DS-542 - Python in Data Science Cource Syllabus

Course Description

Hi, Student!

This is the introductory course to cover a base on python for data science. In this course, we will go over the basics of python, and then we will move our way into the python ecosystem of the tools and packages that's been used to work on data science projects.

The emphasis of the course will follow some of the basic steps of data science:

- Data preperation
- Manipulation
- Data Visualization
- Use of Machine Learning

During this journey, we will explore some theoretical knowledge on some concepts, such as matrices, and while learning the core of python, you will go over on various python packages, such as numpy, pandas, matplotlib, scikit-learn, statmodels, and other machine learning libraries such as keras, tensorflow, etc.

Additional information about this class can be found on SPU's website, which requires student's login.

Prerequisites

This course requires Introduction to Data Science (DS-510), and Data Analysis and Decision Modeling (DS-520) to be taken. You won't have to know python for this course, python will be thought as part of this class.

Check out more information about the prerequisites on curriculum of the SPU's data science.

Learning Outcomes

This course aims to teach you the following concepts.

- You will feel comfortable with reading, writing, and using python as a general purpose programming language.
- You will be able to understand big data science ecosystem and be able to understand its tools and methods using python.

Textbooks & Tools

Python Data Science Handbook by Jake VanderPlas.

```
Title: Python Data Science Handbook
Author(s): Jake VanderPlas
Release date: November 2016
Publisher(s): O'Reilly Media, Inc.
ISBN: 9781491912058
```

Think Python, 2nd Edition by Allen B. Downey.

```
Title: Think Python, 2nd Edition
Author(s): Allen B. Downey
Release date: December 2015
Publisher(s): O'Reilly Media, Inc.
ISBN: 9781491939369
```

Learning with Python 3 (RLE) by Peter Wentworth, Jeffrey Elkner, Allen B. Downey, and Chris Meyers.

```
Title: Learning with Python 3 (RLE)
Author(s): Peter Wentworth, Jeffrey Elkner, Allen B. Downey, and Chris
Meyers
Release date: October 2012
```

There will be more books from a project called Runestone Interactive. I have already created a class for us in this website with the class name as `DS-542`. What you have to do is the following:

1. Register to the `Runestone Interactive` from their website. Follow their tutorial if you have any issues.
2. Enter the course name as `DS-542`.
3. You should be registered to the class!

Grade Determination

The items that student is responsible for this class is the following. Students will be graded according to this chart.

Item	Percentage
Assignments	30 %
Projects	30 %
Final	30 %
Participation	10 %
Bonus (Maybe)	10 %

Below list is subject to change.

Grade Scale

Based on the Grade Determination, a final letter grade will be assigned to student based on the following scale.

Min	Max	Grade
94	100	A
87	92	A-
83	86	B+
80	82	B
77	79	B-
73	76	C+
67	72	C
60	66	C-
51	60	D
0	50	F

Course Outline

During the term of this course, following materials will be covered.

Week #	Content Description	Assignments
Week 1	Introduction to course, syllabus, student's preparation with tools, etc.	Welcome
Week 2	IDEs (Jupyter, iPython, VSCode, etc.) Environment setup, Test driven Development (TDD), Debugging, Unittesting with Python, Continuous Integration Tools, etc.	
Week 3	Introduction to python. Learning python as general-purpose programming language. Section includes python syntax, functions, classes, inheritance, MRO, Python's GIL paradigm, object oriented programming, etc.	Assignment 1
Week 4	Introduction to data manipulation tools, numpy.	Assignment 2
Week 5	Data manipulation tools, pandas part 1 - common pandas concepts, pandas as data manipulation tool.	Assignment 3
Week 5	Data manipulation tools, pandas part 2 - advanved use of pandas, performance improvements, etc.	Assignment 4
Week 6	Midterm	
Week 7	Data visualization tools, matplotlib, seaborn, plotly, etc.	Assignment 5

Week #	Content Description	Assignments
Week 8	Use of machine learning tools, scikit-learn	
Week 9	Use of advanced machine learning tools - keras, tensorflow	
Week 10	Final Exam	
Week 11	Student Project Presentations	
Week 12	Whats next? Use of containers in data science, model deployment, use of cloud systems, etc.	

Above list is subject to change.

Attendence Policy

Student missing more than 2 classes will be withdrawn with an F grade. It is the responsibility of student to notify the instructor well in advance. In case of medical emergency, student should update instructor via email.

Academic Honesty and Student Conduct

Students need to submit **only their own original work** (e.g. Code, PPT, figures, visualizations...). Student need to familiarize themselves with the academic rules of the University. In case a student is found guilty under the act of plagiarism, his/her test or assignment will be graded zero.

If plagiarism occurs twice, student will receive 'F' grade with immediate effect. It is expected that everybody turns off/mute all devices that emit sounds and noises that may interrupt the class (e.g. mobile phones, pagers, watch alarms). If an occasion arises, in which a student may need to leave the class to receive a phone call (important call), use rest room or get drinking water, he or she should silently walk out without disturbing rest of the class.

Working on assignments and project work, that belongs to another course is **STRICTLY NOT ALLOWED**.

Copying your colleagues code is **STRICTLY NOT ALLOWED**. All submissions will be submitted to autodetection plagiarism tool that our university provides!

Agreement between You and the Instructor

I certify that in this course, my contribution and assignments will be my own work, based on my personal study and/or research and that I am acknowledging all material and sources will be used in its preparation, whether they be books, articles, reports, lecture notes, and any other kind of document, electronic or personal communication.

I also certify that that in this course, the assignments I will be submitting will not previously been submitted for assessment in any other unit, except where specific permission has been granted from all unit coordinators involved, or at any other time in this unit, and that I have not copied in part or whole or otherwise plagiarized the work of other students and/or persons.

Student		**Instructor**	
Student's Name	.....	Instructor's Name	.....
Signature	.....	Signature	.....