1. Course Code

2261

2. Course Title

G51e: Data Science

3. Teacher

HAMIDULLAH, Sokout

4. Term

Spring 1

5. Course Requirements (Courses / Knowledge for this course) and Important Information

Basic Knowledge of Statistics with Excel and Computer Programming with Python(But NOT Essential)

6. Course Overview and Objectives

Data Science has a big impact on the business landscape, which is constantly changing in today's world. As the quantity of human connection with digital footprint increases daily, an unimaginable large mass of fine-grained data is generated on a regular basis. Global internet users have climbed to 4.95 billion at the start of 2022, with internet penetration now standing at 62.5 percent of the world’s total population. The data generated behind could be useful to analyze existing issues, reveal previously hidden opportunities and predict future insights. Companies will compete for hundreds of thousands of new workers needed to navigate the digital world. Already, experts estimate that millions of jobs in DATA SCIENCE might remain vacant for the lack of readily available talent.

This course is about the world of data science. In this course we will start with an introduction where we will discuss the world of data science. I explain why data science is important and how it helps businesses stay competitive in this day and age. Then continues by introducing important mathematical and statistical concepts that are essential to do data science. Once we have laid out a solid foundation, we will be ready to learn Data Analysis for Data Science. Starting from the very basics, we will build up your skills and soon you will be able to play with Exploratory Data Analysis with Python. As well as the participants will learn visualization with Python.

7. Course Outline

1. Course orientation and Introduction to Data Science + Case study
2. Data Science Tools and Methodology + Fundamentals of Data and Data Science
3. Descriptive Statistics
4. Exercises
5. Inferential Statistics
6. Exercises
7. Hypothesis testing
8. P-value
9. Data Analysis with Python
10. Data Analysis with Python Exercises
11. Exploratory Data Analysis with Python
12. Exploratory Data Analysis with Python Exercises
13. Data Visualization with Python
14. Data Visualization with Python Exercises (Basics and Interactive)
15. Final Presentation
16. Final Presentation

8. Textbooks (Required Books for this course)

9. Reference Books (optional books for further study)

(1) Learning from Data: An Introduction to statistical reasoning, third edition. International Standard Book Number-13: 978-0-8058-4921-9.
(2) Data Science from scratch, ISBN-978-1-4919-0142-7.
(3) Available resources on Internet

10. Course Goals (Attainment Targets)

(1) Become familiar with foundations of Data and Data Science.
(2) Understanding the basic of statistics for Data Science.
(3) Be able to summarize a data set using descriptive statistics and Inferential statistics.
(4) Understand the Data Science Methodology (i) from problme to appraoch, (ii) working with the data, (iii) Deriving the answer.
(5) Be able to apply Data Analysis and Exploratory Data Analysis with Python
(6) Be able to define data-intensive problems in data science and understand their underlying statistical and computational principles.
(7) Understanding and applying Visulization with Python
(8)

11. Correspondence relationship between Educational goals and Course goals

Educational goals of the school			Course Goals
High level ICT skills	Basic academic skills		(1) (2) (3) (4)
	Specialized knowledge and literacy		(5)(6)(7)
Human skill (Tankyu skill)	Ability to continually improve own strengths		(5)(6)(7)
	Ability to discover and resolve the problem in society	Problem setting	(6)(7)
		Hypothesis planning	(2) (3) (6)
		Hypothesis testing	(2) (3) (6)
		Practice	(2) (3) (4) (5)(6) (7)
	Fundamental Competencies for Working Persons	Ability to step forward	(2) (3) (6)
		Ability to think through	(2) (3) (6)
		Ability to work in a team	(2) (3) (4) (5)(6) (7)
Professional ethics			(5)(6)(7)

12. Evaluation

Goals	Evaluation method & point allocation					
	Examination	Quiz	Reports	Presentation	Deliverables	Other
(1)		○		○		○
(2)		○			○	○
(3)				○	○	○
(4)		○				○
(5)		○		○	○	○
(6)		○		○	○	○
(7)		○		○	○	○
(8)						
Allocation		20		30	40	10

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 choose and summarize a scientific research journal paper or thier own ideas related to the course main objectives and contents. 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	Individuale and group assignment will be assigned for the students with focus on learning goals (2,3, and 5). The evaluation will be based on how the students understand the exercises and participation.
Other	Class Exercises Participation and Team Collaboration

14. Active Learning

Hourly percentage of active learning within the whole class time		70%
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.	Sometimes
3	Outcome presentations and feedbacks.	Sometimes
4	Students actively make decisions on how the class should be conducted.	Sometimes

15. Notes

This course is like a jouney by attending it you will benefit and might enjoy. Please bring your own computer in the class.

16. Course plan

(Notice) This plan is tentative and might be changed at the time of delivery

Lessen 1: (Course Orientation and Introduction to Data Science) **Lecture**

1. Introduction, Evaluation, Scope of the Course.
2. Data Science concepts and case studies.
3. Imerging Technologies shaping the modern data

Lessen 2:Fundametals of Data Science + Data Science Tools and Methodolog **Lecture**

1. Data
   a. Traditional
   b. Big
2. Data Science
   a. Business Intelligence
   b. Traditional Methods
   c. Machine Learning
3. Data Science Methodology

Lessen 3: Statisitcs for Data Science **Lecture + Exercises**

1. Population vs Sample
2. Types of Data

Lessen 4: Descriptive Statistics Exercises **Exercises**

1. Measure of central tendency
   a. Mean
   b. Median
   c. Mode

2. Measure dispersion
   a. Variance, Standard deviation
   b. Covariance, Correlation
   c. Range

Lessen 5: Inferential Statistics **Lecture + Exercises**

1. Distribution
2. Confidence Interval
3. Z-Statistics
4. T-Statistics

Lessen 6: Inferential Statistics Exercises **Exercises**

1. Inferential Statistics Exercises

Lessen 7: Hypothesis Testing **Lecture + Exercises**

1. Hypothesis Testing
   a. Null Hypoythesis
   b. Alternative Hypothesis
   c. Error in Hypothesis

Lessen 8: P-Value + Exercises **Exercises**

1. Concept of P-value
2. Exercises with P-value

Lessen 9: Data Analysis with Python **Lecture + Exercises**

1. Data Analysis with Python
   a. Python Key Packages for Data Analysis
   b. Basics of Data Analysis
   c. Pre-processing of Data with Python

Lessen 10: Data Analysis with Python Exercises **Exercises**

1. Data Analysis with Python Exercises
   a. Explore the dataset
   b. Deal with Missing data
   c. Formating data
   d. Data Standardization
   e. Data Normalization
   f. Data Bining
   g. Labeling data

Lessen 11: Exploratory Data Analysis with Python **Lecture + Exercises**

1. Exploratory Data Analysis with Python
   a. Descriptive Statistics with Python
   b. GroupBy
   c. ANOVA Test
   d. Correlation

Lessen 12: Exploratory Data Analysis with Python Exercises **Exercises**

1. Exploratory Data Analysis with Python Exercises

Lessen 13: Data Visualization with Pyton **Lecture + Exercises**

1. Data Visualization Exercises with Pyton

Lessen 14: Data Visualization Exercises with Pyton **Exercises**

1. Bascis of Visualizations
2. Interactive way of Visualization with Python

Lessen 15: (Presentation by Students) **Presentation (90 min)**

Group Presentation by students

Lessen 16: (Presentation by Students) **Presentation (90 min)**

Group Presentation by students