

```
1 library(dplyr)
2 library(plyr)
3
4 data("iris")
5
6 rowNum <- nrow(iris)
7 colNum <- ncol(iris)
8 rowHead <- head(iris,10)
9 sepallLen <- iris$Sepal.Length
10 sepallLenMean <- mean(iris$Sepal.Length)
11 sepallLenRow <- subset(iris,Sepal.Length > 7.6)
12 sepallLenRowIndex <- which(iris > 7.6)
13 setosaRow <- iris[grepl("setosa", iris$Species), ]
14 setosaRowBoth <- filter(iris, Sepal.Length > 7.6 & Species == "setosa")
15 sepalMaxLength <- ddply(iris, ~Species, summarise, max=max(Sepal.Length))
16 sepalRowIndex <- which(iris >= 7.8)
17
18
19
20 print(paste("a. Get number of rows (Hint: nrow):", rowNum))
21 cat("\n")
22 print(paste("b. Get number of columns (Hint: ncol):", colNum))
23 cat("\n")
24 print("c. Show first 10 rows:")
25 rowHead
26 cat("\n")
27 print("d. Show column Sepal.Length:")
28 sepallLen
29 cat("\n")
30 print(paste("e. Calculate the mean Sepal.Length?:", sepallLenMean))
31 cat("\n")
32 print("f. Show all rows where Sepal.Length > 7.6:")
33 sepallLenRow
34 cat("\n")
35 print("fi. What are the row indexes where Sepal.Length > 7.6? (Hint: which):" )
36 sepallLenRowIndex
37 cat("\n")
38 print("g. Show all rows where Species is setosa:")
39 head(setosaRow)
40 cat("\n")
41 print("h. Show all rows where Sepal.Length > 3.0 and Species is setosa:")
42 setosaRowBoth
43 cat("\n")
44 print("i. Get the largest value of Sepal.Length:")
45 sepalMaxLength
46 cat("\n")
47 print(paste("ii. Get the row index that contains this value:", sepalRowIndex))
48 cat("\n")
49 print("j. What Species corresponds to this largest Sepal.Length?: virginica")
```

```

1 [1] "a. Get number of rows (Hint: nrow): 150"
2
3 [1] "b. Get number of columns (Hint: ncol): 5"
4
5 [1] "c. Show first 10 rows:"
6   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
7 1           5.1           3.5           1.4           0.2  setosa
8 2           4.9           3.0           1.4           0.2  setosa
9 3           4.7           3.2           1.3           0.2  setosa
10 4           4.6           3.1           1.5           0.2  setosa
11 5           5.0           3.6           1.4           0.2  setosa
12 6           5.4           3.9           1.7           0.4  setosa
13 7           4.6           3.4           1.4           0.3  setosa
14 8           5.0           3.4           1.5           0.2  setosa
15 9           4.4           2.9           1.4           0.2  setosa
16 10          4.9           3.1           1.5           0.1  setosa
17
18 [1] "d. Show column Sepal.Length:"
19   [1] 5.1 4.9 4.7 4.6 5.0 5.4 4.6 5.0 4.4 4.9 5.4 4.8 4.8 4.3 5.8 5.7 5.4 5.1
20  [19] 5.7 5.1 5.4 5.1 4.6 5.1 4.8 5.0 5.0 5.2 5.2 4.7 4.8 5.4 5.2 5.5 4.9 5.0
21  [37] 5.5 4.9 4.4 5.1 5.0 4.5 4.4 5.0 5.1 4.8 5.1 4.6 5.3 5.0 7.0 6.4 6.9 5.5
22  [55] 6.5 5.7 6.3 4.9 6.6 5.2 5.0 5.9 6.0 6.1 5.6 6.7 5.6 5.8 6.2 5.6 5.9 6.1
23  [73] 6.3 6.1 6.4 6.6 6.8 6.7 6.0 5.7 5.5 5.5 5.8 6.0 5.4 6.0 6.7 6.3 5.6 5.5
24  [91] 5.5 6.1 5.8 5.0 5.6 5.7 5.7 6.2 5.1 5.7 6.3 5.8 7.1 6.3 6.5 7.6 4.9 7.3
25 [109] 6.7 7.2 6.5 6.4 6.8 5.7 5.8 6.4 6.5 7.7 7.7 6.0 6.9 5.6 7.7 6.3 6.7 7.2
26 [127] 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8
27 [145] 6.7 6.7 6.3 6.5 6.2 5.9
28
29 [1] "e. Calculate the mean Sepal.Length?: 5.84333333333333"
30
31 [1] "f. Show all rows where Sepal.Length > 7.6:"
32   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
33 118           7.7           3.8           6.7           2.2 virginica
34 119           7.7           2.6           6.9           2.3 virginica
35 123           7.7           2.8           6.7           2.0 virginica
36 132           7.9           3.8           6.4           2.0 virginica
37 136           7.7           3.0           6.1           2.3 virginica
38
39 [1] "fi. What are the row indexes where Sepal.Length > 7.6? (Hint: which):"
40 [1] 118 119 123 132 136
41
42 [1] "g. Show all rows where Species is setosa:"
43   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
44 1           5.1           3.5           1.4           0.2  setosa
45 2           4.9           3.0           1.4           0.2  setosa
46 3           4.7           3.2           1.3           0.2  setosa
47 4           4.6           3.1           1.5           0.2  setosa
48 5           5.0           3.6           1.4           0.2  setosa
49 6           5.4           3.9           1.7           0.4  setosa
50
51 [1] "h. Show all rows where Sepal.Length > 3.0 and Species is setosa:"
52 [1] Sepal.Length Sepal.Width Petal.Length Petal.Width Species
53 <0 rows> (or 0-length row.names)
54
55 [1] "i. Get the largest value of Sepal.Length:"
56   Species max
57 1      setosa 5.8

```

```
58 2 versicolor 7.0
59 3 virginica 7.9
60
61 [1] "ii. Get the row index that contains this value: 132"
62
63 [1] "j. What Species corresponds to this largest Sepal.Length?: virginica"
```

2. Consider the answer posted to Quora.com to “Why is R great for Data Science?”. Answer one of the following questions.

One online community where members discuss R is found at a youtube channel named “freeCodeCamp.org” (<https://www.youtube.com/watch?v=V8eKsto3Ug>). This channel specializes in creating online lectures for a wide variety of coding languages, with R being included. All lectures on this channel are made completely free which makes learning extremely friendly to newcomers. To add on to this, all the R tutorial files he goes through in this video are provided via google docs. The age range is various but most community members in this channel are primarily students.