COURSE DESCRIPTION CARD - SYLLABUS

Course name
Big Data in Management [N1Log2>BDwZ]

Course

Field of study
Logistics

Year/Semester
3/5

Area of study (specialization)
–

Profile of study
general academic

Level of study
first-cycle

Course offered in
Polish

Form of study
part-time

Requirements
compulsory

Number of hours

Lecture
8

Laboratory classes
8

Other
0

Tutorials
8

Projects/seminars
0

Number of credit points

4,00

Coordinators

dr inż. Katarzyna Ragin-Skorecka
katarzyna.ragin-skorecka@put.poznan.pl

Lecturers

Prerequisites

Basic knowledge of computer science and statistics.

Course objective

Providing students with basic knowledge in the field of organization, management and processing of large data sets. Developing students' problem-solving skills related to the organization, management and processing of Big Data.

Course-related learning outcomes

Knowledge:
1. Student knows the basic concepts of logistics and its detailed issues and supply chain management With Big Data support. [P6S_WG_05]
2. Student knows the best practices in logistics and its specific issues using Big Data. [P6S_WK_06]

Skills:
1. Student can design, using appropriate methods and techniques, an object, system or process that meets the requirements of logistics and its specific issues and supply chain management with Big Data

usage in enterprise. [P6S_UW_07]
2. Student is able to present, using properly selected means, the problem within the scope of logistics and its specific issues, and supply chain management using Big Data. [P6S_UK_01]

Social competences:
1. Student is able to plan and manage in an entrepreneurial manner in Big Data supporting company. [P6S_KO_01]
2. Student is aware of cooperation and group work on solving problems within logistics and supply chain management with Big Data usage in company. [P6S_KR_02]
3. Student is aware of the responsible filling, correct identification and resolution of dilemmas related to the profession of logistics with usage of Big Data. [P6S_KR_01]

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Lecture: Knowledge acquired during the lecture is verified by completing one problem-solving task and a final test, pass mark: 50% of points.
Exercises: During the exercises, students work in groups on specific tasks that are assessed. The final grade is based on the sum of points received for the tasks.
Laboratory: During the laboratory, students work in groups on specific tasks that are assessed. The final grade is based on the sum of points received for the tasks.

Programme content

Lecture: Introduction to Big Data systems, motivations, definitions, problems in the world of Big Data. Hadoop platform, distributed file systems on the example of HDFS, task scheduling systems in Big Data systems on the example of YARN, data batch processing engines on the example of MapReduce, MapReduce processing optimization techniques, decomposition of complex problems into MapReduce action sequences, Hadoop Streaming. Relational data processing using Spark SQL, DataFrame and Dataset data types, data processing in Spark SQL, processing optimization mechanisms. The use of Big Data in Logistics.
Tutorial: Introduction to Big Data systems, motivations, definitions, problems in the world of Big Data. Hadoop platform, distributed file systems on the example of HDFS, task scheduling systems in Big Data systems on the example of YARN, data batch processing engines on the example of MapReduce, MapReduce processing optimization techniques, decomposition of complex problems into MapReduce action sequences, Hadoop Streaming. Relational data processing using Spark SQL, DataFrame and Dataset data types, data processing in Spark SQL, processing optimization mechanisms. The use of Big Data in Logistics.
Laboratory: Introduction to Big Data systems, motivations, definitions, problems in the world of Big Data. Hadoop platform, distributed file systems on the example of HDFS, task scheduling systems in Big Data systems on the example of YARN, data batch processing engines on the example of MapReduce, MapReduce processing optimization techniques, decomposition of complex problems into MapReduce action sequences, Hadoop Streaming. Relational data processing using Spark SQL, DataFrame and Dataset data types, data processing in Spark SQL, processing optimization mechanisms. The use of Big Data in Logistics.

Course topics

Lecture: introduction to big data, cloud computing, database management systems, processing and analytics of large data sets, use of big data in logistics.
Exercises: data analysis in logistics using Business Intelligence tools (e.g. Power BI).
Laboratory: designing and automating data collection, processing and management processes using modern tools and platform integrations (e.g. Tally.so, Make.com, Airtable, OpenAI API).

Teaching methods

Lecture: multimedia presentation illustrated with examples given on the board, discussion and problem analysis.
Tutorial: problem solving, discussion, team work.
Laboratory: problem solving, discussion, team work.

Bibliography

Basic:
1. Marz N., Warren J., Big Data. Principles and best practices of scalable realtime data systems, Manning Pubications Co., 2015.
2. Zaharia M., Chambers B., Spark: The Definitive Guide, O'Reilly Media, 2018.
3. Odersky M., Spoon L., Venners B., Programming in Scala, 3rd edition, Artima Inc, 2016.
4. Rajaraman A., Ullman J.D., Mining of Massive Datasets, Cambridge University Press, 2012 (http://infolab.stanford.edu/~ullman/mmds.html).

Additional:
1. Horstmann C., Scala for the Impatient, Addison-Wesley, 2016.

Breakdown of average student's workload

	Hours	ECTS
Total workload	100	4,00
Classes requiring direct contact with the teacher	24	1,50
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	76	2,50