SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

ITA 0443 - STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEM

DAY 2 – LAB EXERCISES

Reg No:192124052

Name:M.Nivetha

RESHAPE FUNCTION IN R

Exercise: 1

Construct the following data frame ‘country’.

a) Reshape in R from wide to long:

Reshape the above data frame from wide to long format in R.

- data frame “country” is passed to reshape function
- idvar is the variable which need to be left unaltered which is “countries”
- varying are the ones that needs to converted from wide to long
- v.names are the values that should be against the times in the resultant data frame.
- new.row.names is used to assign row names to the resultant dataset
- direction is, to which format the data needs to be transformed

Program:

```
country <- data.frame( countries = c("A", "B",

  "C"), value.population_in_million =
```

```
  c(100,200,120), value.gdp_perCapital =
  c(2000,7000,15000)
)
print(country) wide_df <- data.frame( countries
= c("A", "B", "C"), value.population_in_million =
c(100,200,120), value.gdp_perCapital =
c(2000,7000,15000)
)
```

Output:

```
countries                  variable  value
1         A  value.population_in_million   100
2         B  value.population_in_million   200
3         C  value.population_in_million   120
4         A         value.gdp_perCapital     2000
5         B         value.gdp_perCapital     7000
6         C         value.gdp_perCapital    15000
```

b) <u>Reshape in R from long to wide:</u>

- data (country_w_to_L) which is in long format, is passed to reshape function
- idvar is the variable which need to be left unaltered, which is “countries”
- timevar are the variables that needs to converted to wide format
- v.names are the value variable
- direction is, to which format the data needs to be transformed

7. MELTING AND CASTING IN R

Exercises :

1. Melt airquality data set and display as a long – format data ?

Program:

```
data("airquality")

airquality_long = melt(airquality, id.vars = c("Month", "Day"),

                       variable.name = "Measurement", value.name = "Value")

head(airquality_long)
```

Output:

```
> head(airquality_long)
  Month Day variable value
1     5   1    Ozone    41
2     5   2    Ozone    36
3     5   3    Ozone    12
4     5   4    Ozone    18
5     5   5    Ozone    NA
6     5   6    Ozone    28
```

2. Melt airquality data and specify month and day to be “ID variables” ?

Program:

```
library(tidyr)

# Load the airquality data

data(airquality)

# Melt the airquality data, keeping the "Month" and "Day" columns as the
identifier variables
```

```
airquality_melted <- melt(airquality, id.vars = c("Month", "Day"))

airquality_melted
```

Output:

	Month	Day	variable	value
1	5	1	Ozone	41
2	5	2	Ozone	36
3	5	3	Ozone	12
4	5	4	Ozone	18
5	5	5	Ozone	NA
6	5	6	Ozone	28
7	5	7	Ozone	23
8	5	8	Ozone	19
9	5	9	Ozone	8
10	5	10	Ozone	NA
11	5	11	Ozone	7
12	5	12	Ozone	16
13	5	13	Ozone	11
14	5	14	Ozone	14
15	5	15	Ozone	18

			Ozone
			Ozone
			Ozone
			Ozone
			Ozone
16	5	16	14
17	5	17	34
18	5	18	6
19	5	19	30
20	5	20	11
21	5	21	Ozone 1
22	5	22	Ozone 11
23	5	23	Ozone 4
24	5	24	Ozone 32
25	5	25	Ozone NA
26	5	26	Ozone NA
27	5	27	Ozone NA
28	5	28	Ozone 23
29	5	29	Ozone 45
30	5	30	Ozone 115
31	5	31	Ozone 37
32	6	1	Ozone NA
33	6	2	Ozone NA
34	6	3	Ozone NA
35	6	4	Ozone NA
36	6	5	Ozone NA
37	6	6	Ozone NA
38	6	7	Ozone 29
39	6	8	Ozone NA
40	6	9	Ozone 71
41	6	10	Ozone 39
42	6	11	Ozone NA
43	6	12	Ozone NA

			Ozone
			Ozone
			Ozone
			Ozone
			Ozone
44	6	13	23
45	6	14	NA
46	6	15	NA
47	6	16	21
48	6	17	37
49	6	18	Ozone 20
50	6	19	Ozone 12
51	6	20	Ozone 13
52	6	21	Ozone NA
53	6	22	Ozone NA
54	6	23	Ozone NA
55	6	24	Ozone NA
56	6	25	Ozone NA
57	6	26	Ozone NA
58	6	27	Ozone NA
59	6	28	Ozone NA
60	6	29	Ozone NA
61	6	30	Ozone NA
62	7	1	Ozone 135
63	7	2	Ozone 49
64	7	3	Ozone 32
65	7	4	Ozone NA
66	7	5	Ozone 64
67	7	6	Ozone 40
68	7	7	Ozone 77
69	7	8	Ozone 97
70	7	9	Ozone 97
71	7	10	Ozone 85

			Ozone
			Ozone
			Ozone
			Ozone
			Ozone
72	7	11	NA
73	7	12	10
74	7	13	27
75	7	14	NA
76	7	15	7
77	7	16	Ozone 48
78	7	17	Ozone 35
79	7	18	Ozone 61
80	7	19	Ozone 79
81	7	20	Ozone 63
82	7	21	Ozone 16
83	7	22	Ozone NA
84	7	23	Ozone NA
85	7	24	Ozone 80
86	7	25	Ozone 108
87	7	26	Ozone 20
88	7	27	Ozone 52
89	7	28	Ozone 82
90	7	29	Ozone 50
91	7	30	Ozone 64
92	7	31	Ozone 59
93	8	1	Ozone 39
94	8	2	Ozone 9
95	8	3	Ozone 16
96	8	4	Ozone 78
97	8	5	Ozone 35
98	8	6	Ozone 66
99	8	7	Ozone 122

			Ozone
100	8	8	89
101	8	9	Ozone 110
102	8	10	Ozone NA
103	8	11	Ozone NA
104	8	12	Ozone 44
105	8	13	Ozone 28
106	8	14	Ozone 65
107	8	15	Ozone NA
108	8	16	Ozone 22
109	8	17	Ozone 59
110	8	18	Ozone 23
111	8	19	Ozone 31
112	8	20	Ozone 44
113	8	21	Ozone 21
114	8	22	Ozone 9
115	8	23	Ozone NA
116	8	24	Ozone 45
117	8	25	Ozone 168
118	8	26	Ozone 73
119	8	27	Ozone NA
120	8	28	Ozone 76
121	8	29	Ozone 118
122	8	30	Ozone 84
123	8	31	Ozone 85
124	9	1	Ozone 96
125	9	2	Ozone 78
126	9	3	Ozone 73
127	9	4	Ozone 91
128	9	5	47
129	9	6	Ozone 32
130	9	7	Ozone 20

Ozone

131 9 8 Ozone 23 20

137 9 14 Ozone 21

138 9 15 Ozone 21

139 9 16 Ozone 46140 9
17 Ozone 21

132 9 9 Ozone 21

133 9 10 Ozone 24

134 9 11 Ozone 44

135 9 12 Ozone 21

136 9 13 Ozone 28

141 9 18 Ozone 13

142 9 19 Ozone 24

143 9 20 Ozone 16

144 9 21 Ozone 13

145 9 22 Ozone 23

146 9 23 Ozone 36

147 9 24 Ozone 7

148 9 25 Ozone 14

149 9 26 Ozone 30

150 9 27 Ozone NA

151 9 28 Ozone 14

152 9 29 Ozone 18

153 9 30 Ozone 20

154 5 1 Solar.R 190

155 5 2 Solar.R 118
156 5 3
Solar.R 149

157 5 4 Solar.R 313

158 5 5 Solar.R NA

159 5 6 Solar.R NA

160 5 7 Solar.R 299

161 5 8 Solar.R 99

162	5	9	Solar.R	19
163	5	10	Solar.R	194
164	5	11	Solar.R	NA
165	5	12	Solar.R	256
166	5	13	Solar.R	290
167	5	14	Solar.R	274
168	5	15	Solar.R	65
169	5	16	Solar.R	334
170	5	17	Solar.R	307
171	5	18	Solar.R	78
172	5	19	Solar.R	322
173	5	20	Solar.R	44
174	5	21	Solar.R	8
175	5	22	Solar.R	320
176	5	23	Solar.R	25
177	5	24	Solar.R	92
178	5	25	Solar.R	66
179	5	26	Solar.R	266
180	5	27	Solar.R	NA
181	5	28	Solar.R	13
182	5	29	Solar.R	252

Ozone

183 5 30 Solar.R
223

184	5	31	Solar.R	279
185	6	1	Solar.R	286
186	6	2	Solar.R	287
187	6	3	Solar.R	242
188	6	4	Solar.R	186
189	6	5	Solar.R	220
190	6	6	Solar.R	264
191	6	7	Solar.R	127
192	6	8	Solar.R	273
193	6	9	Solar.R	291
194	6	10	Solar.R	323
195	6	11	Solar.R	259
196	6	12	Solar.R	250
197	6	13	Solar.R	148
198	6	14	Solar.R	332
199	6	15	Solar.R	322
200	6	16	Solar.R	191
201	6	17	Solar.R	284
202	6	18	Solar.R	37
203	6	19	Solar.R	120
204	6	20	Solar.R	137
205	6	21	Solar.R	150
206	6	22	Solar.R	59
207	6	23	Solar.R	91
208	6	24	Solar.R	250
209	6	25	Solar.R	135

210 6 26 Solar.R 127

211 6 27 Solar.R 47 212 6 28 Solar.R 98

213 6 29 Solar.R 31

214 6 30 Solar.R 138

215 7 1 Solar.R 269

216 7 2 Solar.R 248

217 7 3 Solar.R 236

218 7 4 Solar.R 101

219 7 5 Solar.R 175

220 7 6 Solar.R 314

221 7 7 Solar.R 276

222 7 8 Solar.R 267

223 7 9 Solar.R 272

224 7 10 Solar.R 175

225 7 11 Solar.R 139

226 7 12 Solar.R 264

227 7 13 Solar.R 175

228 7 14 Solar.R 291

229 7 15 Solar.R 48

230 7 16 Solar.R 260

231 7 17 Solar.R 274

232 7 18 Solar.R 285

233 7 19 Solar.R 187

234 7 20 Solar.R 220

235 7 21 Solar.R 7

236 7 22 Solar.R 258

Ozone

237 7 23 Solar.R 295

238 7 24 Solar.R 294

239 7 25 Solar.R 223 240 7 26 Solar.R 81

241 7 27 Solar.R 82

242 7 28 Solar.R 213

243 7 29 Solar.R 275

244 7 30 Solar.R 253

245 7 31 Solar.R 254

246 8 1 Solar.R 83

247 8 2 Solar.R 24

248 8 3 Solar.R 77

249 8 4 Solar.R NA

250 8 5 Solar.R NA

3. Cast the molten airquality data set

Program:

```
library(reshape2)

data("airquality")

# Cast the airquality dataset to a long format

airquality_long <- melt(airquality, id.vars = c("Month",
"Day"))

airquality_long
```

Output:

```
> airquality_long
      Month Day variable value
1         5   1    Ozone 41
2         5   2    Ozone 36
3         5   3    Ozone 12
```

4	5	4	Ozone 18
5	5	5	Ozone NA
6	5	6	28

Ozone
Ozone
Ozone
Ozone
Ozone
Ozone
Ozone

7	5	7	23
8	5	8	19
9	5	9	8
10	5	10	NA
11	5	11	7
12	5	12	16
13	5	13	11
14	5	14	Ozone 14
15	5	15	Ozone 18
16	5	16	Ozone 14
17	5	17	Ozone 34
18	5	18	Ozone 6
19	5	19	Ozone 30
20	5	20	Ozone 11
21	5	21	Ozone 1
22	5	22	Ozone 11
23	5	23	Ozone 4
24	5	24	Ozone 32
25	5	25	Ozone NA
26	5	26	Ozone NA
27	5	27	Ozone NA
28	5	28	Ozone 23
29	5	29	Ozone 45
30	5	30	Ozone 115
31	5	31	Ozone 37
32	6	1	Ozone NA

Ozone NA

			Ozone	
			Ozone	
			Ozone	
			Ozone	
			Ozone	
			Ozone	
			Ozone	
33	6	2	Ozone	NA
34	6	3	Ozone	NA
35	6	4	Ozone	NA
36	6	5	Ozone	NA
37	6	6	Ozone	NA
38	6	7	Ozone	29
39	6	8	Ozone	NA
40	6	9	Ozone	71
41	6	10	Ozone	39
42	6	11	Ozone	NA
43	6	12	Ozone	NA
44	6	13	Ozone	23
45	6	14	Ozone	NA
46	6	15	Ozone	NA
47	6	16	Ozone	21
48	6	17	Ozone	37
49	6	18	Ozone	20
50	6	19	Ozone	12
51	6	20	Ozone	13
52	6	21	Ozone	NA
53	6	22	Ozone	NA
54	6	23		
55	6	24	NA	
56	6	25	NA	
57	6	26	NA	
58	6	27	NA	
			Ozone	NA

Ozone
Ozone
Ozone
Ozone
Ozone
Ozone
Ozone
59 6 28 NA
60 6 29 NA
61 6 30 NA
62 7 1 Ozone 135
63 7 2 Ozone 49
64 7 3 Ozone 32
65 7 4 Ozone NA
66 7 5 Ozone 64
67 7 6 Ozone 40
68 7 7 Ozone 77
69 7 8 Ozone 97
70 7 9 Ozone 97
71 7 10 Ozone 85
72 7 11 Ozone NA
73 7 12 Ozone 10
74 7 13 Ozone 27
75 7 14 Ozone NA
76 7 15 Ozone 7
77 7 16 Ozone 48
78 7 17 Ozone 35
79 7 18 Ozone 61
80 7 19 Ozone 79
81 7 20 Ozone 63
82 7 21 Ozone 16
83 7 22 Ozone NA
84 7 23 Ozone NA
85 7 24 Ozone 80
Ozone NA

			Ozone	
			Ozone	
			Ozone	
			Ozone	
			Ozone	
			Ozone	
			Ozone	
86	7	25	Ozone	108
87	7	26	Ozone	20
88	7	27	Ozone	52
89	7	28	Ozone	82
90	7	29	Ozone	50
91	7	30	Ozone	64
92	7	31	Ozone	59
93	8	1	Ozone	39
94	8	2	Ozone	9
95	8	3	Ozone	16
96	8	4	Ozone	78
97	8	5	Ozone	35
98	8	6	Ozone	66
99	8	7	Ozone	122
100	8	8	Ozone	89
101	8	9	Ozone	110
102	8	10		
103	8	11	NA	
104	8	12	44	
105	8	13	28	
106	8	14	65	
107	8	15	NA	
108	8	16	22	
109	8	17	59	
110	8	18	Ozone	23
111	8	19	Ozone	31
			Ozone	NA

Ozone
Ozone
Ozone
Ozone
Ozone
Ozone
Ozone
112 8 20 Ozone 44
113 8 21 Ozone 21
114 8 22 Ozone 9
115 8 23 Ozone NA
116 8 24 Ozone 45
117 8 25 Ozone 168
118 8 26 Ozone 73
119 8 27 Ozone NA
120 8 28 Ozone 76
121 8 29 Ozone 118
122 8 30 Ozone 84
123 8 31 Ozone 85
124 9 1 Ozone 96
125 9 2 Ozone 78
126 9 3 Ozone 73
127 9 4 Ozone 91
128 9 5 Ozone 47
129 9 6 Ozone 32
130 9 7 Ozone 20
131 9 8 Ozone 23
132 9 9 Ozone 21
133 9 10 Ozone 24
134 9 11 Ozone 44
135 9 12 Ozone 21
136 9 13 Ozone 28

Ozone NA

Ozone
Ozone
Ozone
Ozone
Ozone
Ozone
Ozone

137	9	14	Ozone 9
138	9	15	Ozone 13
139	9	16	Ozone 46
140	9	17	Ozone 18
141	9	18	Ozone 13
142	9	19	Ozone 24
143	9	20	Ozone 16
144	9	21	Ozone 13
145	9	22	Ozone 23
146	9	23	Ozone 36
147	9	24	Ozone 7
148	9	25	Ozone 14
149	9	26	Ozone 30
150	9	27	

Ozone NA

Ozone
Ozone
Ozone
151 9 28 14
152 9 29 18
153 9 30 20
154 5 1 Solar.R 190 155 5 2 Solar.R 118
156 5 3 Solar.R 149
157 5 4 Solar.R 313
158 5 5 Solar.R NA
159 5 6 Solar.R NA
160 5 7 Solar.R 299
161 5 8 Solar.R 99
162 5 9 Solar.R 19
163 5 10 Solar.R 194
164 5 11 Solar.R NA
165 5 12 Solar.R 256
166 5 13 Solar.R 290
167 5 14 Solar.R 274
168 5 15 Solar.R 65
169 5 16 Solar.R 334
170 5 17 Solar.R 307
171 5 18 Solar.R 78
172 5 19 Solar.R 322
173 5 20 Solar.R 44
174 5 21 Solar.R 8
175 5 22 Solar.R 320
176 5 23 Solar.R 25
177 5 24 Solar.R 92
178 5 25 Solar.R 66
179 5 26 Solar.R 266
180 5 27 Solar.R NA
181 5 28 Solar.R 13
182 5 29 Solar.R 252
183 5 30 Solar.R 223
184 5 31 Solar.R 279
185 6 1 Solar.R 286
186 6 2 Solar.R 287
187 6 3 Solar.R 242
188 6 4 Solar.R 186
189 6 5 Solar.R 220
190 6 6 Solar.R 264
191 6 7 Solar.R 127
192 6 8 Solar.R 273
193 6 9 Solar.R 291

194 6 10 Solar.R 323
195 6 11 Solar.R 259
196 6 12 Solar.R 250
197 6 13 Solar.R 148
198 6 14 Solar.R 332
199 6 15 Solar.R 322
200 6 16 Solar.R 191
201 6 17 Solar.R 284
202 6 18 Solar.R 37
203 6 19 Solar.R 120
204 6 20 Solar.R 137
205 6 21 Solar.R 150
206 6 22 Solar.R 59
207 6 23 Solar.R 91
208 6 24 Solar.R 250
209 6 25 Solar.R 135
210 6 26 Solar.R 127
211 6 27 Solar.R 47
212 6 28 Solar.R 98
213 6 29 Solar.R 31
214 6 30 Solar.R 138
215 7 1 Solar.R 269
216 7 2 Solar.R 248
217 7 3 Solar.R 236
218 7 4 Solar.R 101
219 7 5 Solar.R 175
220 7 6 Solar.R 314
221 7 7 Solar.R 276
222 7 8 Solar.R 267
223 7 9 Solar.R 272
224 7 10 Solar.R 175
225 7 11 Solar.R 139
226 7 12 Solar.R 264
227 7 13 Solar.R 175
228 7 14 Solar.R 291
229 7 15 Solar.R 48
230 7 16 Solar.R 260
231 7 17 Solar.R 274
232 7 18 Solar.R 285
233 7 19 Solar.R 187
234 7 20 Solar.R 220
235 7 21 Solar.R 7

```
236     7  22  Solar.R   258
237     7  23  Solar.R   295
238     7  24  Solar.R   294
239     7  25  Solar.R   223
240     7  26  Solar.R 81
241     7  27  Solar.R 82
242     7  28  Solar.R   213
243     7  29  Solar.R   275
244     7  30  Solar.R   253
245     7  31  Solar.R   254
246     8   1  Solar.R 83 247
                8   2  Solar.R
        24
248     8   3  Solar.R 77
249     8   4  Solar.R NA
250     8   5  Solar.R NA
```

4. Use cast function appropriately and compute the average of Ozone, Solar.R , Wind and temperature per month ?

Program:

```
df=airquality

md=melt.data.frame(df)

print(cast(md))
```

8 FILE MANUPULATION IN R

Exercise

1. Consider the following data present. Create this file using windows notepad . Save the file as **input.csv** using the save As All files(*.*) option in notepad.

2. Use appropriate R commands to read **input.csv** file.

Program:

```
file="C:\Users\Uthra\OneDrive\Desktop\R Programming\input.csv"
```

```
data = read.csv(file)

print(data)
```

3. **Analyze the CSV File and compute the following.**

a. Get the maximum salary

Program:

```
file="C:\Users\Uthra\OneDrive\Desktop\R Programming\input.csv"

data = read.csv(file)

print(max(data$salary))
```

b. Get the details of the person with max salary

c. Get all the people working in IT department

```
file="C:\Users\Uthra\OneDrive\Desktop\R Programming\input.csv"

data = read.csv(file) print(data$dept=IT)
```

d. Get the persons in IT department whose salary is greater than 600

e. Get the people who joined on or after 2014

4. Get the people who joined on or after 2014 and write the output onto a file called output.csv