Course Syllabus: Foundation of Data Science

The class, Foundations of Data Science, is designed to be a freshman level data science class that focuses on the fundamentals of data science with some primary introductions of basic machine learning algorithms near the end of the class. Instead of focusing on the theory of machine learning and data analysis, we will get started with data analysis directly. The course content is primarily based on the undergraduate course, The Foundations of Data Science, from UC-Berkeley and the Data Science class from UC-San Diego.

1. Critical information, at a glance

- Textbook for our class will be the freely available awesome textbook, Computational and Inferential Thinking- The Foundations of Data Science", by Ani Adhikari and John DeNero.
- A reference textbook very useful for AI is Artificial Intelligence: A Modern Approach by S. Russell and P. Norvig.
- There will be a reading assignment for most of the days. It is expected that you complete the reading assignment before the start of the lecture.
- You should score at least 55% in the final exam to get a passing grade for this class, regardless of your overall percentage.
- All computing labs are done in pairs. We have 10 labs in total.
- There is an optional final project that will be treated as extra credit.

2. Pre-requisites

The pre-requisite of this class is basic high school algebra and an inquisitive mind. There is no requirement on prior programming experience. Each student is expected to have a computer. Either Windows or Mac is fine.

3. What you will learn from this class

This class will teach you how to explore data in a scientific way and show you the importance of data analysis. It will also teach you skills for programming data analysis code in Python. The topics included in the schedule adopts a breadth-first approach to give you a big picture of data science. Specifically, at the end of this course you will be able to:

- Understand the basics of Python programming
- Understand important statistics concepts such as sampling, hypothesis testing, and confidence intervals.
- Understand experimental design to gather data
- Use appropriate classification and inference tools to analyze data.

4. Grading

Your final grade will be determined via the following percentages:
Lecture participation points: 5%
Labs: 60%
Final: 35%

Important grading policies:

- You must score at least 55% on the final exam to pass the course. If you score lower than 55% on the final, you will receive an F for the course, regardless of your overall average.
- Every student should follow the policy on pair programming.
- According to Fudan University's policy, there is a threshold on the percentage of students who may receive A or A- in a class. Please keep this policy in mind.

5. Attendance Policy

Students are expected to participate in all lectures and missing more than 3 days of lectures will result in an F for the course regardless of lab or final exam scores.