Hong Kong Baptist University

School of Business

Programme:	**Master of Science in Data Analytics and Business Economics Programme**		
Department:	Economics		
Course Code:	ECON 7890	Level:	Taught Postgraduate
Course Title:	Foundations in Big Data Analytics: Programming		
Prerequisites:	Nil.	Medium of Instruction	English
Duration:	39 hours	Units:	3 (3, 3, 0)
Course Description: (including Aims & Objectives)	The vast majority of data science roles are Python-based. This course aims at equipping students with Python programming techniques necessary to manipulate the data, perform feature selection and model optimization, analyze data using machine learning, and evaluate the outputs.		
Texts & References: *(* recommended textbook(s))*	1. *VanderPlas J. (2017) *Python Data Science Handbook: Essential Tools for Working with Data*. O'Reilly Media Inc. 2. McKinney W. (2017). *Python for Data Analysis*. 2nd Edition. O' Reilly Media. 3. Wickham H. and G. Grolemund. (2017). *R for Data Science: Import, Tidy, Transform, and Model Data.* 1st Edition. O' Reilly Media. 4. Wickham H. (2016). *ggplot2: Element Graphics for Data Analysis (Use R!)*. 2nd Edition. Springer. 5. Provost F. and T. Fawcett (2013). *Data Science for Business*. 1st Edition. O' Reilly Media.		
Learning Outcomes:	Upon completion of this course, students should be able to: 1. Demonstrate a detailed understanding of state-of-the-art programming platform for both data storage and data analysis. 2. Identify and illustrate the challenges of programming for big data. 3. Implement solutions to various big data programming problems using a range of data-analytical tools, and evaluate model performance. 4. Present an informed view of the data analytical results.		

Teaching & Learning Activities:

	Learning Outcome Addressed:						Learning Outcome Addressed:				
	1	2	3	4	5		1	2	3	4	5
Lecture	✓	✓	✓	✓		Services Learning					
Guest speakers						Internship					
Case Study						Field study					
Role playing						Company visits					
Student presentation	✓	✓	✓	✓		e-learning					
Project	✓	✓	✓	✓		Independent study					
Simulation game						Others					
Exercises and problems	✓	✓	✓	✓							

Major Assessment Methods: For each Major Assessment Method below, please indicate the specific pedagogical /assessment methods involved (by putting a ✓ in the relevant box(es) on the right-hand side).	Case Study	Role Playing	Student Presentation	Individual project/paper	Group project/paper	Simulation Game	Exercises & problems	Service learning	Internship	Field Study	Company visits	Written examination	Oral examination	Learning Outcomes Addressed
Class Participation/ Discussion (20%)			✓											1 to 4
Assignment(s) (80%)			✓		✓		✓							1 to 4
Test(s) (0%)														
Examination (0%)														
Others *(please specify)* _____ (%)														

Course Web: Course templates are available at BU eLearning (formerly called "BU Moodle"), programme website and Staff Area in School website *(for staff only)*.

Course Content:

	Hours	Learning Outcome no.
I. Introduction of IPython A. Installation B. Basic commands C. Python Modules D.	3	1,2
II. Introduction to Numpy A. Multidimensional array B. Indexing and slicing C. Sorting data D. Basic statistics E.	6	1,3,4
III. Introduction to Pandas A. Data frames and series B. Missing data C. Data aggregation and grouping D. Combining datasets E. F. Application: Breast Cancer	6	1,3,4
IV. Introduction to Matplotlib A. Simple linear plots and scatter plots B. Visualizing errors C. Histogram, binning and density D. Customizing figures E. Mulitple plots F. Seaborn module	6	1,3,4
V. Machine Learning I A. Introduction to scikit-learn B. Hyperparameters and model validation C. Learning curves D. Feature engineering	6	1,3,4

	E. s		
	VI. Machine Learning II A. Linear regression B. Decision trees C. k-Means clustering	6	1,3,4
	VII. Machine Learning III A. Support vector machine B. Naive Bayes classification C. Principal component analysis	6	1,3,4
	Total	39 hrs.	

Contribution to the Mission of the School:	☑ to cultivate academic curiosity, integrity and leadership potential ☑ to enhance all-rounded training ☐ to develop consciousness of values and social responsibility ☑ to disseminate contemporary knowledge ☑ to foster a global vision ☐ to disseminate the research findings of faculty members in the School ☐ to develop awareness in public policy
Contribution to the Learning Goals of the Programme:	The learning goals of the MSc in Data Analytics and Business Economics Programme: ☐ Identify the challenges of digital economy and apply economic principles to think strategically about business decisions. ☐ Define business problem clearly and identify appropriate analytical tools to address the issues. ☑ Demonstrate ability in choosing appropriate algorithms and implementing programming languages for business analysis. ☑ Interpret the data outcomes and deliver the crucial findings and insights effectively through data visualization for business analysis. ☑ Formulate solutions to real-world problems with data analytics and/ or economic principles.
Course Co-ordinator:	Dr. Shui Ki WAN