UNIVERSITI MALAYA

Institute of Mathematical Sciences

PROGRAMME HANDBOOK

Master of Science in Statistics by Coursework

Session 2022/2023

https://ism.um.edu.my/postgraduate-info

Master of Science in Statistics Session 2022/2023 (42 CREDITS)		
1. Programme Core Courses (27 CREDITS)		
Course Code	**Course Name**	**Credits**
SQB7022	Research Methodology for Statistics	3
SQB7023	Research Project in Statistics	10
SQB7024	Statistical Inference	4
SQB7025	Probability Theory	4
SQB7026	Programming in Statistics	3
SQB7027	Statistical Consultancy and Data Analysis	3
2. Programme Elective Courses (15 CREDITS)		
Course Code	**Course Name**	**Credits**
SQB7028	Multivariate Analysis	3
SQB7029	Stochastic Models	3
SQB7030	Bayesian Statistics	3
SQB7031	Generalized Linear Models	3
SQB7032	Experimental Design and Quality Engineering	3
SQB7033	Statistical Time Series	3
SQB7034	Computer Intensive Methods	3
SQB7035	Robust Statistics	3
SQB7036	Data Mining	3
SQB7037	Survival Data Analysis	3

PROGRAMME GOAL

To produce master graduates who have strong statistical knowledge, are capable of analysing and solving problems, and can think critically. Our graduates are capable of adapting to diverse environments and contributing meaningfully towards professions in different fields.

PROGRAMME LEARNING OUTCOMES

At the end of the programme, graduates of the Master of Science in Statistics are expected to demonstrate the following qualities:

1. Comprehend advanced statistical theories with regards to the statistical and mathematical arguments, proofs and abstract concepts.
2. Apply technical and practical skills related to advanced statistics in different fields.
3. Perform professional work with good social skills, and behaving responsibly to the community.
4. Practise professionally and ethically in one's profession.
5. Work collaboratively, and articulate statistical ideas clearly, accurately, and effectively.
6. Solve statistical problems using logical, analytical, and critical thinking.
7. Habitually use statistical thinking to understand and manage information.

SAMPLE PLANNER FOR COURSES (GRADUATION IN TWO SEMESTERS)

COMPONENT	YEAR 1				TOTAL CREDIT HOURS
	SEMESTER 1		SEMESTER 2		
	COURSES	CREDIT HOUR	COURSES	CREDIT HOUR	
Core Courses	SQB7022 Research Methodology for Statistics	3	SQB7024 Statistical Inference	4	**27**
	SQB7025 Probability Theory	4	SQB7027 Statistical Consultancy and Data Analysis	3	
	SQB7026 Statistical Laboratory	3	SQB7023 Research Project in Statistics	5	
	SQB7023 Research Project in Statistics	5			
Elective Courses	*SQB70***	6	*SQB70***	9	**15**
TOTAL CREDIT HOURS		**21**		**21**	**42**

Note: Most of our programme alumni take 3 to 4 semesters to finish their studies. Please consult the course coordinator about course planning.

INSTITUTE OF MATHEMATICAL SCIENCES

The Institute of Mathematical Sciences (ISM) was established as a department in the Faculty of Science when the University of Malaya was founded in Kuala Lumpur in 1959. It has grown into three branches: Pure Mathematics, Applied Mathematics and Statistics. The Institute offers 4 first degree programmes, one master's by coursework programme as well as master's by research and doctoral programmes.

Since the 2017/2018 academic session, ISM has been conducting the Master of Science in Statistics coursework programme. This programme provides opportunities to first degree graduates to acquire advanced statistical knowledge and skills.

STAFF

ISM has a group of experienced lecturers in teaching. They are active in research and have published their work in international as well as local journals. Our research activities encompass a broad spectrum, from findings and knowledge which are abstract in nature, to those with direct applications in the industry. ISM strives to establish and forge close relationships with the industry and other research institutions. This strengthens the quality of teaching and supervision of projects/theses for students at the Bachelor, Master and Doctoral levels.

HEAD:
Assoc. Prof. Dr. Zailan Siri, *BSc(UM), MSc(UPM), PhD(UKM)*

DEPUTY HEADS:
Profesor Dr. Wong Kok Bin, *BSc, MSc, PhD(UM)*
Assoc. Prof. Dr. Adriana Irawati Nur Ibrahim, *BSc(USM), MSc(UM), PhD(Bath)*
Assoc. Prof. Dr. Ng Kok Haur, *BSc, MSc(UPM), PhD(UM)*
Assoc. Prof. Dr. Noor Fadiya Mohd Noor, *BSc, MSc(UTM), PhD(UKM)*
Dr. Muhamad Hifzhudin Noor Aziz, *BSc(UM), MSc, PhD(Glasgow)*

MATHEMATICS UNIT

COORDINATOR (B.Sc. in MATHEMATICS):
Assoc. Prof. Dr. Chooi Wai Leong, *BSc, MSc, PhD(UM)*

PROFESSORS:
Dr. Angelina Chin Yan Mui, *BSc, MSc(UM), PhD(Queensland)*
Dr. Wan Ainun Mior Othman, *BSc(UNCC), MSc(N Carolina State), PhD(USM)*
Dr. Wong Kok Bin, *BSc, MSc, PhD(UM)*

HONORARY PROFESSORS:
Dr. Kurunathan Ratnavelu, *BSc, MSc(UM), PhD(Flinders)*
Dr. Mohd Omar, *BSc(UM), MSc(Hull), PhD(Exeter)*

ASSOCIATE PROFESSORS:
Dr. Chooi Wai Leong, *BSc, MSc, PhD(UM)*
Dr. Deng Chai Ling, *BSc, MSc, PhD(UM)*
Dr. Zailan Siri, *BSc(UM), MSc(UPM), PhD(UKM)*

LECTURERS:
Dr. Amizah Malip, *BSc(IIUM), MSc, PhD(London)*
Dr. Elayaraja Aruchunan, *BSc, MSc(UMS), PhD(Curtin)*

Dr. Kohilavani Naganthran*, BSc(UNISEL), MSc, PhD(UKM)*
Dr. Kwa Kiam Heong, *BSc, MSc(UM), PhD(Ohio)*
Dr. Loo Tee How, *BSc, MSc, PhD(UM)*
Mr. Mohamad Bakri Zubir, *BSc(Bristol), MSc(Exeter)*
Dr. Mohd Zahurin Mohamed Kamali, *BSc, MSc, PhD(UM)*
Dr. Muhamad Hifzhudin Noor Aziz, *BSc(UM), MSc, PhD(Glasgow)*
Dr. Ong Siew Hui, *BSc, MSc, PhD(UM)*
Dr. Oon Shea Ming, *BSc, MSc, PhD(UHP)*
Dr. Ruhaila Md. Kasmani, *BSc(UKM), MSc(UTM), PhD(UM)*
Dr. Shahizat Amir, *BSc(UKM), MPhil, PhD(UM)*
Dr. Siti Suzlin Supadi, *BSc, MSc, PhD(UM)*
Dr. Tan Ta Sheng, *BA, MA, PhD(Cambridge)*

STATISTICS AND ACTUARIAL SCIENCE UNIT

COORDINATOR (B.Sc. in STATISTICS):
Dr. Nur Anisah Mohamed @ A. Rahman, *BSc, MSc(UM), PhD(Newcastle)*

COORDINATOR (B. ACTUARIAL SCIENCE):
Mdm. Nadiah Zabri, *BBA(Wisconsin-Madison), MSc(Kent)*, *AIA*

COORDINATOR (M.Sc. in STATISTICS):
Assoc. Prof. Dr. Khang Tsung Fei, *BSc, MSc(UM), PhD(NUS)*

PROFESSORS:
Dr. Ibrahim Mohamed, *BSc(Bristol), MSc(Reading), PhD(UiTM)*

HONORARY PROFESSORS:
Dr. **Nor Aishah Hamzah**, *BSc(Southampton), MSc(Leeds), PhD(Bristol), DipEd(UKM), MIS(UK)*
Dr. Ong Seng Huat, *BSc, MSc, PhD(UM)*

ASSOCIATE PROFESSORS:
Dr. Adriana Irawati Nur Ibrahim, *BSc(USM), MSc(UM), PhD(Bath)*
Dr. Khang Tsung Fei, *BSc, MSc(UM), PhD(NUS)*
Dr. Ng Kok Haur*, BSc, MSc(UPM), PhD(UM)*
Mr. Raveendran A/L VGK Menon, *BEng(UM), MActuarialSc(Georgia State), ASA, AIAA*
Dr. Rossita Mohamad Yunus, *BSc, MSc(UM), PhD(USQ)*

LECTURERS:
Dr. Dharini A/P Pathmanathan*, BSc, MSc, PhD(UM)*
Dr. Koh You Beng, *BSc(UMS), MSc(UM), PhD(HKU)*
Dr. Lim Sok Li, *BEd, MSc, PhD(USM)*
Dr. Mohd Azmi Haron, *BBA, MBA, PhD(UPM)*
Mdm. Nadiah Zabri, *BBA(Wisconsin-Madison), MSc(Kent)*, *AIA*
Dr. Ng Choung Min, *BSc(UTM), MSc, PhD(UM)*
Dr. Nur Anisah Mohamed @ A. Rahman, *BSc, MSc(UM), PhD(Newcastle)*
Dr. Shaiful Anuar Abu Bakar, *DipActuarialSc, BSc(UiTM), MSc(Heriot-Watt), PhD(Manchester)*

COORDINATOR (B.Sc. Ed. Mathematics):
Dr. Lim Sok Li, *BEd, MSc, PhD(USM)*

RESEARCH AREAS

Research areas at ISM include the following fields: differential geometry, group theory, ring theory, linear preserver problems, functional identities, linear and multilinear algebra, matrix theory, combinatorial graph theory, graph theory, social network analysis, supply chain management, operations research, numerical analysis, computational statistics, robust statistics, probability distribution theory, nonlinear time series, image processing, regression analysis, statistical quality control, mathematical biology, and data science.

COMPUTER FACILITIES

Currently, ISM has a computer lab equipped with 10 laptops, 17 workstations, 121 desktops, 3 laser printers, 1 colour printer, and 4 heavy-duty dot matrix printers, all of which are interconnected through a network system. The lab is also equipped with 4 LCD projectors, 2 visualizers, and 3 scanners. The lab utilises state-of-the-art software such as MATLAB (with various toolboxes), SPSS, Wolfram Mathematica, MathType, Minitab, Microsoft Visual C++, Dev-C++, and S-PLUS. In addition, three of the lecture halls and tutorial rooms are each equipped with an LCD projector and a visualiser.

JOB OPPORTUNITIES

The learning of mathematics increases one's skills in problem solving and analysis. It trains one's mind to process information, form sophisticated ideas, and discern complex arguments. The training to think quantitatively, logically, and analytically is invaluable to one's career.

Since the use of mathematics is all encompassing in human endeavours, a graduate's career opportunity is numerous and not confined to only teaching and research. Our graduates are well-represented in diverse sectors such as finance, business, industry, and information science.

SYNOPSIS OF COURSES

SQB7022 RESEARCH METHODOLOGY FOR STATISTICS

This course is designed to give knowledge and skills to students related to suitable methodologies in statistical research. It includes searching and critically evaluating journal articles on statistical problems. The course will introduce fundamental statistical concepts and techniques useful for research in statistics. The students will be guided in writing research proposal in the area of statistics.

Assessment
Continuous Assessment: 100%

References:

1. Bordens, K.S., Bruce, B. (2011). Research and design methods: a process approach. 8th Ed. McGraw-Hill.
2. Flick, U. (2011), Introducing research methodology: a beginner's guide to doing a research project. Sage Publication, London.
3. Dawson, C. (2019). Introduction to Research Methods: A Practical Guide for Anyone Undertaking a Research Project. 5th Ed. Robinson.
4. Boudah, D. J. (2019). Conducting Educational Research: Guide to Completing a Thesis, Dissertation, Or Action Research Project. Sage Publications.
5. Paul D.. Leedy, Ormrod, J. E., & Johnson, L. R. (2019). Practical research: Planning and design. Pearson Education.

SQB7023 RESEARCH PROJECT IN STATISTICS

Refer to lecturers concerned for project synopsis and reference texts.

Assessment
Continuous Assessment: 100%

References:
Refer to lecturers.

SQB7024 STATISTICAL INFERENCE

Estimating of parameters and fitting of probability distribution; EM algorithm; use of statistical software for parameter estimation; Common families of distribution: location and scale families, exponential family; Principles of data reduction: the sufficiency principle – sufficient statistics, minimal sufficient statistics, ancillary statistics, complete sufficient statistics; Hypothesis tests: likelihood ratio tests, generalized likelihood ratio tests, error probabilities and power function; uniformly most powerful (UMP) test; UMP unbiased test.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Hogg, R.V., Craig, A.T. and Mckean, J.W. (2019). Introduction to Mathematical Statistics. 8th Ed. Pearson.
2. Casella, G. & Berger, R.L. (2021). Statistical Inference. Cengage Learning.
3. Hogg, R. V., Tanis, E. A., & Zimmerman, D. L. (2020). Probability and statistical inference. 10th Ed. Pearson/Prentice Hall..
4. Devore, J.L., & Berk, K.N. (2012). Modern Mathematical Statistics with Applications. Cengage Learning.

SQB7025 PROBABILITY THEORY

Probability measure and space, sigma field. Lebesgue integration. Random variables, measurability, independence. Distribution functions. Inequalities, characteristic functions. Various modes of convergence of sequences of random variables. Classical limit theorems. Examples of applications.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Billingsley, P. (1995). Probability and Measure (3rd ed.). New York: John Wiley.
2. Durrett, R. (2019). Probability: Theory and Examples (5th ed.). Cambridge: Cambridge University Press.
3. Karr, A. F. (1993). Probability. New York: Springer-Verlag.
4. Rosenthal, J.S. (2006). A First Look at Rigourous Probability Theory (2nd ed.). Singapore: World Scientific Publishing Company.

SQB7026 PROGRAMMING IN STATISTICS

Use of functions and commands in statistical packages for exploratory data analysis, modelling and statistical inferences, coding programming language.

Assessment
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Crawley, M. (2019). *Statistics: An Introduction using R* (2nd ed.). Chichester, UK: John Wiley & Sons.
2. Matloff, N. (2011). *The Art of R Programming: A Tour of Statistical Software Design*. San Francisco, CA: No Starch Press.
3. Deitel, P.J., Dietal, H. (2019). *Introduction to Python for Computer Science and Data Science: Learning to Program with AI, Big Data and the Cloud*. UK: Pearson Education.
4. Raschka, S., Mirjalili, V. (2019). *Python Machine Learning*: *Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2* (3rd ed.). Birmingham, UK: Packt Publishing

.

SQB7027 STATISTICAL CONSULTANCY AND DATA ANALYSIS

Introduction to consultancy activities and consulting methods. Related problems and issues. Exposure to the use of primary and secondary data from various sources. Application of suitable statistical methods such as multivariate analysis, regression and time series in the analysis of real data. Producing report and presenting the findings that suit the needs of the clients.

Assessment
Continuous Assessment: 100%

References:
1. Boen, J.R. & Zahn, D.A. (1982). The Human Side of Statistical Consulting. Belmont, CA: Lifetime Learning Publications.
2. Hand, D.J. & Everitt, B.S. (1987). The Statistical Consultant in Action. Cambridge: Cambridge University Press.
3. Lander, J. P. (2017). R for everyone: advanced analytics and graphics. Pearson Education.
4. Wickham, H. & Grolemund, G. (2017). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. Sebastopol, CA: O'Reilly Media

SQB7028 MULTIVARIATE ANALYSIS

Graphical display of multivariate data. Dimensional reduction methods (factor analysis, principal components, cluster analysis, etc.). Discriminant and classification methods (discriminant analysis and cluster analysis). Multivariate linear regression, canonical correlation analysis. Selected topics in advanced multivariate methods (e.g. sparse multivariate methods, multidimensional scaling, and functional data analysis). Introduction to applied spatial statistics. Exploratory spatial data analysis. Selected topics in spatial data analysis (spatial regression, random field, etc.). Models and methods for spatiotemporal and multivariate spatial data. Application of multivariate analysis to real world problems.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Johnson R.A & Wichern, D.W (2014). Applied multivariate statistical analysis. 6th Ed., Pearson new international edition.
2. Härdle, W.K. & Simar, L.A. (2019). Applied multivariate statistical analysis. Springer.
3. Konishi, S. (2014). Introduction to multivariate analysis-linear and nonlinear modelling. Chapman & Hall.
4. Schabenberger, O. & Gotway, C.A. (2017). Statistical methods for spatial data analysis. CRC press.
5. De Smith, M. J., Goodchild, M. F., & Longley, P. (2018). Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Troubador publishing ltd.

SQB7029 STOCHASTIC MODELS

Introduction to Markov chains. Continuous-time Markov chains. Poisson processes. Backward and forward Kolmogorov equations. Birth and death processes with examples. Definition and concepts in renewal processes, distribution for the number of renewal, renewal function and theorems for renewal processes. Examples for various types and applications of renewal processes. Introduction to martingales. Examples of stochastic models for real world applications.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Cox, D. R. and Miller, H. D. (1965). The Theory of Stochastic Processes. Chapman & Hall.
2. Durret, R. (2016). Essentials of Stochastic Process, 3rd ed. Springer.
3. Cox, D. R. (1962). Renewal Theory, Methuen.
4. Jones, P. W., & Smith, P. (2017). Stochastic Processes: An Introduction. CRC Press.
5. Dobrow, R. P. (2016). Introduction to Stochastic Processes with R. John Wiley & Sons.

SQB7030 BAYESIAN STATISTICS

Bayesian inference. Prior formulation. Multi-parameter models. Implementation via posterior sampling. Convergence assessment. Hierarchical and mixture models. Model checking. Model evaluation and comparison. Bayesian decision theory. Examples with different models from different areas of real-world application such as Bayesian linear regression and Bayesian network.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Lee, P. M. (2012). Bayesian statistics: an introduction, 4th edition. Wiley.
2. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis. CRC press.
3. Cowles, M.K. (2013). Applied Bayesian Statistics with R and Open BUGS example, Springer.
4. Turkman, M. A. A., Paulino, C. D., & Müller, P. (2019). Computational Bayesian statistics: an introduction. Cambridge University Press.
5. McElreath, R. (2020). Statistical rethinking: A Bayesian course with examples in R and Stan. CRC press.

SQB7031 GENERALIZED LINEAR MODELS

Generalised Linear Models (GLM) based on exponential family. Linear predictor, link function, canonical link, likelihood, the iterative reweighted least squares algorithm, Fisher information, tests on individual parameters, deviance, residuals. Application of GLMs on real life data such as using models for continuous and discrete response, models for polytomous data, log-linear models, quasi-likelihoods, models for correlated responses, and generalised additive model. Statistical software use for data analysis. Model fitting, selection and diagnostics.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Dobson A.J. & Barnett A.G. Boca Raton (2008). An Introduction to Generalized Linear Models. 3rd Ed., Chapman & Hall/CRC.
2. McCullagh P. & Nelder J.A., Boca Raton (1989). Generalized Linear Models.2nd Ed., Chapman & Hall.
3. Myers R.H., Montgomery D.C., Vining G.G., Robinson T.J. Hoboken, New Jersey (2010). Generalized Linear Models: with Applications in Engineering and the Sciences. 2nd Ed., John Wiley & Sons.
4. Dunn P.K. & Smyth G.K. New York (2018). Generalized Linear Models With Examples in R. Springer.
5. Fox, J., & Weisberg, S. (2018). An R companion to applied regression. Sage publications.

SQB7032 EXPERIMENTAL DESIGN AND QUALITY ENGINEERING

Factorial designs. Blocking and confounding systems. Fractional factorial designs. Response surface methodology. Robust design. Gauge repeatability and reproducibility.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Montgomery, D.C. (2019). Design and analysis of experiments. 10th Ed., John Wiley.
2. Myers, R.H., Montgomery, D.C., Anderson-Cook, C.M. (2016). Response surface methodology: Process and product using designed, experiments. 4th Ed., John Willey.
3. Wu, C.F.J., Hanada, M.S. (2009). Experiments: Planning analysis and optimization. 2nd Ed., John Willey.
4. Dean, A., Morris, M., Stufken, J., & Bingham, D. (Eds.). (2020). Handbook of design and analysis of experiments (Vol. 7). CRC Press.
5. Fowlkes, W.Y. Creveling, C.M. (1995). Engineering methods for robust product design: Using Taguchi methods in technology and product development. Addison-Wesley.

SQB7033 STATISTICAL TIME SERIES

Introduction to time series and forecasting, time series regression, ARIMA models, state space models, dynamic regression models.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Box, G.E.P., Jenkins, G.W., Reinsel, G., & Ljung, G. M. (2015). Time Series Analysis, Forecasting And Control, 4th edition. Prentice Hall.
2. Shumway, R. H., & Stoffer, D. S. (2017). Time Series Analysis And Its Applications: With R Examples. Springer.
3. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles And Practice. OTexts.
4. Brockwell, P.J. & Davis, R. A. (2002). Introduction To Time Series Analysis And Forecasting, 2nd edition. Springer.
5. Pankratz, A. E. (1991). Forecasting With Dynamic Regression Models. New York, USA: John Wiley & Sons.

SQB7034 COMPUTER INTENSIVE METHODS

Error in floating point calculations. Probability function and distribution function approximations. Generating random numbers, including evaluating the quality of the generator and calculation methods in linear algebra: Gaussian elimination, sweep operators. Calculation methods for multiple regression, (not constrained) nonlinear regression and model fitting other than the least squares, bootstrap, Markov chain Monte Carlo and EM algorithm.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements Of Statistical Learning: Data Mining, Inference And Prediction. Springer, New York.
2. Ross, S.M. (2013). Simulation. Academic Press, Amsterdam.
3. Chernick, M.R. (2008). Bootstrap Methods: A Guide For Practitioners And Researchers. Auxilium Pharm.
4. Hjorth, J.S. (1994**).** Computer Intensive Methods in Statistics. Chapman & Hall.
5. Zwanzig, S. and Mahjani, B. (2019). Computer Intensive Methods In Statistics. CRC Press, Boca Raton.

SQB7035 ROBUST STATISTICS

Introduction to robust statistics: a review of basic philosophy and concepts central to the area of robust statistics, as well as a review of some basic robust methods. Topics covered include the L-estimates, monotonic and the redescending M-estimates, with applications to the estimation of univariate location and scale; robust regression methods including the M-estimates, the GM-estimates, the least median of squares method (LMS), the S-estimates and the MM-estimates; robust multivariate methods including the M-estimates, the minimum volume ellipsoid method (MVE), and the MM-estimates. Notions of robustness such as breakdown point and influence function will also be reviewed. Computational issues and analysis of real data sets using R software will also be discussed.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Maronna, R.A., Martin, D.R., and Yohai, V.J. (2019). Robust Statistics: Theory and Methods (with R) (Wiley Series in Probability and Statistics, 2nd ed). John Wiley.
2. Huber, P. J., Ronchetti, E.M. (2011). Robust Statistics. John Wiley.
3. Agostinelli, C., Basu, A., Filzmoser, P., & Mukherjee, D. (Eds.). (2016). Recent advances in robust statistics: theory and applications. Springer India.
4. Rousseeuw, P.J. and Leroy, A. M. (1987). Robust Regression And Outlier Detection. John Wiley.

SQB7036 DATA MINING

Introduction to data mining methods and tools. Data cleaning and processing. Data visualisation. Dimensional reduction methods: principal component analysis, variable selection. Data mining algorithms for supervised learning (logistic regression, support vector machine, discriminant analysis, classification and regression trees, etc.) and unsupervised learning (K-means, hierarchical clustering). Ensemble methods (random forests, bagging, boosting). Performance evaluation of classifiers. Interpretation of data mining results. Examples of data mining for real world applications.

Assessment:
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Bramer, M. (2013). Principles of Data Mining. (2nd ed.). London: Springer-Verlag.
2. Flach, P. (2012). Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge: Cambridge University Press.
3. Hastie, T., Tibshirani, R., Friedman, J. (2017). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). New York: Springer.
4. Wickham, H. & Grolemund, G. (2017). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. Sebastopol, CA: O'Reilly Media.

SQB7037 SURVIVAL DATA ANALYSIS

Basic concepts such as survival and hazard functions. Survival data analysis including life table, Kaplan-Maier; log-rank and Wilcoxon tests. Survival regression modeling including the Cox regression model, several parametric models and the accelerated life time model and risk model. Diagnostic checking of the models. Application to the real dataset.

Assessment
Continuous Assessment: 60%
Final Examination: 40%

References:
1. Lee, Elisa T. (2013). Statistical methods for survival data analysis. 4th Ed., John Wiley.
2. Miller, Rupert G. (1998). Survival analysis. John Wiley.
3. Hosmer, D.W., Lemershow, S and May, S. (2011) Applied Survival Analysis: Regression Modeling of Time to Event Data. 2nd Ed., John Wiley.
4. Collett, D. (2015) Modelling Survival Data in Medical Research. 3rd Ed., Chapman and Hall
5. Karim M.R., Islam M.A. (2019). Reliability and Survival Analysis. Springer Singapore.