MG-GY 8411 (MoT)

Data Engineering

Brooklyn Campus, Fall 2022

Course Description

This course complements *Statistics for Business Analytics* in coordination with *Programming for Business Intelligence* (BI). By focusing on case studies, students will gain the experience needed for electives in several knowledge areas pertinent to Data Science.

Course main topics:

1. By learning how to clean and structure raw data, students can prepare an environment to plugin AI/ML algorithms.
2. By learning how to reduce data dimensionality, students can build the transition from *Big Data* to AI/ML-based modeling techniques.
3. By learning how to split and organize data, students can engineer the information to deliver high-quality AI/ML-based models.

Pre-Requisites

Introductory Probability and Statistics, Linear Algebra, and **Python** (for example, if you already took the course *Programming for BI*).

Instructors

Carlos J. De Oliveira, cnd279@NYU.edu

Teaching Assistant: Wenqi Zhu, wz2366@nyu.edu

Office hours: Tuesday or Wednesday, 6.00 pm – 7:00 pm (by appointment)

Textbook

To be announced

Software

Python on Anaconda

Grading Policy

- **Weekly Assignments - 50%**

  – Graded

  Homework is based on Python coding. Grading:

  - Code OK → A (for **top 3** homework) or A-
  - Code not OK → B+, B, B-, or C
  - No Code (i.e., no homework) → I

- **Final presentation – 50%**

  – Graded

  It is a team or individual presentation. Grading:

  - Best team (or top 3 individual presentations): A
  - Second best team: A-
  - And so on.

Statement of Academic Integrity

Students are expected to follow standards of excellence set forth by New York University. Such standards include respect, honesty, and responsibility. This class does not tolerate violations of academic integrity, including:

- Plagiarism
- Cheating
- Submitting your work toward requirements in more than one course without prior approval from the instructor
- Collaborating with other students for work expected to be completed individually
- Giving your work to another student to submit as their own
- Purchasing or using papers or work online or from a commercial firm and presenting it as your work

Fall 2022 Course Schedule

To be announced