The Hong Kong Polytechnic University
Subject Description Form

Subject Code	AMA503
Subject Title	Statistics and Data Analytics
Credit Value	3
Level	5
Pre-requisite / Co-requisite/ Exclusion	Nil
Objectives	This subject introduces students fundamental theories and methods of statistics and data analysis that are useful for improving processes in semiconductor manufacturing.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: (a) understand the principle of data analytics and data mining; (b) apply various data analytics algorithms and techniques to analyze data related to semiconductor manufacturing; (c) carry out in-depth data analysis for the optimization of semiconductor processing and device fabrication.
Subject Synopsis/ Indicative Syllabus	**Overview of statistics and data analytics**: Introduction to the use of data mining and analysis in semiconductor industries. **Analytics algorithms and techniques**: Data analysis methods include principal component analysis, independent component analysis, hypothesis testing, linear and logistic regression, autoregressive–moving-average, autoregressive integrated moving average and analysis of variance. Understanding of other computational intelligence and advanced analytics techniques. **Programming languages and tools for data analytics**: Python, R programming, Excel, SAS, etc. **Examples of using analytics technique in semiconductor industries**: Data provided by the semiconductor industries for data analysis and optimization purpose.
Teaching/Learning Methodology	This subject is mainly delivered through lectures and programming training. The teaching and learning approach is mainly problem-solving oriented. The approach aims at the development of data analytics techniques and how the techniques can be applied to problem solving. Students are encouraged to adopt a deep study approach by employing high level cognitive strategies, such as critical and evaluative thinking, relating, integrating and applying theories to practice.

Assessment Methods in Alignment with Intended Learning Outcomes	Specific assessment methods/tasks	% weighting	Intended subject learning outcomes to be assessed (Please tick as appropriate)		
			a	b	c
	1. Continuous Assignments	25%	✓	✓	✓
	2. Project	30%			✓
	3. Examination	45%	✓	✓	✓

Last update: November 2022

	Total	100 %	

Continuous Assignments: includes participation in the computer laboratory sessions and turning on time the graded quizzes to check students' progress throughout the semester.

Project: will be assigned during the second half of the semester. Students will use their knowledge gained in the class/lab to tackle problems related to data analytics in semiconductor manufacturing.

Examination: assess the knowledge acquired by the students, as well as to determine the extent to which they have achieved the intended learning outcomes.

Student Study Effort Expected	Class contact:	
	▪ Lecture	26 Hrs.
	▪ Computer Laboratory	13 Hrs.
	Other student study efforts:	
	▪ Assignment/Projects	51 Hrs.
	▪ Reading and self-study	30 Hrs.
	The total student study effort	120 Hrs.
Reading List and References	N.R. Pal and L. Jain (Eds), Advanced Techniques in Data Mining and Knowledge Discovery, Springer, 2005 (Chapter 2, Advanced Methods for the Analysis of Semiconductor Manufacturing Process Data).	
	R. Peck, C. Olsen and J. Devore, Introduction to Statistics and Data Analysis, 3rd Ed, Thomson Higher Education, 2008.	
	S. Bagchi *etal*, Data analytics and stochastic modeling in a semiconductor fab, Appl. Stochastic Models Bus. Ind., 26:1–27, 2010.	
	G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning – with Applications in R, 2nd Ed, Springer 2021 (https://web.stanford.edu/~hastie/ISLR2/ISLRv2_website.pdf).	