ACADEMIC CITY UNIVERSITY

Program Handbook

MSc DATA SCIENCE & ANALYTICS

Graduate Programs Team
Academic City University
September 2024

Contents

1.0. Welcome Message

Akwaaba!

To the Faculty of Computational Science and Informatics at Academic City University! And to our Master of Science in Data Science and Analytics program. The MSc in Data Science & Analytics programme aims to train students to use data science techniques to solve problems and make complex decisions in the real world.

This handbook guides your journey through this comprehensive program, which will equip you with essential knowledge, skills, and relevant tools in today's data-driven world. As part of our mission to make you #futureready, we have structured this curriculum to develop your analytical, problem-solving, and strategic decision-making capabilities. These skills are crucial across industries ranging from business and management to engineering and entrepreneurship.

This program will empower you with advanced technical skills and foster a mindset that values strategic insight, adaptability, and leadership. We hope it equips you with the expertise needed to excel as a data science professional, helping shape the future of industries in an era where data drives innovation and competitive advantage.

Dr Naval Sharma
Ag. Dean, Faculty of Computational Science and Informatics.

2.0. Program Objectives

The Master's program in Data Science and Analytics focuses on building proficiency with the most widely used analytics tools and technologies. By the end of this program, you will be able to independently apply these tools to solve complex business challenges and make strategic, data-informed decisions.

You will understand how data science impacts various industries, equipping you to inform development decisions and drive forward-thinking change. You will also learn to derive meaningful insights from data and communicate these effectively to stakeholders, building predictive models that shape and inform future business strategies.

Additionally, the program will immerse you in a broad spectrum of industry-specific knowledge, preparing you for a successful career in data science and analytics.

3.0. Intended Learning Outcomes

Upon completion of the program, you will possess the skills to:

- Analyse, formulate and resolve complex problems across the diverse business, management, entrepreneurship, and engineering domains.
- Apply data mining software and big data analytics tools to tackle real-world challenges.
- Utilise cutting-edge tools and methodologies for big data analysis, preparing you to stay at the forefront of technological advancements.
- Employ algorithms to develop intelligent systems and processes.
- Exhibit essential teamwork, leadership, and decision-making skills, enhancing organisational effectiveness.

4.0. Our Uniqueness

Academic City's MSc in Data Science and Analytics stands out for its innovative, industry-integrated approach, combining the expertise of seasoned faculty with active collaboration from industry leaders. Designed to provide a blend of academic rigour and real-world applicability, the program includes experiential learning and live projects that address impactful organisational challenges. Students gain hands-on experience with industry-standard tools and techniques, ensuring they are well-prepared to meet the

demands of today's data-driven landscape. Instructors are a collaborative blend of industry partners, seasoned internal experts, and professional collaborators.

The unique blended format offers flexible Thursdays and some Friday virtual sessions, complemented by immersive on-campus learning on Saturdays. This enables a comprehensive learning experience tailored for working professionals and full-time students. Internship opportunities and organizational-based tasks further enrich the program.

5.0. Graduation Requirements

1. Satisfy all general University requirements (including Eligibility for the Degree).
2. Pass all the required courses on the Course Diet.
3. Complete the prescribed number of credit hours in each category of course specified for the programme of study.
4. Attain a minimum Cumulative Grade Point Average (G.P.A) of 2.0.
5. Settle all financial and other obligations to the Academic City University.
6. Should maintain acceptable moral conduct at the University.
7. Should have adhered to all University policies and requirements outlined in the Graduate Student Handbook.

6.0. Structure

Lecture days	Lecture Schedule	Contact Hours
Thursday	5:30 pm - 8:30 pm	3
Friday	5:30 pm - 8:30 pm	3
Saturday	9:00 am - 5:00 pm	6
Total hours per week		12
Total hours per module		48

7.0. Assessment of Learning

Method	Weightage	Detail
Continuous Assessments: Assignments/ Class Tests / Project Work/ Mid Semester Exam	35%	Class / Lab / Home Assignments Closed Book / Open Book / Practical Test & Viva Voce Practice Examinations in a pattern similar to End Semester Exam
Attendance & Participation	5%	
End Semester Exam	60%	End Semester Assessment, Practical work/ Projects.

8.0. Program Diet

MSc. Data Science								
TOTAL PROGRAMME CREDITS					33			
Year / Level	S/N	Course Code	Course Status	Course Name	L	T	P	C
Year 1 / Level 500	SEMESTER 1							
	1	DS5101	Compulsory	Foundations of Data Science and Big Data Analytics	2	2	3	3
	2	DS5102	Compulsory	Big Data Management	2	2	3	3
	3	DS5103	Compulsory	Introduction to Data Science	2	2	3	3
	4	DS5104	Compulsory	Mathematical Foundations for Data Science	2	2	3	3
				Total	8	8	12	12
	SEMESTER 2							
	1	DS5201	Compulsory	Deep Learning and Neural Networks	2	2	3	3
	2	DS5203	Compulsory	Statistical Methods for Data Science	2	1	3	3
	3	DS5202	Compulsory	Special Topics in Data Science	2	1	3	3
				Total	6	5	9	9
Year 2 / Level 600	SEMESTER 1							
	1	DS6101	Compulsory	Graduate Qualifying Seminar	2	3	6	3
	2	DS6102	Elective I	Data Warehousing and Information Retrieval	2	2	3	3
	3	DS6103	Elective I	Financial Decision-Making for Value Creation	2	2	3	3
	4	DS6104	Compulsory	MSc. Thesis Phase 1	2	3	6	3
				Total	6	7	12	12
	SEMESTER 2							
	1	DS6201	Compulsory	MSc. Thesis Phase II	2	3	6	3
Grand Total					22	23	39	36

9.0. Modular Structure

Below is how modules will be structured for the Data Science & Analytics program.

Year	Semester	Module	Courses on Code & Title	
1	1	1	DS5104	Mathematical Foundations for Data Science
1	1	2	DS5103	Introduction to Data Science
1	1	3	DS5101	Foundations of Data Science and Big Data Analytics
1	1	4	DS5102	Big Data Management
1	2	1	DS5203	Statistical Methods for Data Science
1	2	2	DS5201	Deep Learning and Neural Networks
1	2	3	DS5202	Special Topics in Data Science

Year	Semester	Module	Courses on Code & Title	
2	1	1	DS6101	Graduate Qualifying Seminar
2	1	2	DS6102	Data Warehousing and Information Retrieval
2	1	3	DS6103	Financial Decision-Making for Value Creation
2	1	4	DS6104	MSc. Thesis Phase I
2	2	1	DS6201	MSc. Thesis Phase II

10.0. Course Briefs

10.1. Year One Semester One Courses:

10.1.1. DS5101 Foundations of Data Science and Big Data Analytics

Required of DS students

Prerequisite(s): Knowledge of Databases Management Systems, Programming in Python or R or any relevant object-oriented programming language, and Data Mining.

Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.1.1.1. Course Objectives

The course will enable students to:

- Examine sizable and multivariate datasets (referred to as Big Datasets).
- Uncover hidden patterns in Big Datasets.
- Extract unknown correlations and other information useful to support decision-making processes at different levels.

10.1.1.2. Course Content

Topics to be covered include: Overview of Big Data; Using Big Data in Businesses; Data Mining Concepts; Machine Learning Principles; Machine Learning Techniques such as Artificial Neural Networks; Case Based Reasoning; Technologies for Handling Big Data; Big Data is primarily characterised by Hadoop; Understanding Hadoop Ecosystem; Dig Deep to understand the fundamental of MapReduce and HBase; Understanding Big Data Technology Foundations; Databases and Data Warehouses; Using Hadoop to store data; Learn to Process Data using Map Reduce; Testing and Debugging Map Reduce Applications; Learn Hadoop YARN Architecture; Exploring Hive; Exploring Pig; Exploring Oozie; Learn NoSQL Data Management; Cloud Computing, Scalability, Big Data Design, Polyglot systems; Schemaless databases; Key-value stores; Wide-column stores; Document-stores, Distributed Data Management, Transparency layers; Distributed file systems; File formats; Fragmentation; Replication and synchronisation; Sharding; Distributed hash; LSM-Trees, In-memory Data Management, NUMA architectures; Columnar storage; Late reconstruction; Light-weight compression, Distributed Data Processing, Distributed Query Processing; Sequential access; Pipelining; Parallelism; Synchronisation barriers; Multitenancy; MapReduce; Resilient Distributed, Stream management and processing, One-pass algorithms; Sliding window; Stream to relation operations; Micro-

batching; Sampling; Filtering; Sketching; Big Data Architectures - Centralised and distributed functional architectures of relational systems, Lambda architecture.

10.1.1.3. Mode of Delivery

The course would be delivered through lectures, tutorials, laboratory sessions, term papers, presentations, mini-projects, and discussions.

10.1.1.5. Reading Material

The following are recommended reference books:

1. Devlin, B. (2022) Business unintelligence: Insight and Innovation Beyond Analytics and Big Data. 4th edition. Technics Publications;
2. Cukier, K., and Mayer-Schönberger, V. (2020) *Big Data: A Revolution That Will Transform How We Live, Work, and Think.* 3rd edition. Harper Business;
3. Piegorsch, W. W. (2019) Guide to Big Data Analytics: Origination to Opportunities. 4th edition. Wiley Publishers;
4. Alapati, S. R. (2016) Expert Hadoop Administration: Managing, Tuning, and Securing Spark, YARN. Addison-Wesley Publishers
5. Khan, G. F. (2018) Creating Value with Social Media Analytics: Managing, Aligning, and Mining Social Media Text, Networks, Actions, Location, Apps, Hyperlinks, Multimedia, & Search Engines Data. CreateSpace Independent Publishing Platform
6. Ben-David, S. and Shalev-Shwartz, S. (2018) *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.

10.1.2. DS5102 Big Data Management

Required of DS students

Prerequisite(s): Knowledge of Microsoft Excel and basic knowledge of the R language. Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.1.2.1. Course Objectives

By the end of this course, students should be able to:

- Explain the concept of Big Data and Big Data Management
- State the storage capabilities of Big Data
- Understand the processes of extracting Big Data using predictive analytics techniques

10.1.2.2. Course Content

Topics to be covered include: Big data definition and taxonomy; Big data tools - Hadoop & Spark; Setting up the Lab environment; Using Spark with R; SparkSQL in R; SQL Language; SparkSQL; Predictive Analytics on Big Data; Spark Machine Learning with R; Deep learning with H2O and R; Deep learning algorithms with R and H2O; Text analytics; Text analytics and Natural Language Processing and Sentiment analysis; Big Data, Cloud Computing, Scalability, Big Data Design, Polyglot systems; Schemaless databases; Key-value stores; Wide-column stores; Document-stores, Distributed Data Management, Transparency layers; Distributed file systems; File formats; Fragmentation; Replication and synchronisation; Sharding; Distributed hash; LSMTrees, In-memory Data Management, NUMA architectures; Columnar storage; Late reconstruction; Light-weight compression, Distributed Data Processing, Distributed Query Processing; Sequential access; Pipelining; Parallelism; Synchronisation barriers; Multitenancy; MapReduce; Resilient Distributed, Stream management and processing, One-pass algorithms; Sliding window; Stream to relation operations; Micro-batching; Sampling; Filtering; Sketching; Big Data Architectures - Centralised and distributed functional architectures of relational systems, Lambda architecture.

10.1.2.3. Mode of Delivery

The course would be delivered through lectures, tutorials, laboratory sessions, term papers, presentations, mini-projects, and discussions.

10.1.2.4. Reading Material

The following are recommended reference books:

1. Apache Hadoop (2022) HDFS Architecture. Apache Hadoop Publishers;
2. Lemahieu, W., Vanden Broucke, S., and Baesens, B. (2021) Principles of database management: practical guide to storing, managing and analysing big and small data. Cambridge University Press;
3. Verzani, J. (2020) SimpleR – using R for introductory Statistics. CUNY Publishers;
4. Somani, A. K. and Deka, G. C. (2018) Big data analytics: tools and technology for effective planning. CRC Press.
5. Sadalage, P. J. and Fowler, M. (2013) NoSQL distilled: a brief guide to the emerging world of polyglot persistence. RR Donnelley Publication.
6. Harvey, C. (2017) Big Data Management. Datamation Website Publication [Accessed via https://www.datamation.com/big-data/big-data-management/ on 5th January 2023 at 2:19pm]

10.1.3. DS5103 Introduction to Data Science

Required of DS students

Prerequisite(s): Basic knowledge of Statistics, Object Oriented Programming in Python or R Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.1.3.1. Course Objectives

This course will equip students with the knowledge and skills to:

- Develop and utilise data science tools for acquiring, cleaning, analysing, exploring, and visualising data
- Support data-driven inferences and decisions as well as effectively communicate results.
- Enhance further study in data mining, machine learning, and artificial intelligence with R/Python.

10.1.3.2. Course Content

Topics to be covered include: Introduction to data analysis tools in Python and R; Data structures with Pandas; Introductory hypothesis testing and statistical inference; Web scraping and data acquisition via APIs; Classification methods; Logistic regression; k-nearest neighbours; Decision trees; Support vector machines and neural networks; Data visualisation; Clustering methods; Dimensionality reduction; Including principal component analysis; Network analysis; Rating, ranking, and elections; Cleaning and reformatting messy datasets using regular expressions or dedicated tools such as open refine; Natural language processing; Ethics of big data.

10.1.3.3. Mode of Delivery

The course would be delivered through lectures, tutorials, laboratory sessions, term papers, presentations, mini-projects, and discussions.

10.1.3.4. Reading Material

The following are recommended reference books:

1. Géron, A. (2023) Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 3rd edition. Boston: O'Reilly Media Inc.
2. McKinney, W. (2018) Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. 2nd edition. Boston: O'Reilly Media Inc.
3. Chollet, F. (2018). Deep Learning with Python. 1st edition. *Shelter Island, New York:*Manning Publications Co.

4. Harrison, M. (2016) *Learning the Pandas Library: Python Tools for Data Munging, Data Analysis, and Visualization*. *CreateSpace Independent Publishing Platform*.
5. Grus, J. (2019) Data Science from Scratch: First Principles with Python. Boston: O'Reilly Media Inc.;
6. Leskovec, J., Rajaraman, A. and Ullman, J. D. (2020) *Mining of Massive datasets*. *Cambridge: Cambridge University Press*.

10.1.4. DS5104 Mathematical Foundations for Data Science

Required of DS students

Prerequisite(s): Knowledge of linear algebra and multivariable calculus. Students should also have basic Python programming abilities.

Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.1.4.1. Course Objectives

By the end of this course students should be able to:

- Explain the fundamental mathematical concepts relevant to computer science
- Apply the fundamental mathematical concepts to data science and statistical machine learning.

10.1.4.2. Course Content

Topics to be covered include: Typology of problems; Importance of linear algebra, statistics and optimization from a data science perspective; Structured thinking for solving data science problems. Matrices and their properties (determinants, traces, rank, nullity, etc.); Eigenvalues and eigenvectors; Matrix factorizations; Inner products; Distance measures; Projections; Notion of hyperplanes; half-planes, Probability, Statistics and Random Processes: Probability theory and axioms; Random variables; Probability distributions and density functions (univariate and multivariate); Expectations and moments; Covariance and correlation; Statistics and sampling distributions; Hypothesis testing of means, proportions, variances and correlations; Confidence (statistical) intervals; Correlation functions; White-noise process. Optimisation, Unconstrained optimisation; Necessary and sufficient conditions for optima; Gradient descent methods; Constrained optimisation, KKT conditions; Introduction to non-gradient techniques; Introduction to least squares optimisation; Optimization view of machine learning., Linear regression as an exemplar function approximation problem; Linear classification problems, language duality, Simplex algorithms, dynamic programming, semi-defined programs, Support vector machines, portfolio optimisation, feature selection, optimal power flow, recommendation systems

10.1.4.3. Mode of Delivery

The course would be delivered through lectures, tutorials, laboratory sessions, group presentations and discussions.

10.1.4.4. Reading Material

The following are recommended reference books:

1. Phillips, J. M. (2021). *Mathematical Foundations for Data Analysis*. 1st edition. Springer Series in the Data Sciences
2. Shikhman, V., & Müller, D. (2021). Mathematical Foundations of Big Data Analytics.1st edition. Springer Gabler;
3. Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2019). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Vol. 2, pp. 1-758. New York: Springer Series in Statistics.
4. Peyré, G. (2021). Mathematical Foundations of Data Sciences. SPRINGER Series in Data Sciences
5. Dehmer, M., Moutari, S., & Emmert-Streib, F. (2020) *Mathematical Foundations of Data Science Using R*. De Gruyter STEM Publication Series.

10.2. Year One Semester Two Courses

10.2.1. DS5201 Deep Learning and Neural Network

Required of DS students

Prerequisite(s): A strong mathematical background in calculus, linear algebra, and probability & statistics, Machine Learning, Algorithms & Data Structures, as well as programming in Python and C/C++. Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.2.1.1. Course Objectives

This course will help students to:

- Develop artificial neural networks for machine learning.
- Understand some prominent network architectures such as multi-layer feedforward neural networks, convolutional neural networks (CNNs), auto-encoders, recurrent neural networks (RNNs), and generative-adversarial networks (GANs).
- Understand optimisation and regularisation techniques.
- Train neural networks for various applications, such as object detection, facial expression recognition, handwriting analysis, and natural language processing.

10.2.1.2. Course Content

Topics to be covered include Deep reinforcement learning, Visual Question Answering, Visual Dialog, Novel deep methods (Deep internal learning, Deep image prior), Neural network architectures such as Convolutional Neural Networks, Recurrent Neural Networks, LSTMs, Transformers, and learning how to make them better with strategies such as Dropout, BatchNorm, Xavier/He initialisation, Industry applications using Python and TensorFlow, Real-world cases such as speech recognition, music synthesis, chatbots, machine translation, natural language processing, learning via gradient descent; recursive chain rule (backpropagation); if time: bias variance tradeoff, regularisation; output units: linear, softmax; hidden units, deep belief NETS, recurrent neural networks; sequence modelling; backpropagation through time; vanishing/exploding gradient problem; gradient clipping, long-short term memory (LSTM). AI and Deep Learning, recurrent network structures, deep unsupervised and reinforcement learning, and applications to problem domains like speech recognition and computer vision

10.2.1.3. Mode of Delivery

The course would be delivered through lectures, tutorials, laboratory sessions, group presentations and discussions.

10.2.1.4. Reading Material

The following are recommended reference books:

1. Murty, M. N., & Devi, V. S. (2015). *Introduction to pattern recognition and machine learning*. Vol. 5. World Scientific Publishing Co. Pte Ltd.
2. Rokach, L. (2019). Ensemble learning: pattern classification using ensemble methods.

2nd Edition. World Scientific Publishing Co. Pte Ltd.

3. Zhou, Z. H. (2021). *Machine learning*. Springer Nature.
4. Russell, S. and Norvig, N. (2018) *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence.
5. Hastie, T., Tibshirani, R. and Friedman, J. (2017) *The Elements of Statistical Learning*. Corrected 12th printing. Springer.
6. Bishop, C. M. (1995) *Neural Networks for Pattern Recognition*. Oxford University Press.
7. Pal, S. K., Ray, S. S., & Ganivada, A. (2017). *Granular neural networks, pattern recognition and bioinformatics*. Springer International Publishing.

10.2.2. DS5202 Special Topics in Data Science

Required of DS students

Prerequisite(s): Basic Data Science Knowledge and Skills Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.2.2.1. Course Objectives

By the end of the course, students should be able to:

- Examine research topics that are trending in the data science domain.
- Explore relevant papers covering applications, algorithms, systems, and theory-with a focus on the most recent developments will be discussed and analysed.

10.2.2.2. Course Content

Topics to be covered include: Social Data Analytics, Security, Surveillance, distributed & parallel approaches, join processing, and imprecise data/approximation, computational reproducibility in data science, data explorations, recommender systems: matrices, bandits, and neuron, uncertainties, parallel query processing, web analytics, technical informatics, social computing, containerising data sciences, technical paper writing, agent-based simulations, large-scale data analysis, text as data (natural language processing), digital experiments, mass collaboration, and ethics.

10.2.2.3. Mode of Delivery

The course would be delivered through lectures, tutorials, term paper preparation(s), conference participation(s), articles prepared for Journals, group presentations and discussions.

10.2.2.4. Reading Material

The following are recommended reference books:

1. Beheshti, A., Ghodratnama, S., Elahi, M., & Farhood, H. (2022). *Social Data Analytics*. CRC Press.
2. Schwabish, J. (2021). Better data visualizations: *A guide for scholars, researchers, and wonks*. Columbia University Press.
3. Nick Singh (2021): Ace the Data Science Interview: 201 Real Interview Questions Asked By FAANG, Tech Startups, & Wall Street. Ace the Data Science Interview Publication.
4. Abhari, A. (2018) *Topics in data science with practical examples*. CreateSpace Independent Publishing Platform (Abdolreza Abhari).

5. Zhao, Z., De Stefani, L., Zgraggen, E., Binnig, C., Upfal, E., & Kraska, T. (2017, May). Controlling false discoveries during interactive data exploration. In Proceedings of the 2017 ACM international conference on management of data (pp. 527-540).
6. Chung, Y., Krishnan, S., & Kraska, T. (2016). A data quality metric (DQM): how to estimate the number of undetected errors in data sets. *arXiv preprint arXiv:1611.04878*.

10.2.3. DS5203 Statistical Methods for Data Science

Required of DS students

Prerequisite(s): Knowledge of Basic Descriptive and Inferential Statistical Methods Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.2.3.1. Course Objectives

By the end of this course, students should be able to:

- Understand descriptive and inferential statistics,
- Apply data presentation in graphic or illustrative form,
- Understand the calculation of statistical and decision-making parameters.
- Utilise methods from Mathematical Statistics (probability theory and inference theory) to understand random variations and identify patterns in the data collected as well as provide a basic understanding of the statistical biases in the field of data science.

10.2.3.2. Course Content

Topics to be covered include: Data analysis, including descriptive statistics and data visualisation; Probability theory, including basic probability calculations, random variables, and distributions; Statistical methods, including point and interval estimates, hypothesis testing, regression, probabilistic models in data science applications, for instance, exploratory analysis, inferential analysis, causal analysis, logistic regression, Naive Bayes classifiers and topic models for text or Hidden Markov Models for sequences. Comparisons and ANOVA, standard parametric statistical models, multivariate data visualisation, multiple linear regression and classification, classification and regression trees and random forests. The course's important focus is statistical computing and reproducible statistical analysis. Students are introduced to “R”, the widely used statistical language, and obtain hands-on experience implementing a range of commonly used statistical methods on real-world datasets.

10.2.3.3. Mode of Delivery

The course will be delivered through lectures, tutorials, laboratory sessions, case studies, exercises, group presentations, and field trips.

10.2.3.4. Reading Material

The following are recommended reference books:

1. Sohil, F., Sohali, M. U., & Shabbir, J. (2022). *An introduction to statistical learning with applications in R:* by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, New York: Springer Science and Business Media, 2013, $41.98, eISBN: 978-1-46147137- 7.
2. Lay, D. C., Lay, S. R., & McDonald, J. (2016). *Linear algebra and its applications*. Pearson Education.
3. Montgomery, D. C., & Runger, G. C. (2020). *Applied statistics and probability for engineers*. John Wiley & Sons.
4. Petruccelli, J., Nandram, B., and Chen, M. (2019). Applied Statistics for Engineers and Scientists. Pearson;
5. Yaser, S. Abu-Mostafa, Magdon-Ismail, M. and Hsuan-Tien, L. (2019). Learning From Data. *AMLBook Publishers*;
6. Long, J. D., & Teetor, P. (2019). R cookbook: proven recipes for data analysis, statistics, and graphics. O'Reilly Media.
7. Hastie, T., Tibshirani, R. and Friedman, J. (2017). *The Elements of Statistical Learning*. Corrected 12th printing. Springer.

10.3. Year Two Courses:

10.3.1. DS6101 Graduate Qualifying Seminar

Required of DS students

Prerequisite(s): completion of at least 24 credits of the DS degree or consent of the instructor. Corequisite(s): None

Course Credit=4; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.3.1.1. Course Objectives

The objectives of the course are as follows:

- Equip students with problem-solving skills in teams.
- Support students' practice application of data science theories, tools and techniques.
- Help students demonstrate professional development skills, such as communication, teamwork, leadership, collaboration, and storytelling.

10.3.1.2. Course Content

Students are to liaise with relevant industrial government or corporate organisations to identify a data science need(s) and formulate/develop solutions to address the challenge(s).

10.3.1.3. Mode of Delivery

The course will be delivered through industrial visits, group presentations, discussions, and report writing.

10.3.1.4. Reading Material

The following are recommended reference books:

1. Rob, T. J. (2022) Before you write your first word...: what you should know about self-publishing that no one tells you! *T.J. Rob Books Publications*;
2. Bean, R. (2021). Fail fast, learn faster: lessons in data-driven leadership in an age of disruption, big data, and AI. *Hoboken: John Wiley and Sons Inc.;*
3. Kleppmann, M. (2021). Designing data-intensive applications: the big ideas behind reliable, scalable, and maintainable systems. 17th edition. Sebastopol: O'Reilly Media Inc.
4. Atherton, T. (2020). Technical report writing and style guide: how to write even better technical reports. Independent Publisher (W. A. Atherton).

5. Greenhall, M. (2010). Report writing skills training course: how to write a report and executive summary, plan, design, and present your report, in an easy format. *Universe of Learning Publishers*.

10.3.2. DS6102 Data Warehousing and Information Retrieval (Elective)

Required of DS students

Prerequisite(s): Good knowledge of statistics, mathematical modelling and web search engines Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.3.2.1. Course Objectives

The course will enable students to:

- Understand the theory, design, and implementation of text-based and web-based information retrieval systems.
- Explain the key concepts and models relevant to information retrieval and natural language processing on large-scale corpus such as the web and social systems.
- Understand the basic principles and practical algorithms for information retrieval and text mining on web pages, emails, news, and scientific literature.

10.3.2.2. Course Content

Topics to be covered include: Vector space model, crawling, indexing, web search, ranking, recommender systems, embedding and language model, Efficient text indexing, Boolean and vector-space retrieval models, Evaluation and interface issues, IR techniques for the web, including crawling, link-based algorithms, and metadata usage, document clustering and classification, traditional and machine learning-based ranking approaches, retrieval models, retrieval evaluation, search analysis, search engine architecture text categorisation, Basic Tokenizing, Indexing, and Implementation of Vector-Space Retrieval, query operations, text representation, clustering, extraction and integration, deep learning for IR and recommender systems, Social Networks, Natural Language Processing, social search, link analysis and social search.

10.3.2.3. Mode of Delivery

The course would be delivered through lectures, tutorials, laboratory sessions, group presentations and discussions.

10.3.2.4. Reading Material

The following are recommended reference books:

1. Safraz, M. (2019). Critical Approaches to Information Retrieval Research. United States: IGI Global.

2. Ricardo Baeza-Yates and Berthier Ribeiro-Neto (2019) *Modern Information Retrieval*. Reading: Addison-Wesley Publishers
3. W. Bruce; Metzler, Donald; Strohman, Trevor. Addison Wesley (2018). *Search engines: Information retrieval in practice*. Reading: Addison-Wesley Publishers.
4. Buttcher, S., Clarke, C. L., & Cormack, G. V. (2016). *Information retrieval: Implementing and evaluating search engines*. MIT Press.
5. Bhatia, P. (2019). Data mining and data warehousing: principles and practical techniques. Cambridge University Press.

10.3.3. DS6103 Financial Decision-Making for Value Creation (Elective)

Required of DS students Prerequisite(s): None Corequisite(s): None

Course Credit=3; Lecture Hours=2; Tutorial Hours=0; Practical Hours=3

10.3.3.1. Course Objectives

This course will help students to:

- Understand some key financial decisions modern corporations face, as well as the alternative methods that can be employed to optimise the value of the firm's assets.
- Develop and enhance their ability to implement and communicate a firm's financial decisions related to value creation.
- Emphasise the relationships among a firm's strategic objectives, financial accounting and financial statement data, economic events, responses by market participants and other impacted constituencies, and corporate finance theory to uncover the decision-maker's perspective.

10.3.3.2. Course Content

Topics to be covered include: Defining and establishing value objectives; In-depth analysis of the financial principles; Centrality of opportunity cost of capital; Quantifying value principles into Present Value; Assessing costs and benefits of decisions and investments with Net Present Value; Review of Accounting and Financial Analysis; Managerial decision-making and the limits of financial ratio; Risk as the critical benchmark; Understanding trade-offs between operating and asset efficiencies; Integrating cost of capital in performance measurement; Economic Value Added as a decision-making tool; Understanding and correcting managerial misalignment; Effective incentive compensation, Reconciliation of an HR-centric approach with shareholders'; Strategy in the Context of Shareholder Value; Understanding the relationship of strategy and value creation; Reconciliation of strategic and financial objectives; Integrating a balanced scorecard into the value equation; Cash Flow-based framework; Prioritising investment projects; Explaining stock price volatility; Strategic Transactions, mergers and acquisitions; Strategic rationale; Acquisition valuation.

10.3.3.3. Mode of Delivery

The course would be delivered through lectures, tutorials, term papers, team-based discussions and presentations, real-world simulations, and various financial modelling tools.

10.3.3.4. Reading Material

The following are recommended reference books:

1. Lawrie, M. and Tsetsekos, G. (2021). Business leadership in turbulent times: decisionmaking for value creation. Bloomington: Archway Publishing;
2. Davis, S., Copeland, C. and Wertheimer, R. (2020). Lessons from the titans: what companies in the new economy can learn from the great industrial giants to drive sustainable success. New York: McGraw-Hill.
3. Baruch Lev & Feng Gu (2016). The end of accounting and the path forward for investors and managers. Hoboken: John Wiley and Sons Inc.;
4. Bonime-Blanc, A. (2019). Gloom to boom: how leaders transform risk into resilience and value. Abingdon: Routledge;
5. Karlson, L. C. (2015). Corporate value creation: an operations framework for nonfinancial managers. Hoboken: John Wiley and Sons Inc.

10.3.4. DS6104/6201 MSc Thesis Phase I/Phase 2

Required of DS students

Prerequisite(s): Knowledge of Report Writing, Project Documentation and Technical Compositions.

Corequisite(s): None

Course Credit=6; Lecture Hours=0; Tutorial Hours=3; Practical Hours=6

10.3.4.1. Course Objectives

This course will equip students with skills to:

- Embark on original contributions and advancements in the frontiers of knowledge.
- Write and Communicate research findings.

10.3.4.2. Course Content

Students are to choose from various core and tangential domains in Data Science and develop a project-driven thesis to be written and submitted as a capstone project at the end of the programme. The students are expected to choose a research project that they have spent considerable time in considering and researching (literature review), project design (formulation of a hypothesis), data collection (field and or laboratory), analysis (statistical examination of the data), and finally presentation and synthesis (examination of the statistical results in the context of your hypothesis and literature review). Each of these individual parts will consume considerable time and effort. Therefore, they are appointed supervisors/advisors responsible for guiding students through the research process. Supervisors are appointed upon being offered admission into the programme.

10.3.4.3. Mode of Delivery

The course will be delivered through seminars, presentations, and a final thesis defence.

10.3.4.4. Reading Material

The following are recommended reference books:

1. Siochrú, C. Ó. (2022). A Student Guide to Writing Research Reports, Papers, Theses and Dissertations. Taylor & Francis.;
2. Robbo, J. (2022). APA manual: the simplified guide to APA style formatting, citations and referencing for writers, researchers and students. 7th edition. *Independent publishers (Joe Robbo);*
3. Chris Ifeanyi Ezeh (2021). Learn best practices of referencing & citing in different styles: - present excellent research work - master referencing rules with ease - shine in MLA, APA, and Chicago styles. *EuroAfrica Media Network;*

4. Kornuta, H. M., & Germaine, R. W. (2019). A concise guide to writing a thesis or dissertation: Educational research and beyond. Abingdon: Routledge.
5. Turabian, K. L. (2018). A manual for writers of research papers, theses, and dissertations: Chicago style for students and researchers. University of Chicago Press.
6. Greenhall, M. (2010). Report writing skills training course: how to write a report and executive summary, plan, design, and present your report in an easy format. *Universe of Learning Publishers*.

ACADEMIC CITY UNIVERSITY

#AskACity

Academic City University
Haatso- Agbogba, Accra (Ghana)
www.acity.edu.gh
+233 26 269 3870
info@acity.edu.gh
@acitygh