Questions and answers about final take home

QUESTION: *"Can we use char() or strlength() function. Because if we only have sum(), length(), zeros() function we cannot get the length of the string in the text"*
ANSWER: No. You can only use the functions mentioned in the question.

QUESTION: "in exercise 7, there's a special row in bike_trip.csv, where t_end < t_start. Should we just neglect it, or treat it as error data? Or it may mean that the bike was not returned until the second day?"
ANSWER: Treat it as erroneous data.

QUESTION: *"[...] question about question 4b. Are we allowed to use the gscatter plotting function to colour the different decades? Or are we just supposed to use the scatter function."*
ANSWER: You can use either `gscatter()` or `scatter()`, this is up to you, no difference in marking.

QUESTION: "For question 2, is it alright if we include a 'fprintf' statement at the end of our function to display our results?"
ANSWER: Yes.

Clarification Final Exam: file name
There is a typo in the name of the data file "speed-ligth.csv" for exercise 1 ("t" and "h" are swapped).
Correct it if you want, this is up to you but I suggest leaving the file name as it is. This shouldn't be an issue because you will upload the CSV file along with your Matlab script. Your script will read this file (whether you rename it or not).

QUESTION: *"For exercise 2, should our function work on string inputs as well as character vector inputs?"*
ANSWER: This is up to you. Note that you will *not* receive extra credits if it works for both. Just on one type is enough.

QUESTION: *"Are we allowed to use the functions meshgrid, scatter3, and surf in question 5(a), to draw the regression plane and recorded points in a three-dimensional coordinate system?"*
ANSWER: Yes, you can use any function you want for the figure.

QUESTION: *"About Question 3b. For the second column, the 'numeric performance' for each event, do I take the average? Or do I need to do for each participant? If I were to do for each participant, will there not be 33 columns then?*

"
ANSWER: This is the correlation between, for example, the scores of all participants for "long jump" and their total score. And so on for the nine other events. The end result is a 10x2 table.

QUESTION: *"For exercise 2: Are we allowed to use relational operators* (*Eg a = b == 1*), *colon notation* (*Eg a = 1:4*) , *or filtering with logic arrays* (*Eg b =* [*0 1*], *a= c*(*b*))*?*"
ANSWER: Yes.

QUESTION: *"For exercise 2: Does the function have to be case sensitive?"*
ANSWER: Yes.

QUESTION: *"Exercise 3c: Should the PCA be conducted on the entire dataset or just the first ten columns* (*i.e.*, *on the ten events only and not the total score column*, *or the entire dataset*) *?"*
ANSWER: I will not answer directly to this one. I simply note that the total score is calculated directly from the performances of the ten events.

QUESTION: *"Exercise 4b: Should the pairwise distances calculated in part A be used for the MDS*, *or the usual Euclidean pairwise distances?"*
ANSWER: In general, it may be a good idea to use a) to answer b)... ;)

QUESTION: *"For exercise 3a*, *can we code the subplots individually? I'm asking because I find it easier to name all the axis and graphs when I do it that way instead of writing a function."*
ANSWER: In general, a coding style that involves many "copy/paste" when a loop/function style would be possible, will have its marking negatively impacted.

QUESTION: "If the question says to make a matrix(like in 4a) or to make a table(3b), do you want us to display the result in the command window, or can we put a semi-column after the code?"
ANSWER: Do not print out arrays, tables or matrices if they are too large (more than about 50 rows/columns).

QUESTION: *Do we need to install bioinformatics toolbox for question 4?*
ANSWER: No, and you must not use special functions from any other toolkits than the standard ones we have been using for the entirety of this course.

QUESTION: *My question 4a takes a really long time to run*, *but I believe it outputs the desired n by n matrix.*
ANSWER: Your code may take a few seconds to run for this question (actual time depends on the performances of your computer). But it shouldn't take more than a

minute. If it's longer, it means you can code in a more efficient way and your mark will be affected negatively in that case.

QUESTION: *"For the final submission, would you like us to submit a published PDF version of the code as well as the .m and .csv files?"*
ANSWER: Let me be as clear as I can, again:

- upload your Matlab scripts and CSV data files
- YOUR CODE MUST RUN (without error) ON THE GRADER'S COMPUTER. Test with your friends.
- No PDF files.

The mark will be zero for a code that doesn't run. No second chance.

QUESTION: *"For question 2, can we use 'OR' within our if statements to account for both capitalized and lowercase characters? Also, do we have to account for special characters such as '!@#$%'"*
ANSWER: The function must work for any character and that includes '!' '@' etc. Any logical operator (including "OR") is accepted.

QUESTION: "... are we allowed to create a new array using, for example, a = [ ]; and add elements to this array? I'm not sure if this counts as a built-in Matlab function."
ANSWER: Yes, this is allowed

QUESTION: " does it matter if, for example, the elements of the weather column have quotation marks around them--i.e., "Sunny" as opposed to Sunny?"
ANSWER: No, it does not matter, the grading will not take this into account.

QUESTION: "For the correlation table on question 3b, would you like us to present the values on our final sorted table in absolute values or non-absolute values? "
ANSWER: Like it is asked in the question: presented in non-absolute value, but sorted with absolute value.

I have had questions from several students about using Matlab functions that we have not seen in class (for any exercise in the Final Exam).
I just want to note that:

- all answers use functions that we saw in class
- we (me & TA) will not install a new Matlab toolkit for a specific function to run when grading. If the code doesn't run on our computer, we will consider this an error (:zero point), even if you have the correct answer on your computer.
- if you use Matlab functions not seen in class, but from our standard toolkit libraries, the grading may be altered. It will be on a case by case basis. I can be flexible but, again, there is no reason for you to do so. You would take a risk to not have the full point for the question.

QUESTION: *"I was wondering if in the exercise 3c we are only allowed to have one figure or can we have multiple?"*
ANSWER: Just one, as stated in the question.

QUESTION: *"For Question 7a, the hint you gave mentioned that the join-table functions will be useful. I managed to complete this exercise without using join-table functions but several nested loops. Will I be deducted marks for not using join-table functions?"*
ANSWER: It depends on the code, but probably yes, nested loops are unlikely to yield full points for this question.