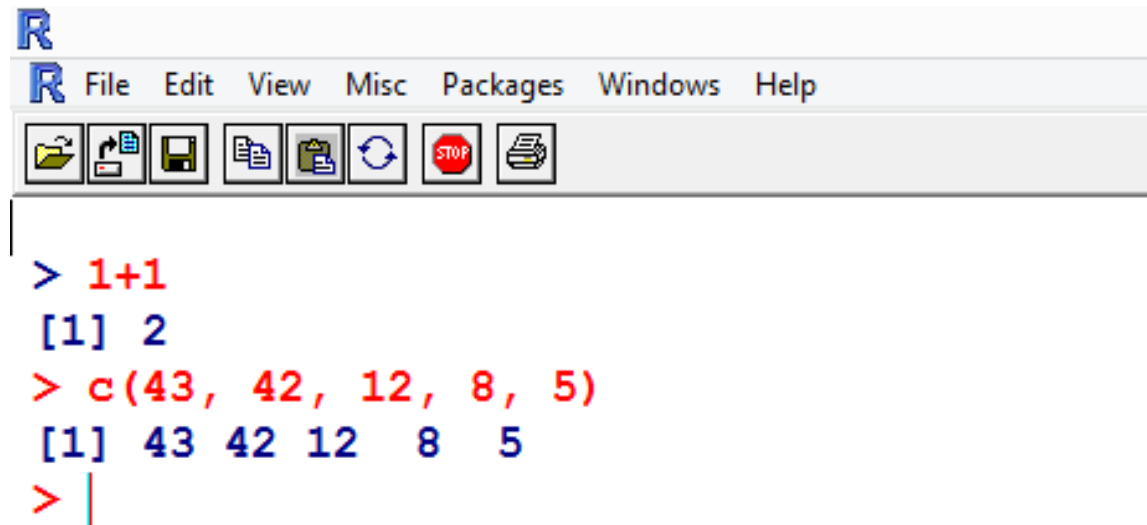


The screenshot shows the R console window. The title bar includes the R logo and menu items: File, Edit, View, Misc, Packages, Windows, and Help. Below the menu bar is a toolbar with icons for file operations (Open, Save, Print, etc.). The console area displays the following commands and output:

```
> 1+1  
[1] 2  
> c(43, 42, 12, 8, 5)|
```

Note that the 'c' command stands for concatenate or combine

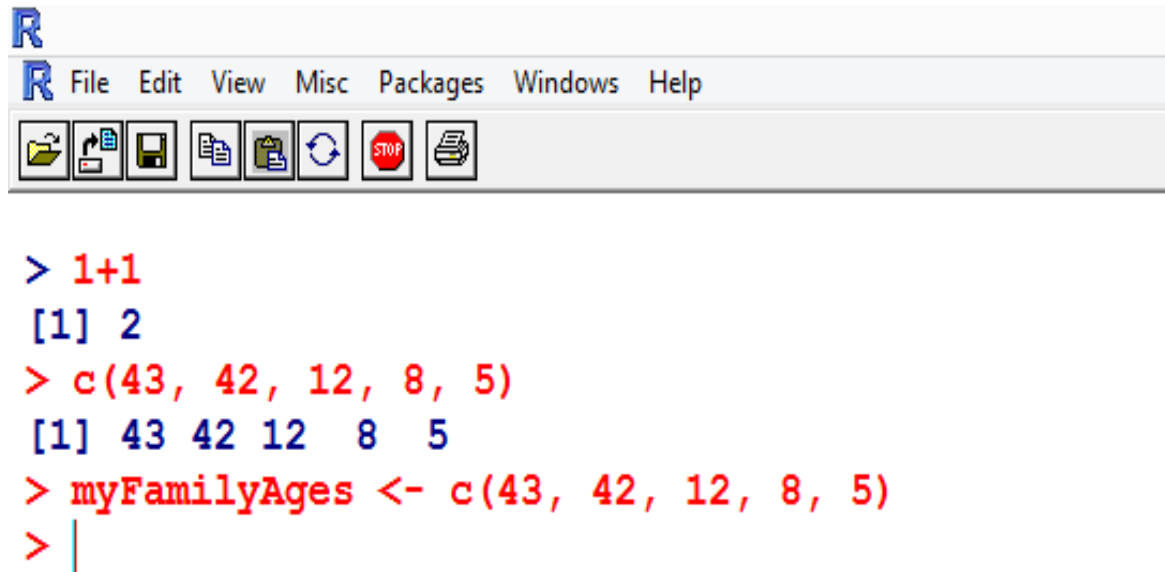
# Creating a Vector



The image shows a screenshot of the R console window. The window has a title bar with the R logo and a menu bar with options: File, Edit, View, Misc, Packages, Windows, and Help. Below the menu bar is a toolbar with icons for file operations (open, save, print, etc.) and a stop button. The console area displays the following R commands and their outputs:

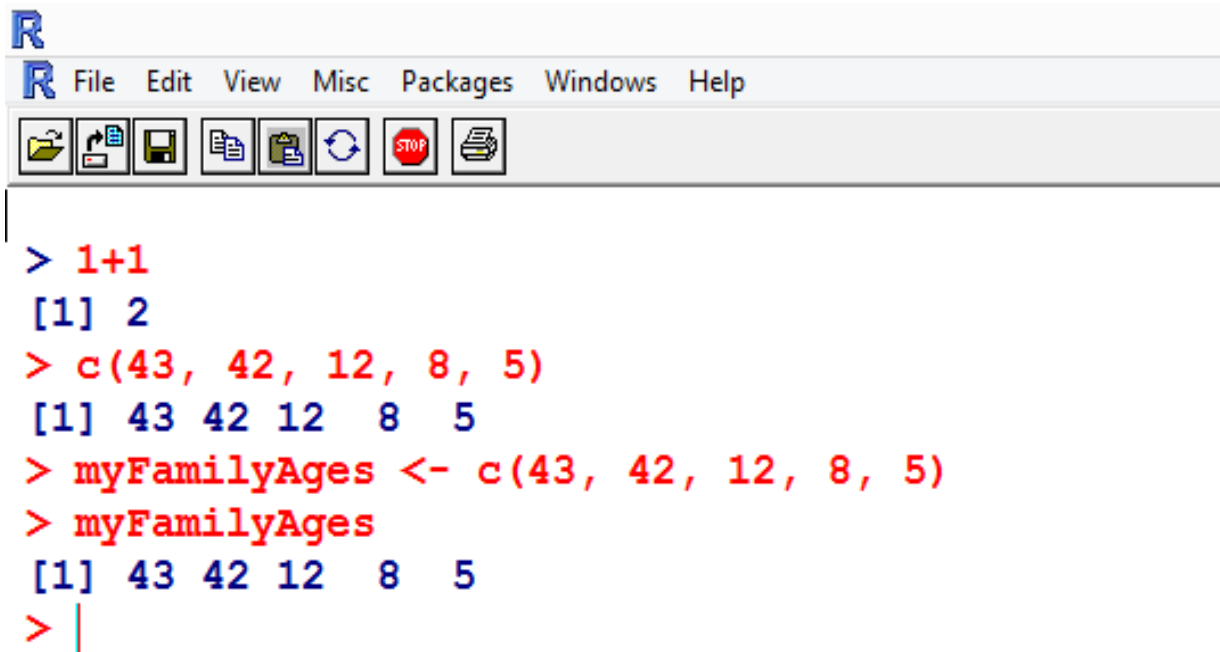
```
> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> |
```

# Storing a Vector

A screenshot of the R console window. The window has a title bar with the R logo and a menu bar with options: File, Edit, View, Misc, Packages, Windows, and Help. Below the menu bar is a toolbar with icons for file operations (open, save, print, etc.). The console area shows the following commands and output:

```
> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> |
```

# Storing a Vector

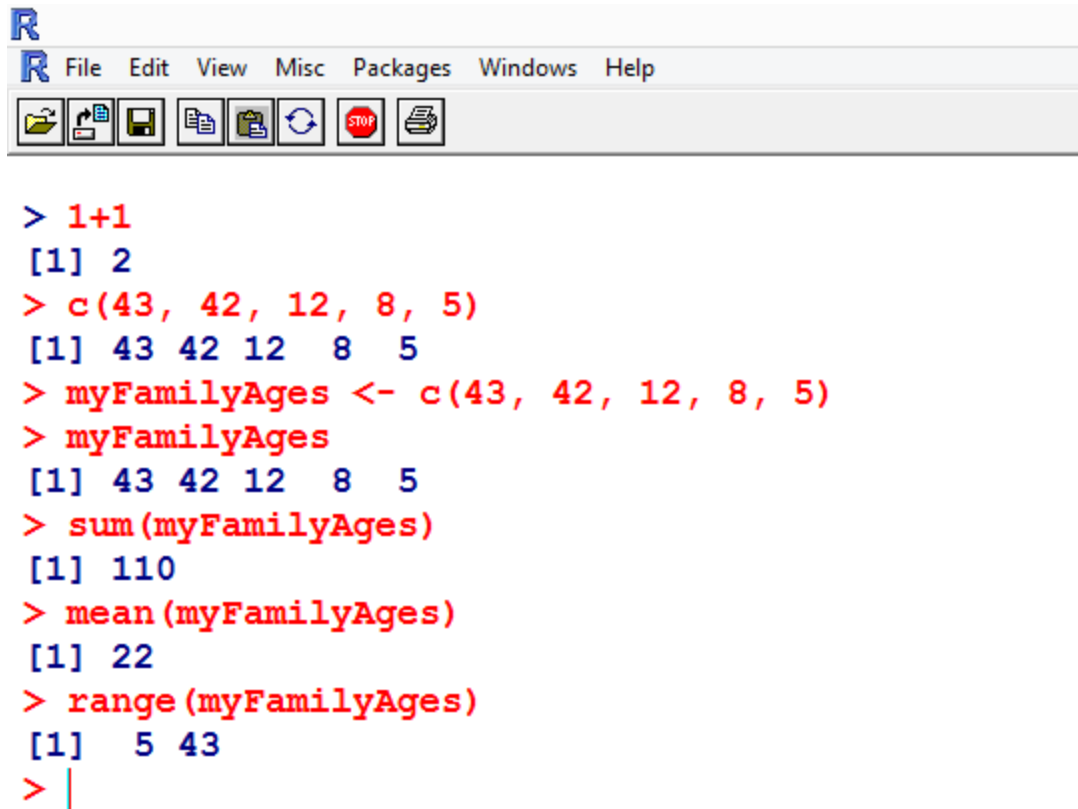


The image shows a screenshot of the R console window. The window has a title bar with the R logo and a menu bar with options: File, Edit, View, Misc, Packages, Windows, and Help. Below the menu bar is a toolbar with icons for file operations (open, save, print, etc.). The console area displays the following R code and its output:

```
> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> |
```

A blue arrow points from the left towards the console output for the `myFamilyAges` variable.

# Some Simple Functions on a Vector



```
R
R File Edit View Misc Packages Windows Help
[Icons: Open, Save, Print, Copy, Paste, Undo, Redo, Stop, Run]

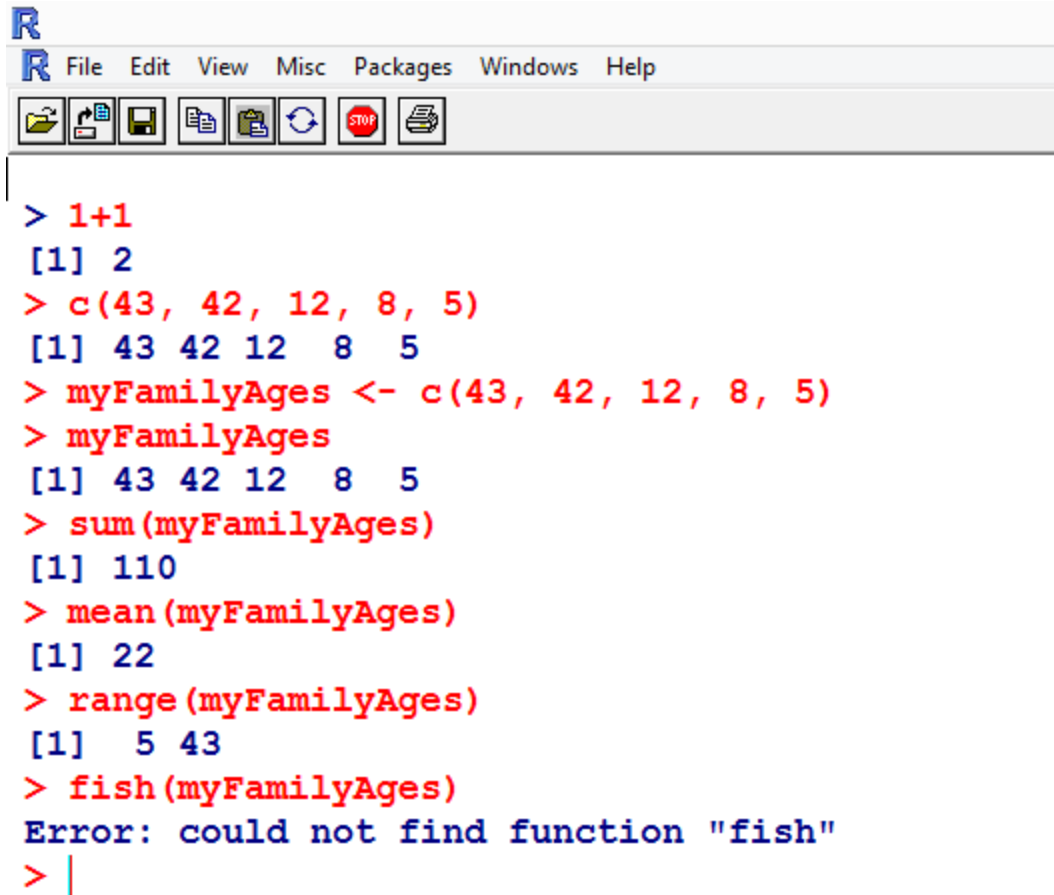
> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> |
```

**sum**

**mean**

**range**

# Errors in R



The screenshot shows the R console window with the following content:

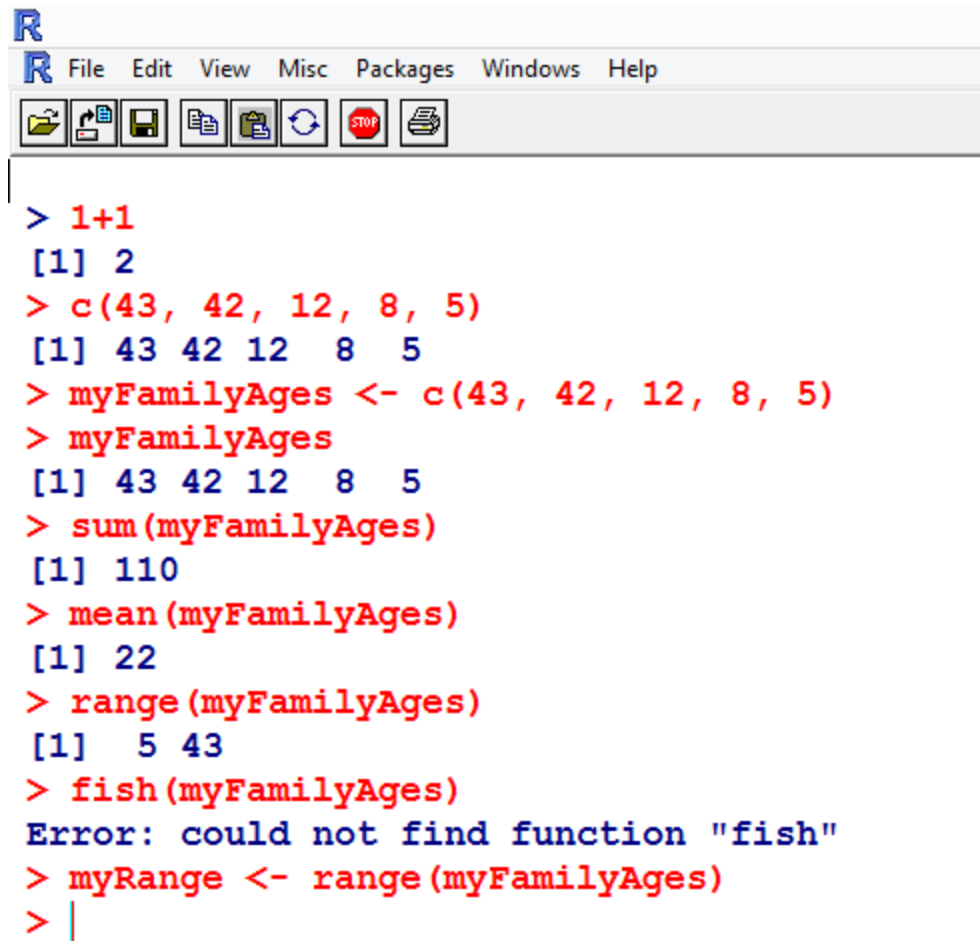
```
R
R File Edit View Misc Packages Windows Help
[Icons: File Explorer, Copy, Paste, Print, Undo, Redo, Stop, Run]

> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> fish(myFamilyAges)
Error: could not find function "fish"
> |
```

A blue arrow points to the error message line.



# Storing Function Results

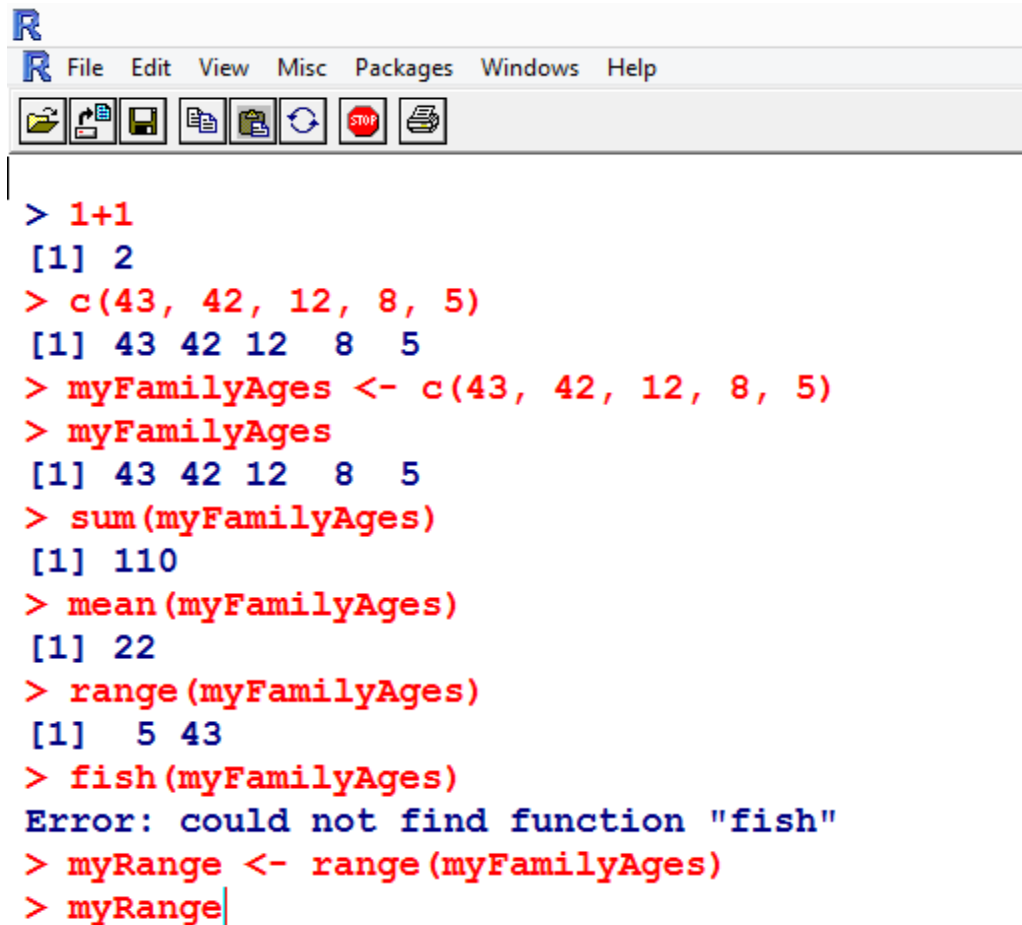


```
R
File Edit View Misc Packages Windows Help
[Icons: Open, Save, Print, Copy, Paste, Undo, Redo, Stop, Run]

> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> fish(myFamilyAges)
Error: could not find function "fish"
> myRange <- range(myFamilyAges)
> |
```

A blue arrow points to the final command in the R console window.

# Storing Function Results




```
R
R File Edit View Misc Packages Windows Help
[Icons: Open, Save, Print, Copy, Paste, Undo, Redo, Stop, Run]

> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> fish(myFamilyAges)
Error: could not find function "fish"
> myRange <- range(myFamilyAges)
> myRange|
```



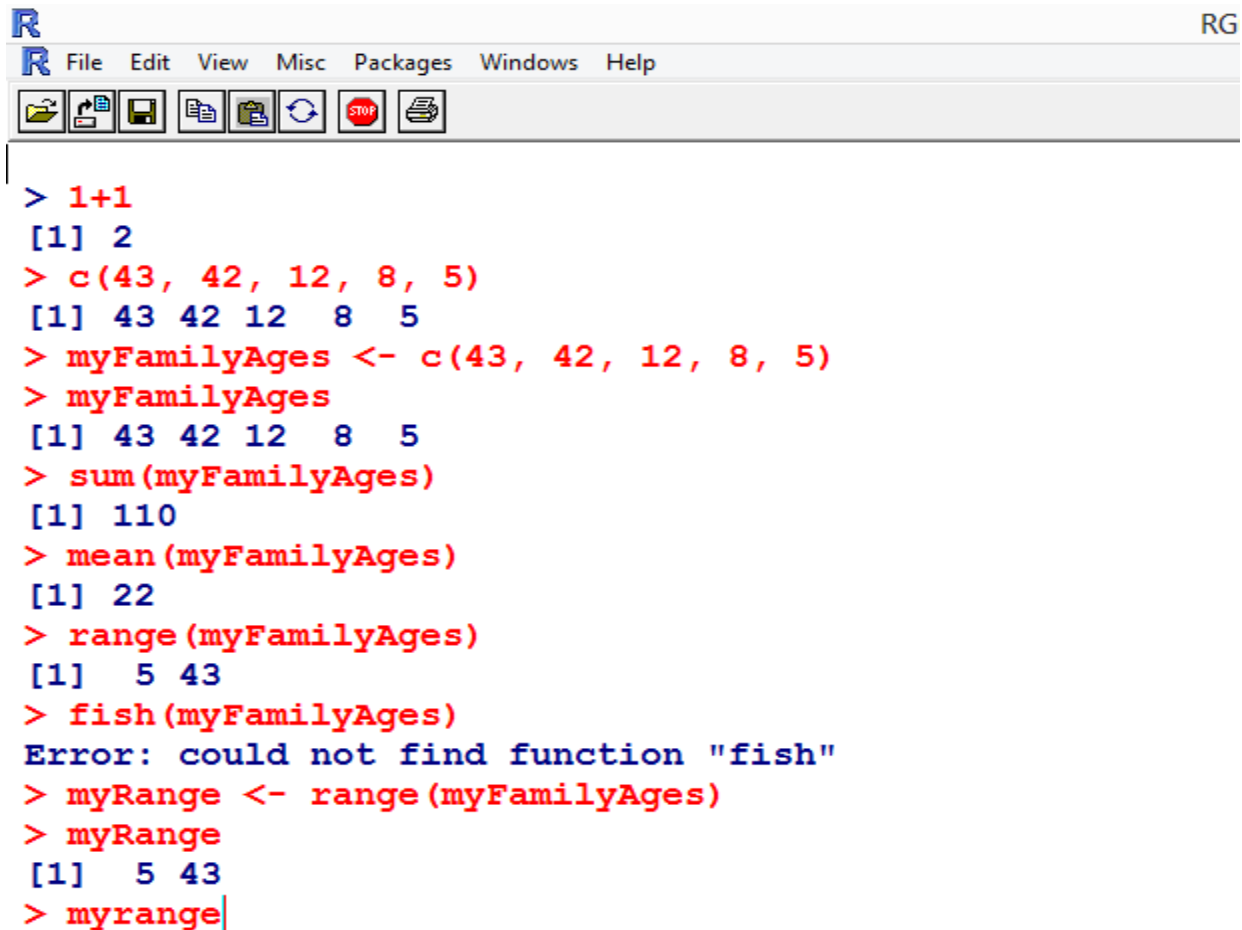
# Storing Function Results



```
R
File Edit View Misc Packages Windows Help

> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> fish(myFamilyAges)
Error: could not find function "fish"
> myRange <- range(myFamilyAges)
> myRange
[1] 5 43
> |
```

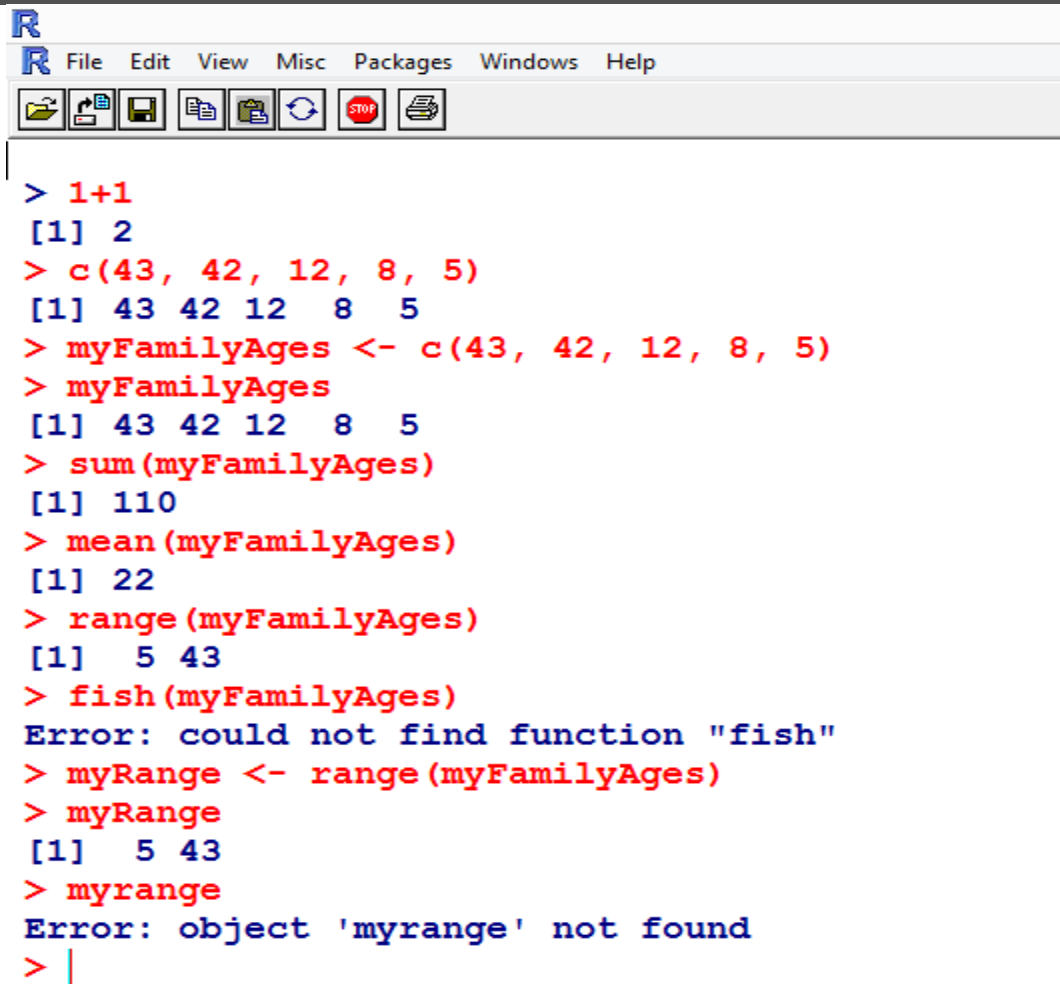
# Caps are NOT Ignored



```
R
File Edit View Misc Packages Windows Help
[Icons]

> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> fish(myFamilyAges)
Error: could not find function "fish"
> myRange <- range(myFamilyAges)
> myRange
[1] 5 43
> myrange|
```

# Caps are NOT Ignored



```
R
R File Edit View Misc Packages Windows Help
[Icons: Open, Save, Print, Run, Stop, Help]

> 1+1
[1] 2
> c(43, 42, 12, 8, 5)
[1] 43 42 12 8 5
> myFamilyAges <- c(43, 42, 12, 8, 5)
> myFamilyAges
[1] 43 42 12 8 5
> sum(myFamilyAges)
[1] 110
> mean(myFamilyAges)
[1] 22
> range(myFamilyAges)
[1] 5 43
> fish(myFamilyAges)
Error: could not find function "fish"
> myRange <- range(myFamilyAges)
> myRange
[1] 5 43
> myrange
Error: object 'myrange' not found
> |
```

# Some Takeaways

- The use of the “c()” function.
- The vector is a basic form of data storage.
- Numbers & text can be collected in lists (Vector)
- A vector can be stored in a named location using the arrow “<-” (e.g. myFamilyAges)
- You can get the data object a named location by typing the name.

# Some Takeaways

- Vector has a length (number of items in it).
- Vector has a mode (type of data).
- Capitalization matters.
- Legal functions include sum, mean, range, and fish is not a legal function.
- Typically more than one variable or data set is used for analysis.

# Data Science

## Data Science Fun Meter







School of Information Studies  
**SYRACUSE UNIVERSITY**