Omkar Gurav
TE IT 8048

Assignment - 5

- Aim:- Perform the following operations using R on the facebook metrics Dataset.
  1. Create data subsets
  2. Merge Data
  3. Sort Data
  4. Transposing Data
  5. Melting Data to long format
  6. Casting Data to wide format

- Theory:-

Q.1. What is R? Explain the features of R.

→ R is a language & environment for statistical computing and graphics. It is a GNU project which is similar to the S language & environment which was developed at Bell Laboratories.

R provides a wide variety of statistical & graphical techniques & is highly extensible.

One of R's strengths is the ease with which well designed publication-quality plots can be produced, including mathematical symbols & formulae where needed.

Features of R:-

① R includes conditionals, loops, user defined recursive functions & input & output facilities.

② R has an effective data handling & storage facility.

③ R provides large, coherent and integrated collection of tools for data analysis.

④ R provides graphical facilities for data analysis & display either directly at the computer or printing on paper.

⑤ R contains suite of operators for different types of calculations on arrays, lists and vectors.

Q.2). What is use of melting & casting?

→ R programming language has many methods to reshape the data using reshape package. melt() and cast() are the functions that efficiently reshape the data.

Melting is done to organize the data. It is performed using melt() function which takes dataset & column values that has to be kept constant. Using melt(), dataframe

is converted into long format and stretches the dataframe.

Casting is used to reshape the molten data using cast() function which takes aggregate function and formula to aggregate the data accordingly. This function is used to convert long format data back into some aggregated form of data based on the formula in cast().

- Conclusion:- Thus we have successfully performed operations using R on the Facebook metrics Data set.

Facebook.R

```
zipf <- "C:/Users/DELL/Downloads/Facebook_metrics.zip"

OutDir <- "C:/Users/DELL/Downloads/Facebook_metrics"


 unzip(zipf, exdir = OutDir)

 Facebook_Data <- read.csv("C:/Users/DELL/Downloads/Facebook_metrics/dataset_Facebook.csv", sep = ";")

 # Subsetting dataframe
 Subset_data <- Facebook_Data[1:7, 1:8]
 head(Subset_data)

 my_data <- Facebook_Data[, 1:4]
 # Merging dataframes using common column
 Merged_data <- merge(Subset_data, my_data, by = "Type")
 View(Merged_data)

 # rbind
 rbind_data_1 <- Facebook_Data[1:40,]
 rbind_data_2 <- Facebook_Data[51:80,]

 rbind_data <- rbind(rbind_data_1, rbind_data_2)
 View(rbind_data)

 # cbind
 cbind_data_1 <- Facebook_Data[1:40, 1:4]
 cbind_data_2 <- Facebook_Data[1:40, 5:8]

 cbind_data <- cbind(cbind_data_1, cbind_data_2)
 View(cbind_data)

 # Sorting dataframe on basis of likes using order()
 Likes_sorted_descending <- Facebook_Data[order(Facebook_Data$like, decreasing = "True"), c(1,2,3,4,17) ]
 head(Likes_sorted_descending)

 # Transposing dataframe
 transpose_data <- t(Subset_data)
```

```
View(transpose_data)


library("reshape")
# Wide format to long format
molten_data <- melt(Subset_data, id = c("Type","Category"))
View(molten_data)

# Long format to wide format
cast_data <- cast(molten_data, fun.aggregate = max)
View(cast_data)
```

Output :

```
> zipf <- "C:/Users/DELL/Downloads/Facebook_metrics.zip"
> OutDir <- "C:/Users/DELL/Downloads/Facebook_metrics"
>  unzip(zipf, exdir = OutDir)
>  Facebook_Data <- read.csv("C:/Users/DELL/Downloads/Facebook_metrics/dataset_Facebook.csv", sep = ";")
>  # Subsetting dataframe
>  Subset_data <- Facebook_Data[1:7, 1:8]
>  head(Subset_data)
  Page.total.likes   Type Category Post.Month Post.Weekday Post.Hour Paid Lifetime.Post.Total.Reach
1           139441  Photo        2         12           4         3    0                       2752
2           139441 Status        2         12           3        10    0                      10460
3           139441  Photo        3         12           3         3    0                       2413
4           139441  Photo        2         12           2        10    1                      50128
5           139441  Photo        2         12           2         3    0                       7244
6           139441 Status        2         12           1         9    0                      10472
>  my_data <- Facebook_Data[, 1:4]
>  # Merging dataframes using common column
>  Merged_data <- merge(Subset_data, my_data, by = "Type")
>  View(Merged_data)
```

RStudio Source Editor

Merged_data

Filter

	Type	Page.total.likes.x	Category.x	Post.Month.x	Post.Weekday	Post.Hour	Paid	Lifetime.Post.Total.Reach	Page.total.likes.y	Category.y	Post.Month.y
1	Photo	139441	2	12	4	3	0	2752	139441	2	12
2	Photo	139441	2	12	4	3	0	2752	135713	3	10
3	Photo	139441	2	12	4	3	0	2752	138414	2	12
4	Photo	139441	2	12	4	3	0	2752	139441	2	12
5	Photo	139441	2	12	4	3	0	2752	138414	1	12
6	Photo	139441	2	12	4	3	0	2752	138895	2	12
7	Photo	139441	2	12	4	3	0	2752	136013	1	10
8	Photo	139441	2	12	4	3	0	2752	139441	2	12
9	Photo	139441	2	12	4	3	0	2752	139441	2	12
10	Photo	139441	2	12	4	3	0	2752	138414	1	12
11	Photo	139441	2	12	4	3	0	2752	138895	1	12
12	Photo	139441	2	12	4	3	0	2752	138895	3	12
13	Photo	139441	2	12	4	3	0	2752	138414	2	12
14	Photo	139441	2	12	4	3	0	2752	136013	3	10
15	Photo	139441	2	12	4	3	0	2752	139441	2	12
16	Photo	139441	2	12	4	3	0	2752	139441	2	12
17	Photo	139441	2	12	4	3	0	2752	136013	1	10
18	Photo	139441	2	12	4	3	0	2752	138414	3	12
19	Photo	139441	2	12	4	3	0	2752	139441	3	12
20	Photo	139441	2	12	4	3	0	2752	138353	1	12
21	Photo	139441	2	12	4	3	0	2752	136393	3	10
22	Photo	139441	2	12	4	3	0	2752	138895	3	12
23	Photo	139441	2	12	4	3	0	2752	138329	1	11
24	Photo	139441	2	12	4	3	0	2752	138329	1	11
25	Photo	139441	2	12	4	3	0	2752	138329	1	11
26	Photo	139441	2	12	4	3	0	2752	138895	1	12
27	Photo	139441	2	12	4	3	0	2752	139441	3	12
28	Photo	139441	2	12	4	3	0	2752	138329	1	11
29	Photo	139441	2	12	4	3	0	2752	138329	1	11
30	Photo	139441	2	12	4	3	0	2752	138329	1	11
31	Photo	139441	2	12	4	3	0	2752	138329	1	11
32	Photo	139441	2	12	4	3	0	2752	138185	1	11
33	Photo	139441	2	12	4	3	0	2752	138353	1	12
34	Photo	139441	2	12	4	3	0	2752	136393	1	10
35	Photo	139441	2	12	4	3	0	2752	138353	2	11
36	Photo	139441	2	12	4	3	0	2752	138185	1	11
37	Photo	139441	2	12	4	3	0	2752	138185	1	11
38	Photo	139441	2	12	4	3	0	2752	138185	1	11

Showing 1 to 40 of 2,220 entries, 11 total columns

```
> # rbind
> rbind_data_1 <- Facebook_Data[1:40,]
> rbind_data_2 <- Facebook_Data[51:80,]
> rbind_data <- rbind(rbind_data_1, rbind_data_2)
> View(rbind_data)
```

RStudio Source Editor

rbind_data

Filter

	Page.total.likes	Type	Category	Post.Month	Post.Weekday	Post.Hour	Paid	Lifetime.Post.Total.Reach	Lifetime.Post.Total.Impressions	Lifetime.Engaged.Users	Lifetime.Post.Consumers	Lifetime.Post.Consumptions	Lifetime.Post.Impressions.by.people.who.have.liked.your.Page	Lifetime.Post.reach.by.people.who.like.your.Pa
1	139441	Photo	2	12	4	3	0	2752	5091	178	109	159	3078	1640
2	139441	Status	2	12	3	10	0	10460	19057	1457	1361	1674	11710	6112
3	139441	Photo	3	12	3	3	0	2413	4373	177	113	154	2812	1503
4	139441	Photo	2	12	2	10	1	50128	87991	2211	790	1119	61027	32048
5	139441	Photo	2	12	2	3	0	7244	13594	671	410	580	6228	3200
6	139441	Status	2	12	1	9	0	10472	20849	1191	1073	1389	16034	7852
7	139441	Photo	3	12	1	3	1	11692	19479	481	265	364	15432	9328
8	139441	Photo	3	12	7	9	1	13720	24137	537	232	305	19728	11056
9	139441	Status	2	12	7	3	0	11844	22538	1530	1407	1692	15220	7912
10	139441	Photo	3	12	6	10	0	4694	8668	280	183	250	4309	2324
11	139441	Status	2	12	5	10	0	21744	42334	4258	4100	4540	37849	18952
12	139441	Photo	2	12	5	10	0	3112	5590	208	127	145	3887	2174
13	139441	Photo	2	12	5	10	0	2847	5133	193	115	133	3779	2072
14	139441	Photo	2	12	5	3	0	2549	4896	249	134	168	3631	1917
15	138414	Photo	2	12	4	5	1	22784	39941	887	337	417	34415	19312
16	138414	Status	2	12	3	10	0	10060	19680	1264	1209	1425	17272	8548
17	138414	Photo	3	12	3	3	0	1722	2981	163	123	148	1868	1050
18	138414	Photo	1	12	2	12	1	53264	111785	1706	1103	1655	92512	39776
19	138414	Status	3	12	2	3	0	3930	7509	130	86	112	5009	2410
20	138414	Photo	3	12	1	11	0	1591	2825	121	88	111	2116	1161
21	138414	Photo	2	12	1	3	0	2848	5066	200	142	184	3561	1963
22	138414	Photo	1	12	7	10	0	1384	2467	15	15	20	2196	1172
23	138414	Link	1	12	7	10	0	3454	6853	118	104	130	6282	3100
24	138414	Photo	3	12	7	3	0	2723	4888	176	118	143	2964	1621
25	138414	Status	2	12	6	10	0	8488	15294	1341	1270	1489	9684	5244
26	138458	Status	2	12	6	3	0	8284	15104	1521	1462	1711	10266	5372
27	138458	Status	2	12	5	11	0	19552	34143	2806	2531	3420	17748	9824
28	138458	Photo	3	12	5	3	0	2478	4306	212	124	149	2612	1443
29	138895	Photo	2	12	5	3	0	9560	18264	973	559	885	9217	4748
30	138895	Video	1	12	4	11	1	36208	61262	1141	1068	1728	30131	14112
31	138895	Photo	2	12	4	2	0	4940	9390	385	306	501	5860	2930
32	138895	Photo	2	12	3	10	0	1683	2929	192	171	221	1585	858
33	138895	Photo	3	12	3	3	0	5280	9578	368	237	345	4480	2422
34	138895	Photo	3	12	2	9	0	3002	5318	268	185	247	3039	1676
35	138895	Photo	1	12	2	3	0	3766	7149	298	260	431	5782	2938
36	138895	Photo	2	12	1	11	0	4512	7808	423	284	431	5183	2954
37	138895	Photo	3	12	1	3	0	2690	4628	252	168	226	3052	1727

Showing 1 to 40 of 70 entries, 19 total columns

```
> # cbind
> cbind_data_1 <- Facebook_Data[1:40, 1:4]
> cbind_data_2 <- Facebook_Data[1:40, 5:8]
> cbind_data <- cbind(cbind_data_1, cbind_data_2)
> View(cbind_data)
```

RStudio Source Editor

cbind_data

Filter

	Page.total.likes	Type	Category	Post.Month	Post.Weekday	Post.Hour	Paid	Lifetime.Post.Total.Reach
1	139441	Photo	2	12	4	3	0	2752
2	139441	Status	2	12	3	10	0	10460
3	139441	Photo	3	12	3	3	0	2413
4	139441	Photo	2	12	2	10	1	50128
5	139441	Photo	2	12	2	3	0	7244
6	139441	Status	2	12	1	9	0	10472
7	139441	Photo	3	12	1	3	1	11692
8	139441	Photo	3	12	7	9	1	13720
9	139441	Status	2	12	7	3	0	11844
10	139441	Photo	3	12	6	10	0	4694
11	139441	Status	2	12	5	10	0	21744
12	139441	Photo	2	12	5	10	0	3112
13	139441	Photo	2	12	5	10	0	2847
14	139441	Photo	2	12	5	3	0	2549
15	138414	Photo	2	12	4	5	1	22784
16	138414	Status	2	12	3	10	0	10060
17	138414	Photo	3	12	3	3	0	1722
18	138414	Photo	1	12	2	12	1	53264
19	138414	Status	3	12	2	3	0	3930
20	138414	Photo	3	12	1	11	0	1591
21	138414	Photo	2	12	1	3	0	2848
22	138414	Photo	1	12	7	10	0	1384
23	138414	Link	1	12	7	10	0	3454
24	138414	Photo	3	12	7	3	0	2723
25	138414	Status	2	12	6	10	0	8488
26	138458	Status	2	12	6	3	0	8284
27	138458	Status	2	12	5	11	0	19552
28	138458	Photo	3	12	5	3	0	2478
29	138895	Photo	2	12	5	3	0	9560
30	138895	Video	1	12	4	11	1	36208
31	138895	Photo	2	12	4	2	0	4940
32	138895	Photo	2	12	3	10	0	1683
33	138895	Photo	3	12	3	3	0	5280
34	138895	Photo	3	12	2	9	0	3002
35	138895	Photo	1	12	2	3	0	3766
36	138895	Photo	2	12	1	11	0	4512
37	138895	Photo	3	12	1	3	0	2690
38	138895	Photo	1	12	7	10	1	19800

Showing 1 to 40 of 40 entries, 8 total columns

```
> # Sorting dataframe on basis of likes using order()
> Likes_sorted_descending <- Facebook_Data[order(Facebook_Data$like, decreasing = "True"), c(1,2,3,4,17) ]
> head(Likes_sorted_descending)
    Page.total.likes  Type Category Post.Month like
245           130791 Photo        2          7 5172
380           111620 Photo        3          4 1998
350           117764 Photo        3          5 1639
169           135428 Photo        1          9 1622
4             139441 Photo        2         12 1572
461            92507 Photo        3          2 1546
> # Transposing dataframe
> transpose_data <- t(Subset_data)
> View(transpose_data)
```

RStudio Source Editor

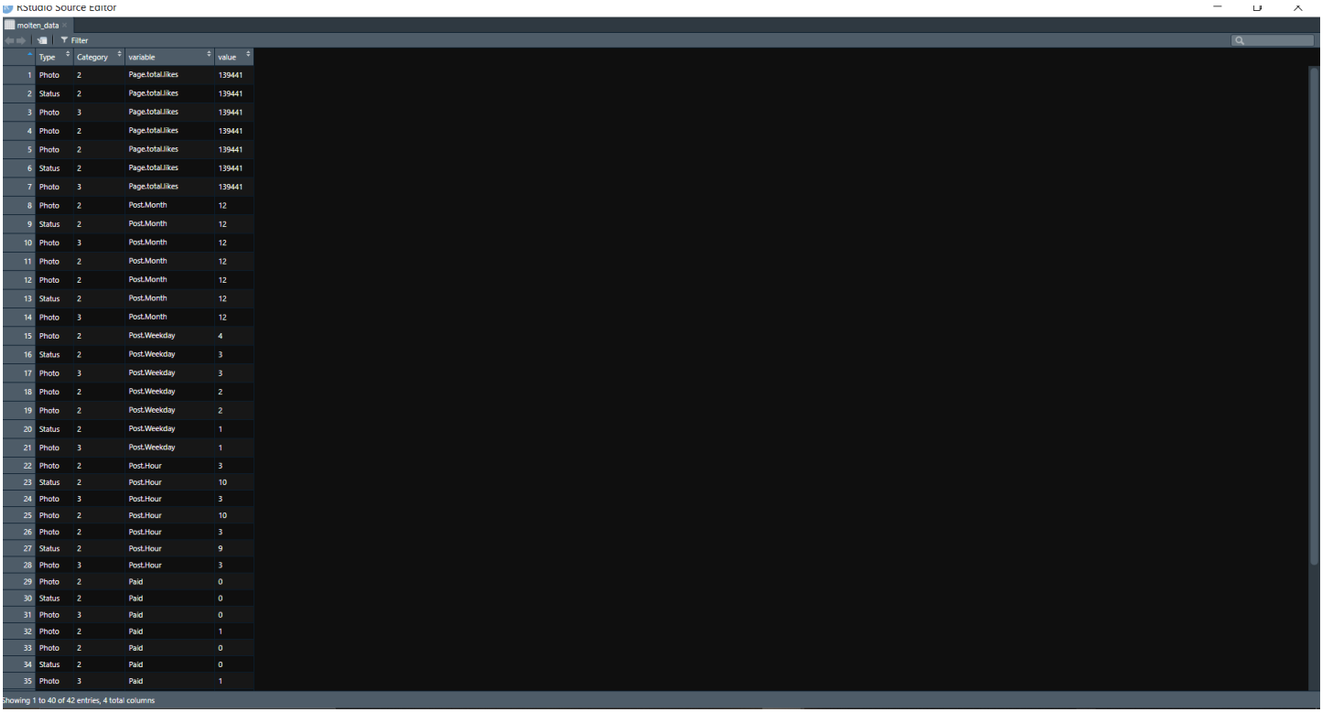
transpose_data

Filter

	1	2	3	4	5	6	7
Page.total.likes	139441	139441	139441	139441	139441	139441	139441
Type	Photo	Status	Photo	Photo	Photo	Status	Photo
Category	2	2	3	2	2	2	3
Post.Month	12	12	12	12	12	12	12
Post.Weekday	4	3	3	2	2	1	1
Post.Hour	3	10	3	10	3	9	3
Paid	0	0	0	1	0	0	1
Lifetime.Post.Total.Reach	2752	10460	2413	50128	7244	10472	11692

Showing 1 to 8 of 8 entries, 7 total columns

```
> library("reshape")
> # Wide format to long format
> molten_data <- melt(Subset_data, id = c("Type","Category"))
> View(molten_data)
```




```
> # Long format to wide format
> cast_data <- cast(molten_data, fun.aggregate = max)
> View(cast_data)
```