

Sheth L.U.J & Sir M.V College  
SAS/SSPS/R  
Practical No.5

Aim: Sorting data using `arrange()` in R.

Output:

The screenshot displays the RStudio interface with a script editor on the left and the Environment pane on the right. The script contains the following R code:

```

1 library(dplyr)
2
3 brain <- read.csv("brain_tumor.csv")
4
5 brain_sorted_mean <- brain %>
6   arrange(Neon)
7   head(brain_sorted_mean, 5)
8
9 brain_sorted_contrast_desc <- brain %>
10  arrange(desc(Contrast))
11  head(brain_sorted_contrast_desc, 5)
12
13 brain_multi_sort <- brain %>
14  arrange(Class, desc(Entropy))
15  head(brain_multi_sort, 10)
16
17 # R4.3.2 ->
18
19 library(dplyr)
20 brain <- read.csv("brain_tumor.csv")
21 brain_sorted_mean <- brain %>
22   arrange(Neon)
23   head(brain_sorted_mean, 5)
24
25 # Image Class Mean Variance Standard Deviation Entropy Skewness Kurtosis Contrast Energy
26 1 Image790 0.0, 0.0763966 3.145628 1.771592 0.2237384 23.79804 385.7968 213.6557 0.4303523
27 2 Image643 0.0, 0.0921154 3.142207 1.828170 0.2475573 22.06995 318.4727 182.8011 0.4358317
28 3 Image693 0.0, 0.0021254 3.142207 1.828170 0.2475573 22.06995 318.4727 182.8011 0.4358317
29 4 Image3184 1.0, 0.1063715 14.50193 3.820293 0.01025748 36.93129 1371.6401 100.5714 0.1465718
30 5 Image2462 1.0, 0.1339841 13.456761 3.722467 0.0526871 33.00861 1099.5635 841.3194 0.2024816
31
32 # Homogeneity Dissimilarity Correlation Coarseness
33 1 0.1832058 0.8703303 0.7339343 7.438341e-135
34 2 0.10760026 0.5903530 0.7543144 7.438341e-155
35 3 0.10760026 0.5903530 0.7543144 7.438341e-155
36 4 0.02343193 0.2270128 7.428571 0.5494262 7.438341e-135
37 5 0.94009112 0.4307953 10.168444 0.8138116 7.438341e-135
38
39 #

```

The Environment pane on the right shows the following objects:

- brain**: data.frame [1000 x 10]
- brain\_sorted\_mean**: data.frame [5 x 10]
- brain\_sorted\_contrast\_desc**: data.frame [5 x 10]
- brain\_multi\_sort**: data.frame [10 x 10]
- Image790**: data.frame [1 x 10]
- Image643**: data.frame [1 x 10]
- Image693**: data.frame [1 x 10]
- Image3184**: data.frame [1 x 10]
- Image2462**: data.frame [1 x 10]

The screenshot shows the RStudio interface with a script editor on the left, a console on the bottom left, and a file explorer on the right.

**Script Editor:**

```

1 library(dplyr)
2
3 brain <- read_csv("brain_tumor.csv")
4
5 brain_sorted_mean <- brain %>%
6   arrange(desc(mean))
7   head(brain_sorted_mean, 5)
8
9 brain_sorted_contrast_desc <- brain %>%
10  arrange(desc(contrast))
11  head(brain_sorted_contrast_desc, 5)
12
13 brain_multi_surt <- brain %>%
14  arrange(Class, desc(entropy))
15  head(brain_multi_surt, 10)
16
17 # R4.52
18
19 image <- 1:10
20
21 image <- 1:10
22
23 image <- 1:10
24
25 image <- 1:10
26
27 image <- 1:10
28
29 image <- 1:10
30
31 image <- 1:10
32
33 image <- 1:10
34
35 image <- 1:10
36
37 image <- 1:10
38
39 image <- 1:10
40
41 image <- 1:10
42
43 image <- 1:10
44
45 image <- 1:10
46
47 image <- 1:10
48
49 image <- 1:10
50
51 image <- 1:10
52
53 image <- 1:10
54
55 image <- 1:10
56
57 image <- 1:10
58
59 image <- 1:10
60
61 image <- 1:10
62
63 image <- 1:10
64
65 image <- 1:10
66
67 image <- 1:10
68
69 image <- 1:10
70
71 image <- 1:10
72
73 image <- 1:10
74
75 image <- 1:10
76
77 image <- 1:10
78
79 image <- 1:10
80
81 image <- 1:10
82
83 image <- 1:10
84
85 image <- 1:10
86
87 image <- 1:10
88
89 image <- 1:10
90
91 image <- 1:10
92
93 image <- 1:10
94
95 image <- 1:10
96
97 image <- 1:10
98
99 image <- 1:10
100

```

**Console Output:**

```

# A tibble: 10 x 10
  image_id mean contrast entropy skewness kurtosis contrast energy
  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 1 0.11308843 11.456761 3.722467 0.01526872 33.00861 1089.5635 941.3194 0.3024638
2 2 0.18520308 0.6709300 3.453480 0.7339345 7.458341e-155
3 3 0.20760026 0.3800590 5.487436 0.7543444 7.458341e-155
4 4 0.18005326 0.38005326 5.487436 0.7543444 7.458341e-155
5 5 0.02154195 0.2279128 7.478571 0.3494262 7.458341e-155
6 6 0.04099151 0.4307953 10.369444 0.8158116 7.458341e-155
7 7 0.04099151 0.4307953 10.369444 0.8158116 7.458341e-155
8 8 0.04099151 0.4307953 10.369444 0.8158116 7.458341e-155
9 9 0.04099151 0.4307953 10.369444 0.8158116 7.458341e-155
10 10 0.04099151 0.4307953 10.369444 0.8158116 7.458341e-155

```

**Table 1: Image Statistics**

Image Class	Mean	Variance	Standard Deviation	Entropy	Skewness	Kurtosis	Contrast	Energy
Image 258	1.04679108	87.76675	9.368383	0.044348175	21.03274	457.2446	3382.574	0.10178670
Image 174	1.17350421	313.41010	37.704522	0.11263470	10.97700	125.1151	1591.167	0.29889923
Image 254	1.04679108	179.23807	33.307609	0.018060599	14.91010	229.4466	1546.996	0.11445647
Image 234	1.02767028	26.87312	3.184122	0.092798252	21.43227	504.9748	1218.839	0.0731804
Image 307	1.03215454	76.16249	8.727132	0.018944790	27.38547	311.5570	1181.760	0.1744682

**Table 2: Class Characteristics**

Class	Mean	Variance	Standard Deviation	Entropy	Skewness	Kurtosis	Contrast	Energy
0	0.01305465	0.3069204	0.727275	0.7576182	7.458341e-155			
1	0.28058077	0.3988772	0.450000	0.9017815	7.458341e-155			
2	0.01305784	0.2758204	0.405556	0.8506709	7.458341e-155			
3	0.001080263	0.2182921	0.450000	0.5747175	7.458341e-155			
4	0.029395941	0.2845713	0.484955	0.8429825	7.458341e-155			

The screenshot displays the RStudio environment. The main window shows an R script with the following code:

```

# RStudio
# Date: 2023-10-10
# Author: [Name]
# Description: [Description]

# Load data
data <- read.csv("data.csv")

# Data cleaning
data <- data[!is.na(data$age), ]
data <- data[!is.na(data$sex), ]
data <- data[!is.na(data$height), ]
data <- data[!is.na(data$weight), ]
data <- data[!is.na(data$income), ]

# Data visualization
plot(data$age, data$income)
plot(data$sex, data$income)
plot(data$height, data$weight)
plot(data$weight, data$income)

# Data analysis
summary(data)
correlation <- cor(data[, c("age", "sex", "height", "weight", "income")])
print(correlation)

# Data modeling
model <- lm(income ~ age + sex + height + weight)
summary(model)

```

The Environment pane on the right shows the following objects:

- data**: A data frame with 1000 rows and 5 columns (age, sex, height, weight, income).
- correlation**: A matrix showing the correlation between the variables.
- model**: A linear model object.

The taskbar at the bottom shows the following applications:

- Microsoft Edge
- Google Chrome
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint
- Microsoft Teams
- Microsoft OneDrive
- Microsoft Outlook
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint
- Microsoft Teams
- Microsoft OneDrive
- Microsoft Outlook

