Data frame exercises Vol. 2

	Working
Alex	Yes
Lilly	No
Mark	No
Oliver	Yes
Martha	Yes
Lucas	No
Caroline	Yes

[In the exercises below we cover the basics of data frames.]

Answers to the exercises are available here.

For other parts of this series please follow the tag: dataframes.

Exercise 1

Consider two vectors:
x=seq(1,43,along.with=Id)
y=seq(-20,0,along.with=Id)
Create a data.frame df:

```
>df
Id Letter x y
1 1 a 1.000000 -20.000000
2 1 b 4.818182 -18.181818
3 1 c 8.636364 -16.363636
4 2 a 12.454545 -14.545455
5 2 b 16.272727 -12.727273
6 2 c 20.090909 -10.909091
7 3 a 23.909091 -9.090909
8 3 b 27.727273 -7.272727
9 3 c 31.545455 -5.454545
10 4 a 35.363636 -3.636364
11 4 b 39.181818 -1.818182
12 4 c 43.000000 0.000000
```

Exercise 2

From the previous one data frame df.
Create this data frame:

Id x.a y.a x.b y.b x.c y.c
1 1 1.00000 -20.000000 4.818182 -18.181818 8.636364 -16.363636
4 2 12.45455 -14.545455 16.272727 -12.727273 20.090909 -10.909091
7 3 23.90909 -9.090909 27.727273 -7.272727 31.545455 -5.454545
10 4 35.36364 -3.636364 39.181818 -1.818182 43.000000 0.000000

Exercise 3

Create two data frame df1 and df2:

> df1
Id Age
1 1 14
2 2 12
3 3 15
4 4 10
> df2
Id Sex Code
1 1 F a
2 2 M b
3 3 M c
4 4 F d

From df1 and df2 create M:

>M
Id Age Sex Code
1 1 14 F a
2 2 12 M b
3 3 15 M c
4 4 10 F d

Exercise 4

Create a data frame df3:

> df3
id2 score
1 4 100
2 3 98
3 2 94
4 1 99

From M and df3 create N:

Id Age Sex Code score
1 1 14 F a 99
2 2 12 M b 94
3 3 15 M c 98
4 4 10 F d 100

Exercise 5

Consider the previous one data frame N:
1)Remove the variables Sex and Code
2)From N, create a data frame:

values ind

1 1 Id
2 2 Id
3 3 Id
4 4 Id
5 14 Age
6 12 Age
7 15 Age
8 10 Age
9 99 score
10 94 score
11 98 score
12 100 score

Exercise 6

For this exercise, we'll use the (built-in) dataset trees.
a) Make sure the object is a data frame, if not change it to a data frame.
b) Create a new data frame A:

```
>A
Girth Height Volume
mean_tree 13.24839 76 30.17097
min_tree 8.30000 63 10.20000
max_tree 20.60000 87 77.00000
sum_tree 410.70000 2356 935.30000
```

Exercise 7

Consider the data frame A:
1)Order the entire data frame by the first column.
2)Rename the row names as follows: mean, min, max, tree

Exercise 8

Create an empty data frame with column types:

```
> df
Ints Logicals Doubles Characters
(or 0-length row.names)
```

Exercise 9

Create a data frame XY

```
X=c(1,2,3,1,4,5,2)
Y=c(0,3,2,0,5,9,3)
> XY
X Y
1 1 0
2 2 3
```

3 3 2
4 1 0
5 4 5
6 5 9
7 2 3

1)looks at duplicated elements using a provided R function.
2) keeps only the uniques lines on XY using a provided R function.

Exercise 10

For this exercise, we'll use the (built-in) dataset Titanic.
a) Make sure the object is a data frame, if not change it to a data frame.
b) Define a data frame with value 1st in Class variable, and value NO in Survived variable
and variables Sex, Age and Freq.

Sex Age Freq
1 Male Child 0
5 Female Child 0
9 Male Adult 118
13 Female Adult 4