Submission Checksheet

Problem statement

For this project you must create a data set by simulating a real-world phenomenon of your choosing. You may pick any phenomenon you wish – you might pick one that is of interest to you in your personal or professional life. Then, rather than collect data related to the phenomenon, you should model and synthesise such data using Python. We suggest you use the numpy.random package for this purpose.

Specifically, in this project you should:

1. Choose a real-world phenomenon that can be measured and for which you could collect at least one-hundred data points across at least four different variables.
2. Investigate the types of variables involved, their likely distributions, and their relationships with each other.
3. Synthesise/simulate a data set as closely matching their properties as possible.
4. Detail your research and implement the simulation in a Jupyter notebook
5. The data set itself can simply be displayed in an output cell within the notebook.

Research

1. provide clear evidence that you developed new knowledge and skills while completing the assignment.
2. providing references
3. demonstrate in discussion how and why you considered and/or adopted the approaches of others to the problem.
4. Make it clear why you adopted the approach you did, comparing it to the other approaches.
5. Make sure to search online for various approaches to the problem, including other people's attempts at solving a similar problem.
6. Summarise the results of this search in the project documentation.
7. Use concepts and skills you learned in other subjects. The ability to tie all your various modules together is important.

Development

1. Consider your findings from researching the problem and use them to develop a plan for the project. A good plan will be evident in the layout and organisation of your code.
2. Demonstrate how you made sure your software or project did what you said it would do at the start. For instance, you might consider devising some small tests that demonstrate that your submission provides the right outputs.

Consistency

1. You should draw up a plan for completing an assignment and make the plan evident in your submission. You should try to stick to this plan, and your project timeline should clearly demonstrate your work towards it.

Add proof of Validation

1. add proof of functional validation of list generated after list section.

Minimum standard

1. The minimum standard for this project is a git repository containing a README, a gitignore file and a Jupyter notebook.
2. The README need only contain an explanation of what is contained in the repository and how to run the Jupyter notebook. Your notebook should contain the main body of work and should list all references used in completing the project.
3. A good submission will be clearly organised and
4. contain concise explanations of the particularities of the data set.
5. The analysis contained within the notebook will be well conceived,
6. interesting, and
7. well researched.

Note that part of this project is about the use of Jupyter notebooks and so you should

1. make use of all the functionality available in the software including

1. images,
2. links,
3. code and
4. plots.

You may use any Python libraries that you wish, whether they have been discussed in class or not.

Documentation

1. Consider various audiences when documenting your work. A technical audience, such as a lecturer, will require one style of description while a non-technical audience, such as a recruiter, will require another. Unfortunately, most of the time your work needs to speak to both.

From:
http://hdip-data-analytics.com/ - **HDip Data Analytics**

Permanent link:
http://hdip-data-analytics.com/submissions/assesment/checksheet

Last update: **2020/06/20 14:39**