Ex. 20 — Create a dataframe consisting of only the first two columns.

In [36]:
```
workshop<-c("R","SPSS",NA,"SPSS","STATA","SPSS")
gender <- factor (c(" Female ", " Male ", NA , " Female ", " Female ",
" Female ") )
q1 <- c(4 , 3 , 3 , 5 , 4 , 5)
q2 <- c(3 , 4 , 2 , 4 , 4 , 4)
q3 <- c(4 , 3 , NA , 5 , 3, 3)
q4 <- c(5 , 4 , 3 , 3 , 4 , 5)
df <- data.frame(workshop,gender,q1,q2,q3,q4)
df
```

workshop	gender	q1	q2	q3	q4
R	Female	4	3	4	5
SPSS	Male	3	4	3	4
NA	NA	3	2	NA	3
SPSS	Female	5	4	5	3
STATA	Female	4	4	3	4
SPSS	Female	5	4	3	5

Ex. 21 — Create a dataframe consisting of only the first and last row.

In [37]:
```
new_df<-df[,1:2]
new_df
```

workshop	gender
R	Female
SPSS	Male
NA	NA
SPSS	Female
STATA	Female
SPSS	Female

In [38]:
```
new_df<-df[c(1,nrow(df)),]
new_df
```

	workshop	gender	q1	q2	q3	q4
1	R	Female	4	3	4	5
6	SPSS	Female	5	4	3	5

Ex. 22 — What happens when you enter as.list(df)? unlist(df)? The as. list() function is used to convert the vector to a list in R,

In [39]:
```
as.list(df)
```

$workshop

$gender

1. R
2. SPSS
3. <NA>
4. SPSS
5. STATA
6. SPSS

▶ **Levels**:

$q1

1. Female
2. Male
3. <NA>
4. Female
5. Female
6. Female

▶ **Levels**:

$q2

1. 4
2. 3
3. 3
4. 5
5. 4
6. 5

$q3

1. 3
2. 4
3. 2
4. 4
5. 4
6. 4

$q4

1. 4
2. 3
3. <NA>
4. 5
5. 3
6. 3

1. 5
2. 4
3. 3
4. 3
5. 4
6. 5

Use unlist() function to convert a list to a vector by unlisting the elements from a list.

In [40]:
```
unlist(df)
```

workshop1	1
workshop2	2
workshop3	<NA>
workshop4	2
workshop5	3
workshop6	2
gender1	1
gender2	2
gender3	<NA>
gender4	1
gender5	1
gender6	1
q11	4
q12	3
q13	3
q14	5
q15	4
q16	5
q21	3
q22	4
q23	2
q24	4
q25	4
q26	4
q31	4
q32	3
q33	<NA>
q34	5
q35	3
q36	3
q41	5
q42	4
q43	3
q44	3
q45	4
q46	5

Ex. 23 — Create a dataframe called df2 where every entry in the q3 and q4 columns is 0.

In [41]:
```
df2<-df
df2[,c("q3","q4")]=0
df2
```

workshop	gender	q1	q2	q3	q4
R	Female	4	3	0	0
SPSS	Male	3	4	0	0
NA	NA	3	2	0	0
SPSS	Female	5	4	0	0
STATA	Female	4	4	0	0
SPSS	Female	5	4	0	0

Ex. 24 Sort df by gender

In [42]:
```
sort_df<-df[order(df$gender,decreasing=FALSE)]
sort_df
```

workshop	q2	q3	q4	gender	q1
R	3	4	5	Female	4
SPSS	4	3	4	Male	3
NA	2	NA	3	NA	3
SPSS	4	5	3	Female	5
STATA	4	3	4	Female	4
SPSS	4	3	5	Female	5

Ex. 25 Does df have any duplicate rows? What about df2 from exercise 23?

In [43]:
```
duplicate_rows<-df[duplicated(df),]
print(duplicate_rows)
```

```
[1] workshop gender   q1       q2       q3       q4
<0 rows> (or 0-length row.names)
```

Duplicate rows in df2

In [34]:
```
duplicate_rows<-df2[duplicated(df2),]
print(duplicate_rows)
```

```
  name age gender
6  Bob  30   Male
```

Ex. 26 To remove all rows having NA, we can use na. omit function.

In [44]:
```
na.omit(df)
```

	workshop	gender	q1	q2	q3	q4
1	R	Female	4	3	4	5
2	SPSS	Male	3	4	3	4
4	SPSS	Female	5	4	5	3
5	STATA	Female	4	4	3	4
6	SPSS	Female	5	4	3	5

In [ ]: