1. Course Code
2262

2. Course Title
G52e: Data Science Exercises

3. Teacher
HAMIDULLAH, Sokout

4. Term
Spring 2

5. Course Requirements (Courses / Knowledge prerequisite for this course)
Data Science Course (2261)

6. Course Overview and Objectives
The major goal of Data Science is to solve companies (institutions, government, private sectors, .. etc) problems using available data. Data Science entails everything that has something to do with data such as: Collecting the data, Cleaning up the data, Visualizing data, Analyzing it, and Creating predictive models based on the data. Playing with data will paves the way to gain insights that could help us with good decision making.   In this course we will starting from the very basics (pre-processing data), we will build up your skills and soon you will be able to solve advanced statistics tasks using machine learning algorithms in Python such as CLUSTERING, REGRESSION, CLASSIFICATION. As well as you learn with integration and visualization with Python.

7. Course Outline	
1	Course Orientation and Data Analysis with Python
2	Data Visualization with Python
3	Data Visualization with Python
4	Data Visualization with Python Exercises
5	Data Preprocessing
6	Data Preprocessing Exercises with Machine Learning (Regression)
7	Machine Learning Algorithmes (Classification)
8	Classification Exercises
9	Machine Learning Algorithms (Clustering)
10	Clustering Exercises

11	Model Selection
12	Exercises
13	Integration Concept
14	Integration Exercises (Homework)
15	Students Presentation
16	Students Presentation

8. Textbooks (Required Books for this course)

9. Reference Books (optional books for further study)

(1) Machine Learning with Python Cookbook, ISBN: 9781491989388.
(2) Advanced Data Analytics using Python, ISBN: 978-1-4842-3449-5.
(3) Available resources on Internet

10. Course Goals (Attainment Targets)

(1)	Understanding and applying Visulization with Python
(2)	Able to preprocess data using Python
(3)	Gain in-depth familiarity with various Machine Learning algorithms (supervised learning algorithms and unsupervised learning algorithms)
(4)	Able to Implement machine learning algorithms to real-world problems, and rigorously evaluate their performance using different methods.
(5)	Understand the concpet of Integration Programming and implement in real-world problems.

11. Correspondence relationship between Educational goals and Course goals

Educational goals of the school			Course Goals
High level ICT skills	Basic academic skills		
	Specialized knowledge and literacy		(1) (2) (3) (4) (5)
Human skill (Tankyu skill）	Ability to continually improve own strengths		(1) (2) (3) (4) (5)
	Ability to discover and resolve the problem in society	Problem setting	
		Hypothesis planning	
		Hypothesis testing	
		Practice	
	Fundamental Competencies for Working Persons	Ability to step forward	(3) (4) (5)
		Ability to think through	(3) (4) (5)
		Ability to work in a team	(3) (4) (5)
Professional ethics			

12. Evaluation

Goals	Evaluation method & point allocation					
	Examination	Quiz	Reports	Presentation	Deliverables	Other
(1)		O		O	O	
(2)		O		O	O	
(3)				O	O	
(4)		O		O		
(5)					O	
(6)						
(7)						
(8)						
Allocation		20		30	50	

13. Evaluation Criteria

Examination	
Quiz	Every week multiple choice quizzes are used to evaluate the understanding of students and motivate them for further learning.
Reports	
Presentation	In the final presentation, students will be asked to work on a real-world problem using Machine Learning algorithm. The evaluation will be based on participation in the group presentation, the structure of presentation, the relavance of the argument, time managment and relation to the course contents.
Deliverables	Individual and group assignment will be assigned for the students with focus on learning goals (1,3,4, and 5). The evaluation will be based on how the students understand the exercises and participation.
Other	

14. Active Learning

Hourly percentage of active learning within the whole class time	80%
1 Active learning such as problem solving assignment using the knowledge and skills acquired in class.	All the time
2 Active learning such as group works and discussions.	All the time
3 Outcome presentations and feedbacks.	Sometimes
4 Students actively make decisions on how the class should be conducted.	Not at all

15. Notes

Please bring your computers in the class.

16. Course plan	
(Notice) This plan is tentative and might be changed at the time of delivery	
Lessen 1: (Course Orientation)	(90 min)
1. Course Orientation and Introduction to Data Science Exercises 2. Data Analysis with Python Part III	
Lessen 2: (Data Visualization with Python)	(90 min)
1. Data Analysis Exercise 2. Data Visualization with Python	
Lessen 3: (Data Visualization with Python)	(30 min Lecture (60 min Exercises)

1. Data Visualization with Python
   a. Bar charts
   b. Histogram
   c. Regression	

Lessen 4: (Data Visualization with Python)	(30 min Lecture (60 min Exercises)

1. Data Visualization with Python
   a. Word Cloud
   b. Waffle Chart
   c. Exercise	

Lessen 5: (Data Preprocessing)	(90 min)
Data Preprocessing with Python 1. Importing Libraries 2. Uploading Data set 3. Dealing with missing data 4. Encoding categorical data 5. Splitting dataset (Traing & Testing) 6. Feature Scalling	

Lessen 6: (Data Preprocessing Exercises)	(90 min)
1. Machine Learning Algorithmes (Regression) 2. Exercises	
Lessen 7: (Machine Learning Algorithmes)	(90 min)
Classification	
Lessen 8: (Classification Exercises)	(90 min)
Exercises	
Lessen 9: (Machine Learning Algorithmes)	(90 min)
Clustering	
Lessen 10: (Clustering Exercises)	(90 min)
Exercises	
Lessen 11: (Model Selection)	(90 min)
1. Model Selection Concepts 2. Model Selection Exercises	
Lessen 12: (Exercises)	(90 min)
Exercises	
Lessen 13: (Integration Exercises)	(90 min)
1. Integration Concept 2. Integration Exercise	
Lessen 14: Integration Exercises (Homework)	(90 min)
Integration Exercises	
Lessen 15: (Student Presentation)	(90 min)
Students Presentation	
Lessen : Students Presentation	(90 min)
Students Presentation	