

EXP 3:

- a. Explain what is factor. If `z <- factor(c("p", "q", "p", "r", "q"))` and levels of z are ("p", "q", "r")
- write an R expression that will change the level "p" to "x", "q" to "y", "r" to "z".
 - write an R expression that will change the level "p" to "w" so that z is equal to: "w", "q", "w", "r", "q".
 - Write a r program to convert given PH levels of soil into ordered factor. consider the following vector as input
`soil_PH=c(3,7,10,3,4,5,3,4,2,6,10,4,5,3)`
 - Write a program to create a factor for Agriculture field corresponding to swiss dataset in Display its structure.

```
> z=factor(c("p", "q", "p", "r", "q"))
> print(z)
[1] p q p r q
Levels: p q r
>
> levels(z)=c("x","y","z")
> print(z)
[1] x y x z y
Levels: x y z
>
> levels(z)=c("w","p","r")
> print(z)
[1] w p w r p
Levels: w p r
>
>
> oil_PH=c(3,7,10,3,4,5,3,4,2,6,10,4,5,3)
> print(oil_PH)
[1] 3 7 10 3 4 5 3 4 2 6 10 4 5 3
>
> fc=factor(oil_PH,ordered = TRUE)
> print(fc)
[1] 3 7 10 3 4 5 3 4 2 6 10 4 5 3
Levels: 2 < 3 < 4 < 5 < 6 < 7 < 10
>
> print(swiss$Agriculture)
[1] 17.0 45.1 39.7 36.5 43.5 35.3 70.2 67.8 53.3 45.2 64.5 62.0 67.5 60.7 69.3 72.6 34.0 19.4 1
5.2
[20] 73.0 59.8 55.1 50.9 54.1 71.2 58.1 63.5 60.8 26.8 49.5 85.9 84.9 89.7 78.2 64.9 75.9 84.6 6
3.1
[39] 38.4 7.7 16.7 17.6 37.6 18.7 1.2 46.6 27.7
> 11k=factor(c(swiss$Agriculture))
> print(11k)
[1] 17 45.1 39.7 36.5 43.5 35.3 70.2 67.8 53.3 45.2 64.5 62 67.5 60.7 69.3 72.6 34 19.4 1
5.2
[20] 73 59.8 55.1 50.9 54.1 71.2 58.1 63.5 60.8 26.8 49.5 85.9 84.9 89.7 78.2 64.9 75.9 84.6 6
3.1
[39] 38.4 7.7 16.7 17.6 37.6 18.7 1.2 46.6 27.7
47 Levels: 1.2 7.7 15.2 16.7 17 17.6 18.7 19.4 26.8 27.7 34 35.3 36.5 37.6 38.4 39.7 43.5 ... 89
.7
> 11k=factor(c(swiss$Agriculture))
> print(11k)
[1] 17 45.1 39.7 36.5 43.5 35.3 70.2 67.8 53.3 45.2 64.5 62 67.5 60.7 69.3 72.6 34 19.4 1
5.2
[20] 73 59.8 55.1 50.9 54.1 71.2 58.1 63.5 60.8 26.8 49.5 85.9 84.9 89.7 78.2 64.9 75.9 84.6 6
3.1
[39] 38.4 7.7 16.7 17.6 37.6 18.7 1.2 46.6 27.7
47 Levels: 1.2 7.7 15.2 16.7 17 17.6 18.7 19.4 26.8 27.7 34 35.3 36.5 37.6 38.4 39.7 43.5 ... 89
.7
>
```

b. Give an introduction to DataFrame.

- i. Write a short note on accessing an elements from DF with Example**
- ii. Write some built in functions on dataframe with example.**

```
> id=c(1,2,3,4)
> name=c("Gress","Loki","Hitman","Groot")
> division=factor(c("A","B","A","A"))
> attendance=c("20%","50%","40%","80%")
>
> classdata=data.frame(id,name,division,attendance,stringsAsFactors = FALSE)
>
> classdata[3,2]
[1] "Hitman"
> classdata["name"]
  name
1 Gress
2  Loki
3 Hitman
4  Groot
>
> print(classdata$name)
[1] "Gress" "Loki" "Hitman" "Groot"
> print(classdata$attendance)
[1] "20%" "50%" "40%" "80%"
> |
```

```
> id=c(11,12,13,14)
> namea=c("Gress","Loki","Hitman","Groot")
> divisiona=factor(c("A","B","A","A"))
> attendancea=c("20%","50%","40%","80%")
>
> classdataa=data.frame(namea,divisiona,attendancea,stringsAsFactors = FALSE)
> classdataa
  namea divisiona attendancea
1 Gress          A         20%
2  Loki          B         50%
3 Hitman          A         40%
4  Groot          A         80%
> |
```

```
> str(airquality)
'data.frame': 153 obs. of 6 variables:
 $ Ozone : int 41 36 12 18 NA 28 23 19 8 NA ...
 $ Solar.R: int 190 118 149 313 NA NA 299 99 19 194 ...
 $ Wind : num 7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6
 $ Temp : int 67 72 74 62 56 66 65 59 61 69 ...
 $ Month : int 5 5 5 5 5 5 5 5 5 5 ...
 $ Day : int 1 2 3 4 5 6 7 8 9 10 ...
> head(airquality)
  Ozone Solar.R Wind Temp Month Day
1 41 190 7.4 67 5 1
2 36 118 8.0 72 5 2
3 12 149 12.6 74 5 3
4 18 313 11.5 62 5 4
5 NA NA 14.3 56 5 5
6 28 NA 14.9 66 5 6
> tail(airquality)
  Ozone Solar.R Wind Temp Month Day
148 14 20 16.6 63 9 25
149 30 193 6.9 70 9 26
150 NA 145 13.2 77 9 27
151 14 191 14.3 75 9 28
152 18 131 8.0 76 9 29
153 20 223 11.5 68 9 30
```

```

> nrow(airquality)
[1] 153
> ncol(airquality)
[1] 6
> summary(airquality)
      Ozone          Solar.R          Wind          Temp          Month          Day
Min.   : 1.00   Min.   : 7.0   Min.   : 1.700   Min.   :56.00   Min.   :5.000   Min.   : 1.0
1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400   1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
Median : 31.50   Median :205.0   Median : 9.700   Median :79.00   Median :7.000   Median :16.0
Mean   : 42.13   Mean   :185.9   Mean   : 9.958   Mean   :77.88   Mean   :6.993   Mean   :15.8
3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500   3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
Max.   :168.00   Max.   :334.0   Max.   :20.700   Max.   :97.00   Max.   :9.000   Max.   :31.0
NA's   :37      NA's   :7

> merge(classdata,classdataa)
   id  name division attendance namea divisiona attendancea
1  1  Gress         A         20%  Gress         A         20%
2  2  Loki          B         50%  Gress         A         20%
3  3  Hitman        A         40%  Gress         A         20%
4  4  Groot         A         80%  Gress         A         20%
5  1  Gress         A         20%  Loki          B         50%
6  2  Loki          B         50%  Loki          B         50%
7  3  Hitman        A         40%  Loki          B         50%
8  4  Groot         A         80%  Loki          B         50%
9  1  Gress         A         20%  Hitman        A         40%
10 2  Loki          B         50%  Hitman        A         40%
11 3  Hitman        A         40%  Hitman        A         40%
12 4  Groot         A         80%  Hitman        A         40%
13 1  Gress         A         20%  Groot         A         80%
14 2  Loki          B         50%  Groot         A         80%
15 3  Hitman        A         40%  Groot         A         80%
16 4  Groot         A         80%  Groot         A         80%

> is.data.frame(classdataa)
[1] TRUE
> pp=list(111,"Poco","A","10%")
> rbind(classdata,pp)
   id  name division attendance
1  1  Gress         A         20%
2  2  Loki          B         50%
3  3  Hitman        A         40%
4  4  Groot         A         80%
5 111  Poco          A         10%

> cbind(classdata,POE=c(44,44,44,44))
   id  name division attendance POE
1  1  Gress         A         20%  44
2  2  Loki          B         50%  44
3  3  Hitman        A         40%  44
4  4  Groot         A         80%  44
> |

```