Exam Guidelines on attempting Questions

Introduction to Data Science I - CSD 114

Semester I, 2021/2022

Some Instructions on how to get started

For this Data Science Examination, you are only required to complete the Question one (1) as a COMPULSORY Question and then choose two more questions in the remaining questions to to make three questions in total.

In answering the questions in the Exam, full marks are given for explanations, clear, readable, well commented code and not just right answers or copied code from internet. Good luck!

The Exam Questions have already been structured in a Jupyter notebook that will be shared with you on Github.com/

Your are advised to work within that notebook for all questions, push your work/code/answers on your individual github.com accounts and share the link to access it for Marking/Grading. The notebook will be auto-graded.

To get started with the Exam, you will need to download the starter code or Jupyter notebook and follow its contents to the directory where you wish to complete the Exam.

If needed, make use of (the) change directory (cd) command and adopt relative file path referencing to change to the directory where you maybe working from before starting to attempt the questions. You can also find instructions for relative in the Jupyter Notebook in the "Environment Setup Instructions" of the Examination.

Throughout the Exam, you will be using the a notebook with all questions to attempt and empty cells provided.

1. Introduction:

   - During this Examination, you will carryout a range of tasks and data pre-processing and get to see it work on data. Before starting on this Data Science Programming Assessment.
   - We strongly recommend reading and reviewing the questions CLEARLY and the associated concepts.
   - Exam1.xlsx. - An Excel worksheet containing dataset for you to load and get started
   - Exam2.xlsx - An excel worksheet containing a dataset for choice question(s) depending.
   -

   (a) Submission and Grading: After completing or attempting your questions, commit your work to your github.com account, name the Repository "Final Data Science Exam". Then be sure to submit a link of the repository containing your solutions to your witi e-learning account and on email: alon.muhame@witug.org. The following is a breakdown of how each part of this exam will be scored.

   (b) Finally, under any circumstances that any student has shared or copied another students code on any question will be awarded ZERO(0) marks. Students are not allowed to discuss in groups or individually, but only to do their own research.

Question	Content	Marks
Question 1	**Matrices; NumPy Basics; Arrays and Vectorized Computation; Array Manipulation, Data Pre-processing and manipulation**	(25 Marks)
Question 2	**Data Loading, Storage, and File Formats: Data Pre-processing and analysis, descriptive stats module in python**	(15 marks)
Question 3	**Graphing, Plotting in Matplotlib, seaborn, plotly, Descriptive statistics using stats module in python, and Introduction to algorithms:** ..	(20 marks)
Question 4	**Built-in Data Structures, Functions, and Files**	(15 marks)

i. Here below are some helpful commands for somehow making your work easier to navigate. (Comment: Usage of these commands wont attract any mark, of course)

ii. These commands are supplied to save your time only, in case you need to use them.

Command	Action
(h) elp	Display command list
help (command)	Show documentation for ¡command¿
c(ontinue)	Resume program execution
q(uit)	Exit debugger without executing any more code
b(reak) (number)	Set breakpoint at (number) in current file
b (path/to/file.py:number)	Set breakpoint at line (number) in specified file
s(tep)	Step into function call
n(ext)	Execute current line and advance to next line at current level
u(p)/d(own)	Move up/down in function call stack
a(rgs)	Show arguments for current function
debug (statement)	Invoke statement (statement) in new (recursive) debugger
l(ist) (statement)	Show current position and context at current level of stack
hline w(here)	Print full stack trace with context at current position