Vectors

1. Subscripts. Create a 20 element numeric vector
    1.a. Access element 4
    1.b. Access elements 11 to 15
    1.c. Access elements 6, 9, 12, 15, 18
    1.d. Access elements 3, 4, 5, 4, 3
    1.e. Access all elements greater than the mean
    1.f. Access all elements other than 2, 8, 13
    1.g. Create a 100 element vector with values < 20. Subscript the 20 element vector by the 100 element vector.
    1.h. Create a 100 element vector with some values > 20. Subscript the 20 element vector by the 100 element vector.
    1.i. Perform 1a) - 1d) and 1f) for a character vector

2. Ordering. Create a 10 element numeric vector
    2.1. Return the vector of indices of each element if the array were to be sorted in ascending order. Do the same for descending order
    2.2. Return the rank of each element in the vector. Average the rank if you find a tie.
    2.3. Without using function sort(), return the vector from 1) in ascending order. Do the same for descending.

3. Statistics. Create a 25 element numeric vector
    3.1. Calculate the length, sum, product, min and max of the data set
    3.2. Calculate the mean, median, variance and standard deviation of the data set
    3.3. Calculate the cumulative sum of the vector

4. Applying functions.
    4.a) To a 15 element vector of angles in degrees, obtain the tangent. Then obtain the arc-tangent of the result.
    4.b) Obtain the exponent of the series: 3,6,9,12,15. Return the result as a vector.

5. Subdivision/sampling. Create an 18 element numeric vector.
    5.a) Split the vector into 3 groups and assign each group a name
    5.b) Split the vector into 4 groups. Then split it into 5 groups
    5.c) Obtain a sample of 5 from the vector without replacement
    5.d) Obtain a sample of 2 from the 1st group in 5a) with replacement

Arrays

1. Subscripts: Create a 3x2x5 dimensional array
    1.a) Access element 12.
    1.b) Access elements c[1:3,1:2,1:2]
    1.c) Access elements lesser than the mean
    1.d) Access all data except for index 3 dimension 1

2. Outer product
    2.a) Obtain the outer product of a 3x4 and a 4x2 array using /
    2.b) Obtain the outer product of the same arrays using *

3. Applying functions
    3.a) Apply a mathematical function - log10 - to a 3x4x2 array. Do it by each dimension.
    3.b) Apply a statistical function - var - to the same array. Do it by each dimension.

Matrices

1. Subscripts. Create a 6 x 7 numeric matrix
   - 1.a) Obtain element 3, 4
   - 1.b) Obtain elements of row 4 and column 2
   - 1.c) Create an index matrix to obtain elements (1,2), (4,3), (2,5) and (6,1)

2. Diagonal matrix
   - 2.a) Create a diagonal matrix out of a 5 element numeric vector
   - 2.b) Create a 4x4 character matrix and return the diagonal as a vector
   - 2.c) Create a 6x6 identity matrix

3. Matrix multiplication
   - 3.a) Obtain the product of 2x3 matrix and a 3x5 matrix
   - 3.b) Obtain the product of a 2x3 matrix and a 2x5 matrix

4. Cross Product
   - 4.a) Obtain the product of a 3x4 matrix and a 3x5 matrix
   - 4.b) Obtain the product of a 3x2 matrix and a 6x2 matrix

5. Inverse: Obtain the inverse of the following matrices:

```
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    4
```

```
     [,1] [,2] [,3]
[1,]    1    0    5
[2,]    0    1    0
[3,]    0    0    1
```

```
     [,1] [,2] [,3]
[1,]    0    1    0
[2,]    1    0    0
[3,]    0    0    1
```

6. Solve the following linear equations:

$$\begin{aligned} 8x + 7y &= 38 \\ 3x - 5y &= -1 \end{aligned}$$

$$\begin{aligned} 5x + 2y &= 4 \\ -2x + y &= 11 \end{aligned}$$

$$\begin{aligned} x + y &= 1 \\ 2x + 3y &= 3 \end{aligned}$$

7. Using data frame *mtcars* in package *datasets*, perform a linear regression between mileage [mpg], against the number of cylinders [cyl], horse power [hp] and weight [wt].

Data Frames

1. Use the following data to create a data frame

```
Phone, Maker, Price, Country, No_Sold, OperSys, No_Apps, Carrier
iPhone, Apple, 399, USA, 2687161, iOS, 3000000, AT&T
Galaxy, Samsung, 350, Korea, 256121, Android, 5716247, Verizon
Razr, Motorola, 200, USA, 26511, Android, 12381, Sprint
Pearl, Blackberry, 399, Canada, 125819, Blackberry, 123701, Rogers
Optimus One, LG, 299, Korea, 123291, Android, 12312, AT&T
Lumia 800, Nokia, 299, Finland, 23432, Microsoft, 87699, Verizon
```

2. Create a data frame out of the first 3 columns

3. Accessing one or more rows
    - 3.a) Obtain rows 2, 3 and 5
    - 3.b) Obtain the first 3 rows. Then obtain the last 2 rows
    - 3.c) Pull all rows for Android phones

4. Adding columns
    - 4.a) Add a column: “Camera_Res” to this data frame. Create this column out of a vector
    - 4.b) Add a column: “Camera_Res” to this data frame. Create this column out of a list

5. Adding rows
    - 5.a) Add a row to this data frame. Create this row out of a vector
    - 5.b) Add rows from the data frame below to this data frame

```
Phone, Maker, Price, Country, No_Sold, OperSys, No_Apps, Carrier, Camera_Res
Xperia S, Sony, 300, Japan, 79792, Android, 121211, AT&T, 6 Megapixels
One X, HTC, 250, Taiwan, 99191, Android, 1312, Verizon, 6 Megapixels
```

6. Combine columns 1, 3 and 4 from both data frames [before performing 5b)] to form a new data frame

7. Obtaining summaries
    - 7.a) From Edgar Anderson’s *iris* data, obtain the average of Sepal/Petal lengths and widths by species
    - 7.b) From Edgar Anderson’s *iris* data, obtain the standard deviation of Sepal/Petal lengths and widths by species

8. Changing columns: In the phone data in 1)
    - 8.a) Provide Price in Euros
    - 8.b) Provide revenue data [Price x No Sold]

Factors

1. Evaluating functions: Use table *PlantGrowth* in package *datasets*
    1.a) Obtain the mean weight per group. Then obtain the standard deviation per group
    1.b) Obtain the median weight per group

2. Creating levels: Create a numeric vector of length 30 and range 10. Divide this into 5 groups and generate a factor

3. Generate a numeric factor with 6 levels and 7 replicas. Then assign names to the levels

4. Re-order the Species data in Edgar Anderson's *iris* data based on median Petal length

5. Create a vector of length 3. Create a factor of length 24. Return elements of the vector for the entire range of the factor.

Text data

1. Length and parts of a string
    1.a) Obtain the length of a string
    1.b) Obtain the lengths of a character vector of 10 elements
    1.c) Obtain 5 characters starting with the 4th, from a string.
    1.d) In the sentence “How much does Ann shovel in an hour?”, replace “Ann” with “Joe”

2. Concatenating strings
    2.a) Combine the sentences: “Hi! How do you do?” and “My name is Tom”.
    2.b) Create a paragraph out of 3 sentences.

3. Pattern match: “She sells sea shells by the sea shore”
    3.a) With respect to this sentence, provide a list of words with the letter e in each.
    3.b) Provide the positions of all words with the letters ea in them
    3.c) Repeat 3a) and 3b) for the letter s, matching case

4. Alphabet: Obtain the word “pulchritude” from a numeric vector

5. Pattern replacement: `“R is an open source programming language and software environment for statistical computing and graphics. The R language is widely used among statisticians for developing statistical software[2][3] and data analysis”`

    5.a) Replace the letters “an” above. Use a phrase whose length > 2
    5.b) Replace the first occurrence of the letters “ti”. Use a 3 letter phrase.

Dates

1. Creation
    1.a) Assign the date: Oct 2, 2012 to a variable
    1.b) Create a vector of 10 dates using format DD-MM-YYYY. Repeat using format MM-DD-YYYY
    1.c) Return the date for 13 days from Jul 21, 2009. Repeat using arithmetic.

2. Formatting
    2.a) Return the vector from 1b) in format MM/DD/YYYY
    2.b) Return the System date in format DD/MM/YYYY

3. POSIX
    3.a) Store the date: Jan 24, 2011 as a double
    3.b) Store the date: Feb 1 - 7, 2004 as a list
    3.c) Return the vector from 1b) in the format “January xx, xxxx”
    3.d) Return the vector from 1b) in the format “Jan xx, xxxx”

4. ISO8601 specifies that default date format is YYYY-MM-DD. True or False?