

# DATA SCIENCE

MASTER PROGRAMME (2021-2023 EDITION)

## STRUCTURE AND CONTENT DESCRIPTION

### 1 Core subjects

#### 1.1 Quantitative Fundamentals (45 hrs. / 6 ECTS)

The program provides a comprehensive set of tools to support advanced studies in the area of quantitative finance, management and logistics or to prepare for employment in the financial or investment industries. You will be introduced with quantitative data fundamentals and also learn where to start when it comes to analysing quantitative data. Topics include: programming for data analysis, computational linear algebra, and quantitative methods.

#### 1.2 Data Analytics in Business (90 hrs. / 12 ECTS)

This course prepares students to understand business analytics and become leaders in these areas in business organizations. This course teaches the scientific process of transforming data into insights for making better business decisions. It covers the methodologies, issues, and challenges related to analysing business data. At the beginning of the course, students will learn the foundation of analytical tools. The students will apply them to analyse real-world issues. The course will conclude with the challenges and opportunities of the big data.

#### 1.3 Data Integration for Business Intelligence (45 hrs. / 6 ECTS)

Data integration is the collection of processes, tools and strategies by which fragmented data sources are aligned to support business goals. Data integration covers topics such as data preparation, data transformation and data modelling, it is the synonym to ETL (Extract-Transform-Load) process, which is an important part of data preparation to automated analysis. From technical point of view this subject covers: introduction to relational databases and SQL language, computer ETL tools (e.g. Pentaho, Tableau Prep), data cleaning tools and techniques (e.g. Open Refine).

#### 1.4 Programming for Data Science (45 hrs. / 9 ECTS)

Using IT tools for automatic data processing is one of the most wanted skill of data scientist. R is one of most popular language in Data Science and also one of the top programming language, that is mentioned in almost all job offers for business analytics. R package has also the great community of users that will help the young analyst to solve complex problems. During the course students will learn: R fundamentals, basic syntax, data types and structures, conditional instructions and loops, basic statistical measures, graphics creation and web scrapping with R. The outcome of the course will be the development of the analytical project basing on data stored in open data sets or provided by social media.

#### 1.5 Advanced Data Visualization (45 hrs. / 7 ECTS)

The increasing number of data sources and growing data sets provide huge opportunities to use such data and its analysis to solve many problems. In this course students will learn how to become

a professional at communicating business-relevant implications of data analyses and how to make effective visualizations using most popular visualization software (e.g. Tableau, Power BI). Using examples from real world business cases we will demonstrate best practices for data visualization and data storytelling. This subject includes data visualization techniques, data analysis methods and some data mining concepts. At the end of the capstone project, the students are expected to gather data, create data models and design visual dashboards and data story (analytical decision support system), which can be used in real managerial processes.

## 1.6 Marketing Analytics and Decision Making (30 hrs. / 4 ECTS)

This subject introduces the theories and techniques of marketing analytics in the context of various marketing decision-making environments. Since businesses problems often have too many alternative solutions, you will learn how optimization can help you identify the best option. In this course, you discover how to develop quantitative models that leverage business data, statistical computation to forecast sales and marketing impact for customer relationship management, market segmentation, value creation, communication. It places particular emphasis on the application of statistical and computational models to decision-making issues likely to confront marketing managers today and in the future.

## 1.7 Communication and Presentation Skill (15 hrs. / 3 ECTS)

Effective communication & presentation skills are key to business success. In this course students learn how to improve their verbal and non-verbal communication skills to build relationships and establish trust. Students also learn how to prepare and make engaging presentations.

## 1.8 Agile Project Management (30 hrs. / 5 ECTS)

This subject will explore the history and philosophy of Agile project management, including the Scrum framework. Students will learn more about Scrum, exploring its pillars and values and relating essential Scrum team roles. During this course we will show how to manage and refine a product backlog, develop User Stories, implement Agile's value-driven delivery strategies, and how to effectively organize the five important Scrum events for a Scrum Team. The "Agile Project Management" course also includes instructions on most popular software tools (e.g. Redmine, Atlassian Jira) to apply hands-on learning of the topics.

## 1.9 Data-driven Business Management Project (45 hrs. / 10 ECTS)

Data-driven approach allows the managers to analyse the phenomena objectively and make the rational decisions, therefore the course is the case study of utilizing data sources and data transformation into managerial information. During the course students working in groups will develop the analytical project, that will consist of the following steps: formulating the aims, collecting the data from the open repositories, processing and visualization of the results with the known tools like R package, Business Intelligence tools, Graphical Information systems (GIS) in order to solve the Business problem and translate data into decisions. Finally, basing on the analysed data the final conclusions and recommendations will be developed.

# 2 Elective courses

## 2.1 Machine Learning OR NoSQL Databases (45 hrs. / 5 ECTS)

**Machine Learning** course is intended to help students understand the capabilities, challenges, and consequences of automated learning and prepare them to participate in the artificial intelligence (AI) technology. This course will introduce students through a series of practical case studies to the

main areas of machine learning including prediction, classification and clustering. Students will learn how to analyse complex datasets, create systems that adapt over time, and how to design intelligent applications which can be used for predictions from data sets.

**NoSQL Databases** course covers the main features of typical NoSQL databases and basic techniques for implementing NoSQL databases. In this course, students will learn: how to identify the type of NoSQL databases according to business requirements, how to apply NoSQL data modelling, and how to optimize query processing with NoSQL technology.

## 2.2 Robotic Process Automation OR Advanced Programming (45 hrs. / 10 ECTS)

**Robotic Process Automation** (RPA) is an automation technology based on bots or AI digital workers, used for processing the data without manual intervention. Those tools watch the task performed by the user in order to automatically repeat it in the application's graphical user interface. The aim of RPA is to increase the productivity and the effectiveness, therefore soon many jobs can be automated (human labour will be replaced). During the course students will learn: RPA fundamentals, how to use the popular RPA tool, basics of business process flow and modelling, bots creation, case and error management. At the end of the course students will develop the RPA project. The project will consist of the following steps: selecting the right process for automation, mapping the business process, choosing the tool for automation, building the digital worker, testing the bot performance.

**Advanced programming** course is an advanced level module in which some advanced programming concepts in Python are taught. Python is one of the most popular programming language because, its easy and flexible, has a very strong community, offers hundreds of libraries and frameworks and is an effective tool for big data analysis and machine learning. After completing this module, the students will be able to prepare frameworks for automated data analysis and write sophisticated Python applications. The topics covered by the course cover installation and configuration of the environment; introduction to Python and its libraries; variables and data structures; data reading and writing; basic calculations; symbolic calculations; data processing and statistics; charts and visualizations. The main outcome will be the analytical project basing on the data stored in open data sets or provided by social media.

## 2.3 Qualitative data analysis OR Explainable Artificial Intelligence (45 hrs. / 8 ECTS)

**Qualitative data analysis.** This course will introduce students to techniques, tools, and frameworks for qualitative data analysis. Qualitative research often involves an iterative process. We will focus on the ingredients required for this process: data collection and analysis. Students will learn about methods for analysing qualitative data with the aid of software.

**Explainable Artificial Intelligence.** Artificial Intelligence (AI) allow problem solving automatically without any human intervention and is very successful. This course gives an introduction to Explainable AI, providing an overview of relevant concepts such as interpretability, transparency and black-box machine learning methods. The course provides an overview of state-of-the-art methods for generating explanations, and touches upon issues related to decision-support, human interaction with AI/intelligent systems and their evaluation.

See: <https://rekrutacja.umcs.pl/en/ma-data-science/> for more details.