HW_03

DerekRBratcher

10/7/2021

Homework 3

1. (2 points) Suppose we have a dataset A (see code below) where each column represents multiple measures of nitrogen concentration in a particular lake. We want to get the average value for each lake. Do this in two ways: a `for` loop and a vectorized function `colMeans()`.

```
set.seed(12) # to be reproducible
A = matrix(data = runif(n =1:500), nrow = 50, ncol = 10)
colnames(A) = paste("lake", 1:10, sep = "_")
```

- find mean via `for` loop

```
for(i in 1:ncol(A)){
  print(mean(A[ ,i]))
}
```

```
## [1] 0.4601492
## [1] 0.4992815
## [1] 0.5987037
## [1] 0.4580486
## [1] 0.4719578
## [1] 0.4965216
## [1] 0.5110536
## [1] 0.4577936
## [1] 0.5193423
## [1] 0.4856413
```

- find mean via `colMeans()`

```
colMeans(A)
```

```
##    lake_1    lake_2    lake_3    lake_4    lake_5    lake_6    lake_7    lake_8
## 0.4601492 0.4992815 0.5987037 0.4580486 0.4719578 0.4965216 0.5110536 0.4577936
##    lake_9   lake_10
## 0.5193423 0.4856413
```

2. (2 points) From the for loop lecture, we see the following example of using apply():

```
x = array(1:27, dim = c(3, 3, 3))
apply(X = x, MARGIN = c(1, 2),
      FUN = paste, collapse = ", ")
```

```
##      [,1]        [,2]        [,3]
## [1,] "1, 10, 19" "4, 13, 22" "7, 16, 25"
## [2,] "2, 11, 20" "5, 14, 23" "8, 17, 26"
## [3,] "3, 12, 21" "6, 15, 24" "9, 18, 27"
```

- Now, use `for` loops to get the same task done (hint: nested loops). The results should be the same.

```
x1 <- x[ 1:3, 1:3, 1]
x2 <- x[1:3, 1:3, 2]
x3 <- x[1:3, 1:3, 3]
for (i in x1) {
  for (i in x2)
    for (i in x3) {
    }
  }
   print(as.array(paste(x1, x2, x3, sep = ", ")))
```

```
## [1] "1, 10, 19" "2, 11, 20" "3, 12, 21" "4, 13, 22" "5, 14, 23" "6, 15, 24"
## [7] "7, 16, 25" "8, 17, 26" "9, 18, 27"
```

```
#how do i direct my loop to print by dimension??
```

3. (2 points) The Fibonacci Sequence is the series of numbers that the next number is the sum of the previous two numbers: 0, 1, 1, 2, 3, 5, 8 ... Use a for loop to get the first 30 numbers of the Fibonacci Sequence. This question should demonstrate the need for loops because there is no easy way to use vectorized functions in this case.

```
#set up a numeric vector of 40 elements and assign to fibonacci
Fibonacci<- numeric(30)
#make the first two elements 1
Fibonacci[[1]]<- Fibonacci[[2]]<- 1
#set up loop applying formula at 3-30
for (i in 3:length(Fibonacci)) {
  Fibonacci[i]<- Fibonacci[i-2] + Fibonacci[i-1]
}
print("The First 30 Fibonaci numbers")
```

```
## [1] "The First 30 Fibonaci numbers"
```

```
 print(Fibonacci)
```

```
##  [1]      1      1      2      3      5      8     13     21     34     55
## [11]     89    144    233    377    610    987   1597   2584   4181   6765
## [21]  10946  17711  28657  46368  75025 121393 196418 317811 514229 832040
```

4. (2 points) In the example data below, extract those ranking numbers with *regular expression*. The results should have the number(s) and . if it follows after the numbers immediately (i.e., `1.`, `12.`, `105.`, `105.3`, etc.). Remove empty strings from the final results. You should get 107 strings for your results. add 54. Electronic - Tighten Up and 55. Enigma - Sadeness (Part 1)

```
top105 = readLines("http://www.textfiles.com/music/ktop100.txt")
top105 = top105[-c(64, 65)] # missing No. 54 and 55
top105
```

```
##   [1] "From: ed@wente.llnl.gov (Ed Suranyi)"
##   [2] "Date: 12 Jan 92 21:23:55 GMT"
##   [3] "Newsgroups: rec.music.misc"
##   [4] "Subject: KITS' year end countdown"
##   [5] ""
##   [6] ""
##   [7] "On Jan. 1, 1992, the \"Modern Rock\" station KITS San Francisco (\"Live-105\")"
##   [8] "broadcast its list of the \"Top 105.3 of 1991.\"  Here is the countdown"
##   [9] "list:"
##  [10] ""
##  [11] "1. NIRVANA                      SMELLS LIKE TEEN SPIRIT"
##  [12] "2. EMF                          UNBELIEVABLE"
##  [13] "3. R.E.M.                       LOSING MY RELIGION"
##  [14] "4. SIOUXSIE & THE BANSHEES      KISS THEM FOR ME"
##  [15] "5. B.A.D. II                    RUSH"
##  [16] "6. RED HOT CHILI PEPPERS        GIVE IT AWAY"
##  [17] "7. ELECTRONIC                   GET THE MESSAGE"
##  [18] "8. ERASURE                      CHORUS"
##  [19] "9. SCHOOL OF FISH               3 STRANGE DAYS"
##  [20] "10. NORTHSIDE                   TAKE FIVE"
##  [21] "11. JESUS JONES                 INTERNATIONAL BRIGHT YOUNG THING"
##  [22] "12. DIVINYLS                    I TOUCH MYSELF"
##  [23] "13. SIMPLE MINDS                SEE THE LIGHTS"
##  [24] "14. OMD                         PANDORA'S BOX"
##  [25] "15. JAMES                       SIT DOWN"
##  [26] "16. U2                          MYSTERIOUS WAYS"
##  [27] "17. PSYCHEDELIC FURS            UNTIL SHE COMES"
##  [28] "18. MOTORCYCLE BOY              HERE SHE COMES"
##  [29] "19. MATERIAL ISSUE              VALERIE LOVES ME"
##  [30] "20. R.E.M.                      SHINY HAPPY PEOPLE"
##  [31] "21. B.A.D. II                   THE GLOBE"
##  [32] "22. NED'S ATOMIC DUSTBIN        HAPPY"
##  [33] "23. SEVEN RED SEVEN             THINKING OF YOU"
##  [34] "24. BILLY BRAGG                 SEXUALITY"
##  [35] "25. ALISON MOYET                IT WON'T BE LONG"
##  [36] "26. PRIMUS                      JERRY WAS A RACE CAR DRIVER"
##  [37] "27. VOICE OF THE BEEHIVE        MONSTERS & ANGELS"
##  [38] "28. BLUR                        THERE'S NO OTHER WAY"
##  [39] "29. HAVANA 3 A.M.               REACH THE ROCK"
##  [40] "30. THE FIXX                    HOW MUCH IS ENOUGH"
##  [41] "31. TOP                         NUMBER ONE DOMINATOR"
##  [42] "32. THE WONDER STUFF            CAUGHT IN MY ..."
##  [43] "33. TRANSVISION VAMP            B WITH U"
##  [44] "34. ROBYN HITCHCOCK             SO YOU THINK YOU'RE IN LOVE"
##  [45] "35. CHAPTERHOUSE                PEARL"
```

```
##  [46] "36. GARY CLAIL                  HUMAN NATURE"
##  [47] "37. MOODSWINGS                  SPIRITUAL HIGH"
##  [48] "38. THIS PICTURE                NAKED RAIN"
##  [49] "39. SHAMEN                      MOVE MOUNTAINS"
##  [50] "40. RATCAT                      THAT AIN'T BAD"
##  [51] "41. KITCHENS OF DISTINCTION     DRIVE ..."
##  [52] "42. STING                       ALL THIS TIME"
##  [53] "43. CANDY FLIP                  RED HILLS ROAD"
##  [54] "44. THE PIXIES                  LETTER TO MEMPHIS"
##  [55] "45. JUDYBATS                    NATIVE SON"
##  [56] "46. THE OCEAN BLUE              CERULEAN"
##  [57] "47. VOICE FARM                  FREE LOVE"
##  [58] "48. SIOUXSIE & THE BANSHEES     SHADOWTIME"
##  [59] "49. SEAL                        CRAZY"
##  [60] "50. RIGHT SAID FRED             I'M TOO SEXY"
##  [61] "51. MORRISSEY                   SING YOUR LIFE"
##  [62] "52. ERASURE                     LOVE TO HATE YOU"
##  [63] "53. MANIC ST. PREACHERS         STAY BEAUTIFUL"
##  [64] "56. SISTERS OF MERCY            DETONATION"
##  [65] "57. KIRSTY MACCOLL              WALKING DOWN MADISON"
##  [66] "58. THE PRIMITIVES              THE WAY YOU ARE"
##  [67] "59. TEENAGE FANCLUB             STAR SIGN"
##  [68] "60. THE FARM                    ALL TOGETHER NOW"
##  [69] "61. THE DYLANS                  PLANET LOVE"
##  [70] "62. TOO MUCH JOY                CRUSH STORY"
##  [71] "63. MINISTRY                    JESUS BUILT MY HOTROD"
##  [72] "64. PRIMAL SCREAM               MOVIN' ON UP"
##  [73] "65. WIR                         SO AND SLOW IT GROWS"
##  [74] "66. THE MISSION U.K.            HANDS ACROSS ..."
##  [75] "67. INTERNATIONAL BEAT          ROCK STEADY"
##  [76] "68. SQUEEZE                     SATISFIED"
##  [77] "69. NITZER EBB                  FAMILY MAN"
##  [78] "70. I START COUNTING            STILL SMILING"
##  [79] "71. VIOLENT FEMMES              AMERICAN MUSIC"
##  [80] "72. THE MILLTOWN BROTHERS       WHICH WAY ..."
##  [81] "73. HAPPY MONDAYS               BOB'S YER UNCLE"
##  [82] "74. CAMOUFLAGE                  HEAVEN I WANT YOU"
##  [83] "75. MOCK TURTLES                CAN YOU DIG IT?"
##  [84] "76. CROWDED HOUSE               IT'S ONLY NATURAL"
##  [85] "77. POPINJAYS                   VOTE ELVIS"
##  [86] "78. CARTER U.S.M.               THIS IS HOW ..."
##  [87] "79. THE LA'S                    I CAN'T SLEEP"
##  [88] "80. ST. ETIENNE                 ONLY LOVE CAN BREAK YOUR HEART"
##  [89] "81. ENYA                        CARRIBEAN BLUE"
##  [90] "82. PRESENCE                    IN WONDER"
##  [91] "83. PET SHOP BOYS               WHERE THE STREETS HAVE NO NAME   (tie)"
##  [92] "83. SPIREA-X                    SPEED REACTION                   (tie)"
##  [93] "84. THE WENDY'S                 HALFPIE"
##  [94] "85. KATE BUSH                   ROCKET MAN"
##  [95] "86. CANDY SKINS                 SHE BLEW ME AWAY"
##  [96] "87. ORB                         PERPETUAL DAWN"
##  [97] "88. BIRDLAND                    SHOOT YOU DOWN"
##  [98] "89. TIN MACHINE                 BABY UNIVERSAL"
##  [99] "90. SINGLE GUN THEORY           FROM A MILLION"
```

```
## [100] "91. NED'S ATOMIC DUSTBIN        GREY CELL GREEEN                (tie)"
## [101] "91. XYMOX                       PHOENIX OF MY HEART             (tie)"
## [102] "92. LUSH                        DE-LUXE                        "
## [103] "93. SCATTERBRAIN                DOWN WITH THE SHIP"
## [104] "94. EON                         SPICE"
## [105] "95. SMITHEREENS                 TOP OF THE POPS"
## [106] "96. G. W. McLENNAN              EASY COME, EASY GO"
## [107] "97. KLF                         LAST TRAIN TO TRANSCENTRAL      (tie)"
## [108] "97. HOODOO GURUS                MISS FREELOVE '69               (tie)"
## [109] "98. ANTHRAX                     BRING THE NOISE   "
## [110] "99. MARY'S DANISH               JULIE'S BLANKET"
## [111] "100. MEAT PUPPETS               SAM"
## [112] "101. SMASHING PUMPKINS          SIVA"
## [113] "102. ELVIS COSTELLO             OTHER SIDE OF ..."
## [114] "103. SEERS                      PSYCHE OUT"
## [115] "104. THRILL KILL CULT           SEX ON WHEELZ"
## [116] "105. MATTHEW SWEET              I'VE BEEN WAITING"
## [117] "105.3  LATOUR                   PEOPLE ARE STILL HAVING SEX"
## [118] ""
## [119] "Ed"
## [120] "ed@wente.llnl.gov"
## [121] ""
```

```
top_105_ranking_numbers <-stringr::str_sub(top105,start = 1, end =  5)
top_105_ranking_numbers <- top_105_ranking_numbers |>  gsub(pattern = "[[:alpha:]]",replacement =  " ",
top_105_ranking_numbers<- sub("@", "", top_105_ranking_numbers, fixed = TRUE)
 top_105_ranking_numbers <- top_105_ranking_numbers |>  gsub(pattern = "[[:blank:]]",replacement =  "",
top_105_ranking_numbers<- stringr::str_replace_all( top_105_ranking_numbers,":", "")

library(stringi)
top_105_ranking_numbers<- stringi::stri_remove_empty(top_105_ranking_numbers, na_empty = FALSE)
top_105_ranking_numbers<- sub("3..", "3.", top_105_ranking_numbers, fixed = TRUE)
top_105_ranking_numbers<- sub("5..", "5.", top_105_ranking_numbers, fixed = TRUE)
top_105_ranking_numbers
```

```
##   [1] "1."    "2."    "3."    "4."    "5."    "6."    "7."    "8."    "9."
##  [10] "10."   "11."   "12."   "13."   "14."   "15."   "16."   "17."   "18."
##  [19] "19."   "20."   "21."   "22."   "23."   "24."   "25."   "26."   "27."
##  [28] "28."   "29."   "30."   "31."   "32."   "33."   "34."   "35."   "36."
##  [37] "37."   "38."   "39."   "40."   "41."   "42."   "43."   "44."   "45."
##  [46] "46."   "47."   "48."   "49."   "50."   "51."   "52."   "53."   "56."
##  [55] "57."   "58."   "59."   "60."   "61."   "62."   "63."   "64."   "65."
##  [64] "66."   "67."   "68."   "69."   "70."   "71."   "72."   "73."   "74."
##  [73] "75."   "76."   "77."   "78."   "79."   "80."   "81."   "82."   "83."
##  [82] "83."   "84."   "85."   "86."   "87."   "88."   "89."   "90."   "91."
##  [91] "91."   "92."   "93."   "94."   "95."   "96."   "97."   "97."   "98."
## [100] "99."   "100."  "101."  "102."  "103."  "104."  "105."  "105.3"
```

5. (2 points) For the vector with length of 107 you got from question 4, remove all trailing .. (hint: `?sub`). Then convert it to a numeric vector and find out which numbers have duplications (i.e., a tie in ranking). Don't count by eyes, use R to find it out (hint: `table(), sort(); or duplicated(), which(), [` subsetting; there are more than one way to do so).

```
top_105_ranking_numbers<- as.numeric(top_105_ranking_numbers)
table(duplicated(top_105_ranking_numbers))
```

```
##
## FALSE  TRUE
##   104     3
```

```
table(top_105_ranking_numbers)
```

```
## top_105_ranking_numbers
##     1     2     3     4     5     6     7     8     9    10    11    12    13
##     1     1     1     1     1     1     1     1     1     1     1     1     1
##    14    15    16    17    18    19    20    21    22    23    24    25    26
##     1     1     1     1     1     1     1     1     1     1     1     1     1
##    27    28    29    30    31    32    33    34    35    36    37    38    39
##     1     1     1     1     1     1     1     1     1     1     1     1     1
##    40    41    42    43    44    45    46    47    48    49    50    51    52
##     1     1     1     1     1     1     1     1     1     1     1     1     1
##    53    56    57    58    59    60    61    62    63    64    65    66    67
##     1     1     1     1     1     1     1     1     1     1     1     1     1
##    68    69    70    71    72    73    74    75    76    77    78    79    80
##     1     1     1     1     1     1     1     1     1     1     1     1     1
##    81    82    83    84    85    86    87    88    89    90    91    92    93
##     1     1     2     1     1     1     1     1     1     1     2     1     1
##    94    95    96    97    98    99   100   101   102   103   104   105 105.3
##     1     1     1     2     1     1     1     1     1     1     1     1     1
```

duplicates @ 83,91,97.