Factor:

In a data frame, character variables are automatically changed or converted into factor, and the number of levels can be determined as the number of different values in such a vector.

Factor takes a limited number of different values, such variables are referred to as categorical variables. So, Factor represents the categorical data, the factor can be ordered or unordered and are an important class for statistical analysis and for plotting. Factor variables are very useful to many different types of graphics.

Storing data factors insures that the modeling functions will treat such data correctly. The factor can store both integers and strings. These are very useful in the columns which have a limited number of unique values such as “Male, Female” and “True, False” etc.

Factors in R has two varieties

- ordered
- unordered.

Factors are stored as a vector of integer values, with a corresponding set of character values to use when the factor is shown. factor() function is used to create a factor. The required argument to factor is a vector of values, which will be returned as a vector of factor values. Numeric and Character variables both can be made into factors, but a factor’s levels will always be character values.

Factor levels

Getting a dataset you will look that it contains factors with specific factor levels. By the way, sometimes you will willing to change the names of these levels for clarity or any other reasons. R permits you to do this with the function levels().

Examples:

for this is mtcars data:

```
>Str (mtcars)
$ cyl : num 6 6 4 6 8 6 8 4 4 6
```

Here, have shown an example, suppose this is mtcars data where has 32 car brands. Each of those cars we have eleven attributes like horsepower, cylinder, displacement, mileage or all those things. If you see the cylinder, either it has 6 as cylinder or 4 as cylinder or 8 as a cylinder. So, since if you have a minimum number of unique value of particular attributes then that’s an ideal candidate for factors. Because it does not take 6.5 or 4.32 or 5.67. It takes either 4 or 6 or 8 so we want to change this to a factor.

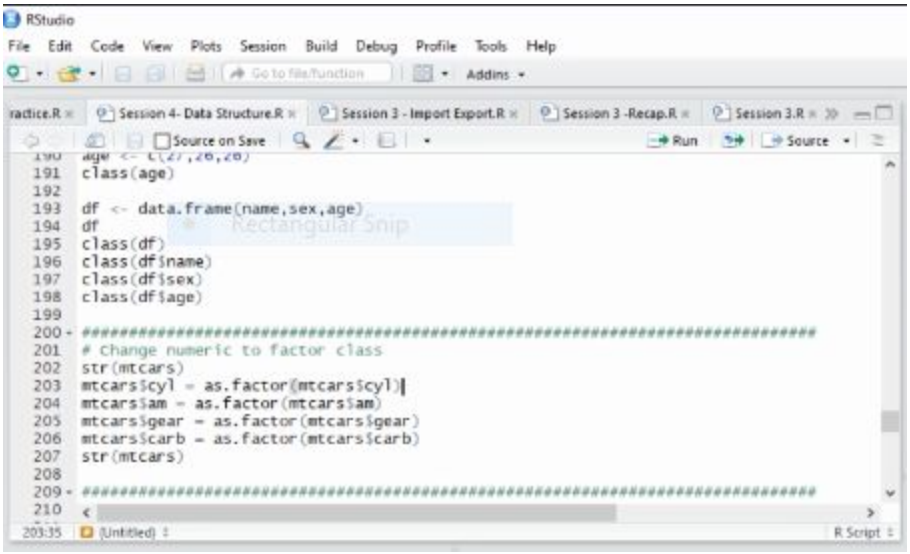
```
Str (mtcars)
mtcars$cyl = as.factor(mtcars$cyl)
```

How to use factor() function? Just use as.factor() function, here the first query is which one you want to change? Suppose, here change the “mtcars$cylinder”.You can access it any column using the dollar function. Using that “as.factor()” function,the “mtcars$cylinder” is converting into Factor.

```
Str (mtcars)
$ cyl : num 6 6 4 6 8 6 8 4 4 6
```

Now, looking carefully at the above example, look the structure of “mtcars” when it changes to factor then everything is changed into the numeric function. Changing this when you look at the cylinder, it is converted to a factor. After changing it has three level 4,6 and 8.

str (mtcars)
(four column or four attributes change to numericals factors)
$ am : factor w/2 levels “0” , “1” : 2 2 2 1 1 1 1 1 1 1 . . .
If it is changed whether manually or automatically. What is the number of gear of the mtcars or how many numbers of carburetors is there, then you see the number of the structure of mtcars. Now you see all those columns or all those attributes as change into factor. You can notice that automation has two levels if it is manual or automatic.



str (mtcars)
(four column or four attributes change to numericals factors)
$ gear : factor w/3 levels “3” , “4” , “5” : 2 2 2 1 1 1 1 2 2 2 . . .
You have three gears levels.
str (mtcars)
(four column or four attributes change to numericals factors)
$ carb: factor w/6 levels “1” , “2” , “3” , “4”,. . : 4 4 1 1 2 1 4 2 2 4. . .
Similarly, you have 6 levels of carbs. That’s how you can change it.
How to change the name of the level?
If you change the name of the level, then you have to create a new variable called gender vector and then you want to store “Male”, “Female” , “Female” , “Male” , “Male”. It may be a data for five employees. Thereafter, you see that the five things have been recorded. If you see the class of gender vector then it is a character vector.
gender vector <-c (“Male”, “Female” , “Female” , “Male” , “Male”)
gender vector
class (gender vector)

- But, if you want to change the gender vector to factor

#Convert gender vector to factor
factor-gender-vector <-as.factor(gender vector)
factor-gender-vector # factor gender has two levels Male and Female

then use “as.factor” and then see it has changed to a factor and it’s showing the level is “Female Male”.
Now you want to change the name of the factors using level()
levels(factor-gender-vector)<-c(“f” , “M”)
levels: Female and Male(earlier)
Now, it changes to “F & M”
That’s how you want to do it.
How to do that? In this case, level() function helps you. Using level() function which you want to change, just give this. Suppose you want to change “factor-gender and vector”. Here suppose you want to change F and M for Female and Male. once you can do it and you can again see this, this levels have changed. Previously it showed Female and Male now it changes to F and M. in this process you want to do it.
This brings an end to this post, I encourage you to re read the post to understand it completely if you haven’t and THANK YOU.