Summary
The objective of this course is to introduce doctoral students to computational methods for data-driven empirical research in management.

Content
The course complements courses in statistics and econometrics with a programmatic understanding of how to acquire, store, manipulate, measure, plot, analyze, and classify data. The course requires students to program in Python. The basics of Python language will be reviewed during the first session, but students who are unfamiliar with Python should review the Python 3 tutorials at: https://wiki.python.org/moin/BeginnersGuide/NonProgrammers and complete the 4-hour tutorial by DataCamp at: https://www.datacamp.com/courses/intro-to-python-for-data-science prior to the start of class. You may also want to review the official Python 3 documentation at: https://docs.python.org/3/

By the end of the course, students will understand how to work with data in python to conduct doctoral research in economics and management. The ultimate learning objective of the course is to build a toolkit that will elevate the empirical quality of each student’s dissertation.

Keywords
Data Processing, Visualization, Cloud Computing, Data Analysis, Text Analysis, Simulation, Machine Learning.

Assessment methods
Students will be evaluated based on five, take-home assignments – each worth 20% of the overall course grade. Each assignment should take 5 to 6 hours to complete. Each assignment is due before the start of the next class; the final assignment is due one week after the final class.

Create a new jupyter notebook for each assignment and commit it to your git repository. Name your jupyter notebooks sequentially as:

a1.ipynb  a2.ipynb  a3.ipynb  a4.ipynb  a5.ipynb

Resources
Notes/Handbook
Please contact the TAs of the course for more information or a detailed syllabus:
maximilian.hofer@epfl.ch
george.abiyounes@epfl.ch