科目ナンバリング	U-LAS11 20006 SE55						
授業科目名 <英訳>	Data Analysis Practice II-E2 Data Analysis Practice II-E2			担当者所属 職名・氏名	医学研究科　准教授　PATAKY, Todd		
群	自然科学科目群			分野(分類)	データ科学(発展)	使用言語	英語
旧群	B群	単位数	2単位	週コマ数	1コマ	授業形態	演習（対面授業科目）
開講年度・開講期	2026・後期	曜時限	金3	配当学年	全回生	対象学生	全学向

[授業の概要・目的]

This course aims to provide students with practice using a variety of data analysis techniques, in a manner that emphasizes data analysis reporting. These data analysis techniques include: descriptive statistics, classical hypothesis testing and machine learning. We will focus on how to organize and present analysis results from relatively complex dataset. No prior knowledge of statistics or data science is required. Computer programming experience is useful but not required.

[到達目標]

This course provides an understanding of data analysis methods and how to organize and report data analysis results. Students will learn the basics of data science, statistics and computer programming. Students will learn how to organize and report data analysis results in a concise, information-dense manner. Students will learn how to use the Python programming language (python.org), Jupyter Notebooks (jupyter.org) and Markdown (markdownguide.org) and GitHub (github.com). The semester-long goal of this course is to produce a Final Project, which involves (1) analysis of a real-world dataset using several analysis techniques, and (2) creation of a full report of your findings, in a user-friendly format, similar to real-world report that you might one day produce for a data analysis customer.

[授業計画と内容]

The following weekly topics will be covered:

1) Jupyter I: Introduction
2) Python I: Basics
3) Python II: Visualizing Data
4) Jupyter II: GitHub Integration
5) Python III: Getting Data
6) Classical Stats I: Descriptive Statistics & Correlation
7) Jupyter III: Organizing Code
8) Classical Stats II: Hypothesis Testing
9) Classical Stats III: Meaning of Probabilities
10) Machine Learning I: Classification
11) Machine Learning II: Regression
12) Jupyter IV: Organizing Reports
13) Machine Learning III: Preprocessing & Dimensionality Reduction
14) Machine Learning IV: Clustering
15) Feedback

Total: 14 lectures + 1 feedback week

Data Analysis Practice II-E2(2)へ続く↓↓↓

Data Analysis Practice II-E2(2)

[履修要件]

There are no specific requirements for this class. However, students must be willing to work with open-source software, which is relatively poorly documented compared to commercial software. The class instructor will help with problems, but students are also encouraged to find solutions to their problems through internet searches.

Additionally, skills in the following would be helpful:
- Computer programming: Python experience (or experience with any other language)
- HTML editing: Markdown (or any other high-level HTML-generation language)
- Statistics: basic hypothesis testing, basic machine learning, etc.

[成績評価の方法・観点]

Students are expected to produce all in-class demonstrations independently, and to independently complete regular assignments.

Evaluation will be based on the following criteria:

- Assignments (80%) [10 @ 8% each]
- Final Project (20%)

TOTAL: 100%

Note that several of the assignments pertain directly to the Final Project. The Final Project will consist of a cumulation of work done throughout the semester.

[教科書]

An open, electronic textbook will be electronically distributed to students and will be used in all classes. All other necessary materials will also be distributed electronically and will be discussed in class.

[参考書等]

（参考書）

Joel Grus 『Data Science from Scratch: First Principles with Python』（O'Reilly Media）ISBN:978-1491901427（Lectures will loosely follow this textbook's content. This textbook is OPTIONAL, but will be useful for reviewing concepts and for independent study.）
"Data Science from Scratch" is a useful reference book, but is not required for this class. Lecture notes and all other materials will be made available electronically.

（関連ＵＲＬ）

https://github.com/joelgrus/data-science-from-scratch(Software (data and code) for “Data Science from Scratch” by Joel Grus)
https://www.jupyter.org(Jupyter will be used extensively for both lectures and assignments.)

[授業外学修（予習・復習）等]

This course has a variety of out-of-class assignments (including a Final Project) and no exam. Students who do not pay attention to the lecture content during class will likely have difficulties completing the assignments.

The lecture content will be made available prior to the lecture. It is recommended that students review this content prior to the lecture.

Data Analysis Practice II-E2(3)へ続く↓↓↓

Data Analysis Practice II-E2(3)

[その他（オフィスアワー等）]

OFFICE HOURS:
Immediately before / after class or by appointment (pataky.todd.2m @ kyoto-u.ac.jp)

[主要授業科目（学部・学科名）]