
College of Computing and Informatics

# STUDY PLAN PROJECT

# MASTER OF DATA SCIENCE

September 2023

## Introduction

The College of Computing and Informatics offers Master of Science in Data Science program that aims to qualify students with high academic skills in aspects related to data science and usage of data analysis software, providing students with the latest tools and methods in big data technologies for the next generation. The program focuses on combining the cognitive and applied aspects in the field of data science, machine learning and artificial intelligence; and practically apply these technologies in problem solving.

## The Importance and Reasons for Creating the Program

- Data science is considered as the most exciting specialty in the twenty-first century, as a result of the great development in usage of Internet technologies, social networking applications and the Internet of things, therefore, we now have huge amounts of data that are difficult to handle and analyze by the traditional statistical methods. Thus, the specialty of data science has become called the oil of the twenty-first century.
- The application of modern science and artificial intelligence techniques to analyze data and extract knowledge patterns has become one of the biggest challenges in the current century. The labor market is still suffering from a severe shortage of qualified personnel to meet the need for work.
- Therefore, the College of Computing and Informatics in the Saudi Electronic University presenting an integrated program for the Master of Data Science, which was built and prepared according to international standards and conform with the latest techniques and methods to qualify students to meet the major challenges in the field of data science successfully and creatively.

## Program Objectives

1- Balance between data science studies theory and practical work.
2- Develop both academic and professional skills in the domain of data science and big data analytics.
3- Prepare learners for the data science profession or continued study.
4- Implementing best practices to develop comprehensive project management plan.
5- Prepare the learner to meet the business needs in areas where data science skills are required in various sectors.

# Duration of Study in the Program

Four semesters for MSc Data Science.

## Program Learning Outcomes

1- Develop algorithmic, computational, and statistical models in data science.

2- Extract, transform, integrate, load, and access large data sets.

3- Evaluate opportunities to employ data science solutions for business forecasting and analytics.

4- Synthesize principles of descriptive, predictive, and prescriptive analytics to address challenges.

5- Create deep learning programs to support the analysis of complex datasets.

6- Differentiate between the major theories of machine learning and neural networks.

7- Visualize data for exploration, analysis, and communication.

8- Use machine learning and optimization models to decision making.

9- Apply problem-solving strategies to data analytics.

10- Articulate analytical conclusions and recommendations in written and visual formats.

11- Assemble computational pipelines to support data science from widely available tools.

12- Understand management, ethical, privacy, and accountability issues in data science.

## Career Opportunities for Graduates of the Program

1- Statistician

2- Data Administrator

3- Computer Systems Analyst

4- Data Scientist

5- Software Developer

6- Data Analyst

7- Big Data Engineer

8- Financial Data Analyst

9- Machine Learning Engineer

10- Data Manager

11- Business Intelligence Engineer

12- Big Data Administrator

13- Data Mining Analyst

14- Data Engineer

15- Big Data Architect

16- Data Visualization Developer

## Vision

## Mission

## Program Study Plan

The Master of Data Science program contains 12 courses, distributed over 4 semesters.The program is only offered in English.

Specialization Requirements: (36 Credits)

1. **CS501:** Research Methods in Computational Studies (3 credits)
2. **DS510:** Statistics for Data Science (3 credits)
3. **DS540:** Advanced Python for Data Science (3 credits)
4. **DS520:** Big Data Processing and Analytics (3 credits)
5. **DS560:** Advanced Data Mining (3 credits)
6. **DS630:** Artificial Intelligence for Data Science (3 credits)
7. **DS610:** Advanced Applied Statistics for Data Science (3 credits)
8. **DS620:** Data Visualization (3 credits)
9. **DS550:** Machine Learning Algorithms for Data Science (3 credits)
10. **DS650**: Predictive Analytics for Business (3 credits)
11. **DS660:** Deep Learning Techniques (3 credits)
12. **DS698:** Capstone Project in Data Science (3 credits)

## Program Structure

| # | Course Code | Course Title | Credit hours | Pre-requisites |
|---|---|---|---|---|
| 1 | DS501 | Research Methods in Computational Studies | 3 | |
| 2 | DS540 | Advanced Python for Data Science | 3 | |
| 3 | DS510 | Statistics for Data Science | 3 | |
| 4 | DS560 | Advanced Data Mining | 3 | |
| 5 | DS520 | Big Data Processing and Analytics | 3 | DS510, DS540 |
| 6 | DS630 | Artificial Intelligence for Data Science | 3 | DS540 |
| 7 | DS550 | Machine Learning Algorithms for Data Science | 3 | DS630, DS520 |
| 8 | DS610 | Advanced Applied Statistics for Data Science | 3 | DS510 |
| 9 | DS620 | Data Visualization | 3 | DS560 |
| 10 | DS660 | Deep Learning Techniques | 3 | DS630 |
| 11 | DS650 | Predictive Analytics for Business | 3 | DS610, DS560 |
| 12 | DS698 | Capstone Project in Data Science | 3 | Department Approval |
| Total Credits | | | 36 | |

# Program Structure by Levels

## First Year

### Level One

| # | Course Code | Course Title | Credit Hours | Pre-Requisites |
|---|---|---|---|---|
| 1 | DS501 | Research Methods in Computational Studies | 3 | |
| 2 | DS540 | Advanced Python for Data Science | 3 | |
| 3 | DS510 | Statistics for Data Science | 3 | |

### Level Two

| # | Course Code | Course Title | Credit Hours | Pre-Requisites |
|---|---|---|---|---|
| 1 | DS560 | Advanced Data Mining | 3 | |
| 2 | DS520 | Big Data Processing and Analytics | 3 | DS510, DS540 |
| 3 | DS630 | Artificial Intelligence for Data Science | 3 | DS540 |

## Second Year

### Level Three

| # | Course Code | Course Title | Credit Hours | Pre-Requisites |
|---|---|---|---|---|
| 1 | DS550 | Machine Learning Algorithms for Data Science | 3 | DS630, DS520 |
| 2 | DS610 | Advanced Applied Statistics for Data Science | 3 | DS510 |
| 3 | DS620 | Data Visualization | 3 | DS560 |

### Level Four

| # | Course Code | Course Title | Credit Hours | Pre-Requisites |
|---|---|---|---|---|
| 1 | DS660 | Deep Learning Techniques | 3 | DS630 |
| 2 | DS650 | Predictive Analytics for Business | 3 | DS610, DS560 |
| 3 | DS698 | Capstone Project in Data Science | 3 | Department Approval |

## Program Courses Descriptions

| | |
|---|---|
| Course Title | Research Methods in Computational Studies |
| Course Code | DS501 |
| Pre-requisite(s) | |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course provides an overview of the important concepts of research design, data collection, statistical and interpretative analysis, and final report presentation. The focus of this course is not on mastery of statistics but on the ability to use research in Computational Studies. Students will prepare a preliminary research design for projects in their subject matter areas and how to accurately collect, analyze and report data. Students will focus on the steps needed to design an individual research project or thesis. The course provides real world active learning assignments that seek to integrate the knowledge and skills gained through undergraduate course work. The course focuses on scientific writing, and oral, written, and graphical presentation of data and research results. |
| | |

| | |
|---|---|
| Course Title | Statistics for Data Science |
| Course Code | DS510 |
| Pre-requisite(s) | |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course provides an overview of data analysis, data production, and statistical inference. Areas of study include: surveys and designed experiments, randomization, causation, regression, and inference using hypothesis tests. This course also explores using statistical methods for data analysis to improve enterprise performance and quality, effectiveness, and marketability. Statistical software will be utilized to conduct a predictive analysis, analyze the results, and document the findings. The preparation of input data for analysis using R analytical package is also performed. |
| | |

| | |
|---|---|
| Course Title | Advanced Python for Data Science |
| Course Code | DS540 |
| Pre-requisite(s) | |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | In this course students will gain an advanced knowledge of programming, design, and testing concepts using Python. Students are introduced to the fundamentals of Python scripting and will become proficient in writing modular Python classes. At the core of class method development, students will write Python methods using lists, dictionaries, conditional logic, and looping controls. Students will also cover how to manipulate and analyze un-curated datasets, utilizing basic statistical analysis and machine learning methods, and visualization results. |
| | |

| | |
|---|---|
| Course Title | Big Data Processing and Analytics |
| Course Code | DS520 |
| Pre-requisite(s) | DS510, DS540 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | In this course students will identify the tools and techniques for analyzing big data for organizations to follow in creating and sustaining an effective data science function. Students will apply forecasting, simulation, and data modeling for complex problem analysis in medium to large organizations including the use of Apache Hadoop and Spark, and NoSQL Databases. |
| | |

| | |
|---|---|
| Course Title | Advanced Data Mining |
| Course Code | DS560 |
| Pre-requisite(s) | |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | In this advanced course students will investigate various statistical approaches used for data mining analyses. Students will prepare data suitable for analysis from an enterprise data warehouse using SQL and document results. Students will also create a data mining analysis project to demonstrate their understanding of the concepts. |
| | |

| | |
|---|---|
| Course Title | Artificial Intelligence for Data |
| Course Code | DS630 |
| Pre-requisite(s) | DS540 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course explores recent advances in artificial intelligence and incorporates multiple ideas from basic machine learning and assumes familiarity with machine learning concepts. Topics range from human-computer interfaces, computational methods for intelligent control of autonomous agents, programming for pattern recognition, planning for flexible and reactive systems. Core techniques and applications may include game playing, multi-agent coordination; negotiation planning, logical representation, minimax search, Markov decision processes, and other relevant approaches. |
| | |

| | |
|---|---|
| Course Title | Advanced Applied Statistics for Data Science |
| Course Code | DS610 |
| Pre-requisite(s) | DS510 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | In this course, students will develop a level of competency in applying R for data science. The course covers the basic and intermediate topics in R including variables and basic operations, vectors, matrices, data frames and lists. In addition, students will dive deeper into the graphical capabilities of R and create data visualizations. |
| | |

| | |
|---|---|
| Course Title | Data Visualization |
| Course Code | DS620 |
| Pre-requisite(s) | DS560 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course teaches the essential and practical skills in data visualization and knowledge representation, including computer graphics, visual data representation, physical and human vision models, numerical representation of knowledge and concepts, animation techniques, pattern analysis, and computational methods. Students will gain essential and practical skills in visualization. |
| | |

| | |
|---|---|
| Course Title | Machine Learning Algorithms for Data Science |
| Course Code | DS550 |
| Pre-requisite(s) | DS520, DS630 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course focuses on the concepts and constructs of data structures and algorithms that are widely used in machine learning in data science. Data structures is a key computer science discipline that focuses on understanding how to efficiently and effectively organize data. This course will present a number of advanced conceptual and algorithmic topics related to software maintainability, efficiency, and algorithm analysis for machine learning. The topics presented in this course will range from introducing abstract data types (ADTs) such as bags, stacks, queues, deques, and priority queues, to further analyzing the efficiency associated with the ADTS and other algorithms. |
| | |

| | |
|---|---|
| Course Title | Predictive Analytics for Business |
| Course Code | DS650 |
| Pre-requisite(s) | DS560,DS610 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course covers the fundamental predictive analytics and data mining approaches applied in business. It introduces basic concepts and techniques to discover patterns in data, identify variables with the most predictive power, and develop predictive models. Advanced predictive models from business cases will be examined. |
| | |

| | |
|---|---|
| Course Title | Deep Learning Techniques |
| Course Code | DS660 |
| Pre-requisite(s) | DS630 |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This course covers the fundamental predictive analytics and data mining approaches applied in business. It introduces basic concepts and techniques to discover patterns in data, identify variables with the most predictive power, and develop predictive models. Advanced predictive models from business cases will be examined. |
| | |

| | |
|---|---|
| Course Title | Capstone Project in Data Science |
| Course Code | DS698 |
| Pre-requisite(s) | Department Approval |
| Credit hours | 3 |
| Contact hours | 3 |
| Course Description | This capstone course provides students with the opportunity to demonstrate competency on the key domains of data science. Students will integrate concepts learned throughout the entire program to develop a comprehensive project in a specific domain of analytics, such as web analytics, social media analytics, big data analytics, or healthcare analytics. Students will undertake a data science problem from data collection and model construction through analysis and presentation of results and recommendations for specific business decisions culminating in a final, publishable paper. |
| | |