DATA FRAMES

```
>
pd=data.frame("Name"=c("senthil","senthil","sam","sam"),"Month"=c("jan","f
eb","jan","feb"),"BS"=c(140.2,139.4,138.3,142.2),"BP"=c(90,88,87,85))

> pd

     Name Month    BS BP
1 senthil   jan 140.2 90
2 senthil   feb 139.4 88
3     sam   jan 138.3 87
4     sam   feb 142.2 85


> print(pd[1:2,])

     Name Month    BS BP
1 senthil   jan 140.2 90
2 senthil   feb 139.4 88

> print(pd[,1:2])
     Name Month
1 senthil   jan
2 senthil   feb
3     sam   jan
4     sam   feb

> print(pd[1:2])

     Name Month
1 senthil   jan
2 senthil   feb
3     sam   jan
4     sam   feb


CREATE A DATA FRAME USING DATA FROM A FILE
>newdf=read.table(path="PATH OF THE FILE")
IN THE PATH NEED TO USE '/' INSTEAD OF '\'

>

> pd2=subset(pd,Name="senthil"|BP>80)

> pd2

     Name Month    BS BP
1 senthil   jan 140.2 90
2 senthil   feb 139.4 88
3     sam   jan 138.3 87
4     sam   feb 142.2 85

> pd2=subset(pd,Name="senthil"|BP>86)

> pd2

     Name Month    BS BP
1 senthil   jan 140.2 90
2 senthil   feb 139.4 88
3     sam   jan 138.3 87
4     sam   feb 142.2 85

>
```

```
> mytable=data.frame()
> mytable=edit(mytable)
> mytable
```

```
>
pd=rbind(pd,data.frame(Name="RAM",Month="April",BS=139.0,BP=88))
> pd
     Name Month    BS BP
1 senthil   jan 140.2 90
2 senthil   feb 139.4 88
3     sam   jan 138.3 87
4     sam   feb 142.2 85
5     RAM April 139.0 88
```

```
> pd=cbind(pd,addr=c("hyd","hyd","sec","sec","srnagar"))
> pd
     Name Month    BS BP    addr
1 senthil   jan 140.2 90     hyd
2 senthil   feb 139.4 88     hyd
3     sam   jan 138.3 87     sec
4     sam   feb 142.2 85     sec
5     RAM April 139.0 88 srnagar
```

Delete ROW and COLUMN

```
> pd
     Name Month    BS BP    addr
1 senthil   jan 140.2 90     hyd
2 senthil   feb 139.4 88     hyd
3     sam   jan 138.3 87     sec
4     sam   feb 142.2 85     sec
5     RAM April 139.0 88 srnagar

> pd1=pd[-5,-5]
> pd1
     Name Month    BS BP
1 senthil   jan 140.2 90
2 senthil   feb 139.4 88
3     sam   jan 138.3 87
4     sam   feb 142.2 85

>
```

how to manipulate the rows in the data frame and what is called as a factory issue. R has inbuilt characteristic to assign the data types to the data you enter.When you enter numeric variables, it knows all the numeric variables that are available when you enter character variables it takes whatever the character variables you are giving as categories or factors levels.

```
> vec1=c(1,2,3)
> vec1
[1] 1 2 3
> vec2=c("c","R","java")
> vec2
[1] "c"    "R"    "java"
> vec3=c("prodedural","STATISTICAL","scaleup")
> vec3
[1] "prodedural"  "STATISTICAL" "scaleup"
> df=data.frame(vec1,vec2,vec3)
> df
```

```
  vec1 vec2        vec3
1    1    c  prodedural
2    2    R STATISTICAL
3    3 java     scaleup
> df[3,1]=3.1
> df
  vec1 vec2        vec3
1  1.0    c  prodedural
2  2.0    R STATISTICAL
3  3.1 java     scaleup
> df[3,3]="OOPS"
Warning message:
In `[<-.factor`(`*tmp*`, iseq, value = "OOPS") :
  invalid factor level, NA generated
> df
  vec1 vec2        vec3
1  1.0    c  prodedural
2  2.0    R STATISTICAL
3  3.1 java        <NA>
```

And it assumes that these are the only factors that are available for now; when you want to change the element in the third row third column to others; what happens is it will display warning message saying that, this others categorical variable is not available and it replaces that with the NA you can notice that the place where we want others to be there we are having a NA and we can also see the use of word factor in the warning message, how to get rid of the factor issue is the question now.

```
> df=data.frame(vec1,vec2,vec3,stringsAsFactors = F)
> df[3,3]="OOPS"
> df
  vec1 vec2        vec3
1    1    c  prodedural
2    2    R STATISTICAL
3    3 java        OOPS

>
```

New entries in R when you are entering should be consistent with the factor levels that are already defined if not those error message will be printed out. If you do not want this issue to happen what you have to do is while defining the data from itself you need to pass another argument, which says strings as factors is false by default this argument is true that is the reason why you get this warning message when you want to change the string characters into new string characters as an element..

Now try doing the same manipulation you want to change the third row third element to others and print the data frame you can see that there is no NA anymore and we achieved what we want.