NATIONAL UNIVERSITY OF SINGAPORE
NUS Business School
Department of Analytics & Operations

DAO2702/DSC2008 Programming for Business Analytics

Session: Semester 2, 2020/2021

Instructor: Xiong Peng bizxio@nus.edu.sg

Description:

This module is an introductory course to business analytics and data science. It covers basic Python programming and preliminary statistics, with a great emphasis on addressing practical business problems and real datasets. Data science is an interdisciplinary field that requires business insights and expertise, proficiency in programming, as well as a strong background in mathematics and statistics. Therefore, lectures and tutorials in this semester would focus on trainings in the following perspectives:

- Python programming and Pythonic coding styles
- Analytical and visualization packages
- Math and statistics
- Practical business insights and problem solving skills

Syllabus:

1. Basics of Python programming
    1. Data structures and flow control
    2. Functions and packages
2. Data analysis with Python
    1. Analytical tools: NumPy, SciPy, Pandas
    2. Data visualization: Matplotlib
    3. Data collection and cleaning
3. Statistical inference
    1. Sampling and inference
    2. Confidence intervals
    3. Hypothesis testing
4. Linear regression
    1. Model assumptions and interpretations
    2. Categorical variables and modelling nonlinearity
    3. Package Statsmodels for regression analysis

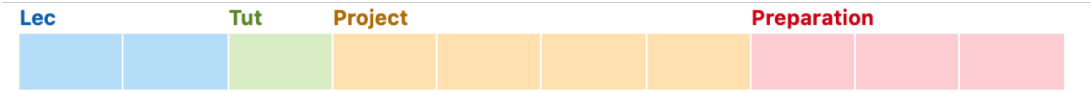
Software:

Anaconda: installation

Class Materials:

- Offline lecture video uploaded in LumiNUS multimedia channels
- Jupyter Notebook files as lecture notes
- Jupyter Notebook files as tutorial case studies
- Jupyter Notebook files as exercises
- Slides as supplementary
- The folder "Advanced topics" provides supplementary reading materials. They are note tested but may be helpful for your project.

Workload:
- Students are required to review lectures covered in each week
- Students are supposed to work on exercises for each week. The solutions will be uploaded in the next week, and there is no need to submit the assignment.
- All tutorial case studies are related to topics covered in previous lectures.



Reference Books:
Python programming:
- Python for data analysis, by Wes Mckinney
- The hitchhiker's guide to Python, by Kenneth Reitz and Tanya Schlusser
- Python data science handbook, by Jake VanderPlas

Statistics:
- Introductory econometrics, by Jeffrey M. Wooldridge
- An introduction to statistical learning, by Trevor Hastie et al.
- Storytelling with data, by Cole Nussbaumer Knaflic

Assessments:
Continuous Assessment:
Class Participation 10%
Group Project 20% for report and 15% for presentation
- Team work
- You may choose your own teammates (in the same tutorial session) by filling a survey. **All team members must fill the survey before the deadline, otherwise we will randomly assign you to a team.**
- A report (no more than 8 pages/4 pieces of double sided paper) and a formal presentation (10 to 15 minutes)
- Your grade of the project would also be affect by the peer evaluation of your teammates. **Zero mark for zero contribution!**
- **Zero mark for plagiarism.**

Final Examination: 55%

Schedule:

Week 1.
Course Overview
Introduction to Programming and Business Analytics

Week 2.
Variables, data types, and control flow I

Week 3.
Control flow II and strings

Week 4.
Built-in compound data types

Week 5.
Functions, modules, and packages

Week 6.
Data arrays and data visualization

Week 7.
Basics of Pandas

Week 8.
Facts from data

Week 9.
Confidence intervals and hypothesis testing

Week 10.
Introduction to regression analysis

Week 11.
Regression analysis for explanatory modeling

Week 12.
Nonlinearity and categorical variables

Week 13.
Review