Elements of Statistics and Probability

Exercises using R

1. Using the vectors

`x <- c(-4, 1, 0, 0, 5, -3, 2)` and `y <- c(1, 1, -5, 4, 3, -2, 0)` implement the following operations:

   i. Create a new vector including the first, second and sixth element of `x`
   ii. Create a new vector that includes the first four elements of `y`
   iii. Create a new vector with the negative elements of `x`
   iv. Create a new vector that includes all elements of `x` apart from the third one
   v. Create a new vector with the elements of `x` satisfying the conditions `x<0` and `y≠1`
   vi. Create a new vector with the elements of `y` satisfying the conditions `y<0` or `x≤0`
   vii. Create a new vector with the first two elements of `y` replicated twice.

2. Suppose that the vectors `x` and `y` include the grades in the assignment and the written examination of a course for a specific student. Calculate the final score of the student in the following cases:
   i. The final score is the average of the two grades.
   ii. If the assignment's grade is 7 or higher the final score is the grade of the written examination, otherwise it is 4. The student also gets a bonus of one credit if the assignment's grade is more than 8.
   iii. The final score is the minimum of the two grades
   iv. If the assignment's grade is at least 5 then the final score is the average of the two grades, otherwise the final score is just the grade of the written examination.

3. Let $A = \begin{pmatrix} 1 & 2 & 6 \\ 5 & 2 & 5 \\ 6 & 1 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$.
   i. Define the above tables in R.
   ii. What is the appropriate operation in order to get $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ from A?

iii. Calculate the following quantity:

$$5A^{-1} + 3(AA^T)^{-1} - 2BB^T + I_3 + \begin{pmatrix} 5 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

4. The following table includes data about the performance of 20 students:

Chemistry	Physics	Mathematics	Literature	Sex	Year
93	42	98	34	Male	1
71	67	68	33	Male	1
77	59	36	24	Male	1
78	70	92	24	Male	1
77	59	44	31	Male	1
81	50	45	22	Male	2
88	50	58	23	Female	2
74	51	31	32	Female	2
67	45	70	31	Female	2
78	64	46	26	Female	2
77	49	41	75	Male	1
67	49	46	81	Male	1
63	48	65	87	Female	1
83	51	62	100	Female	1
73	56	20	81	Female	1
70	47	22	100	Female	2
78	53	92	77	Male	2
95	56	56	89	Male	2
88	49	28	100	Male	2
75	71	94	77	Male	2

a) Enter the data into a data frame.
b) For each gender, compute the average score at each course.
c) Find the max score at each course.
d) Find the max score for all courses.
e) Compute the average score for each student and rank the students according to their mean score. Repeat the above computation and ranking for each year separately.
f) Who is the best student? Standardize the scores in order to be comparable and compute the average standardized score for each student.
g) What is the percentage of students who succeeded (score ≥ 50) at all courses?
h) What is the average score and variance for the successful students?