QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department

Study plan No.	**2024/2025**		**University Specialization**		Data science and Artificial Intelligence
Course No.	0135321		**Course name**		Big Data
Credit Hours	**3**		**Prerequisite Co-requisite**		Databases
Course type	√ MANDATORY UNIVERSITY REQUIREMENT	☐ **UNIVERSITY ELECTIVE REQUIREMENTS**	☐ **FACULTY MANDATORY REQUIREMENT**	☐ **Support course family requirements**	☐ **Mandatory requirements** / Elective requirements
Teaching style	☐ **Full online learning**		☐ **Blended learning**		☐ **Traditional learning**
Teaching model	☐ **2Synchronous: 1asynchronous**		☐ **2 face to face : 1synchronous**		☐ **3 Traditional**

Faculty member and study divisions information (to be filled in each semester by the subject instructor)

Name	Academic rank	Office No.	Phone No.	E-mail	
Dr. Dara Aqel	**Associate professor**			**d.aqel@zuj.edu.jo**	
Division number	**Time**	**Place**	**Number of students**	**Teaching style**	**Approved model**

Brief description

of big data including collecting, analyzing, processing, storing, and managing large amounts of data. The students will be able to apply advanced analytical tools to extract meaningful patterns from large and complex datasets. They will also utilize big data technologies and platforms such as Hadoop and Spark. This course presents an introduction to data warehouse, types of data warehouses, ETL, star architecture, snow flake architecture, implementing data warehouse using SQL, introduction to big data, OLAP vs RTAP, Map Reduce, Hadoop, Spark, Machine learning using spark, Streamline Data Ingestion using AWS, Hive, NoSQL databases.

Learning resources

Course book information (Title, author, date of issue, publisher ... etc)	1- Big Data: Concepts, Technology, and Architecture. Balamurugan Balusamy, Nandhini Abirami R, Seifedine Kadry, Amir Gandomi. John Wiley & Sons. 368 Pages, 2021.			
Supportive learning resources (Books, databases, periodicals, software, applications, others)	**1.** Data Mining, Concepts and Techniques, Jiawei Han, Jian Pei, Hanghang Tong, 4rd edition, Elsevier Science, 2022. **2.** Big Data Concepts, Technologies, and Applications. Husain, Mohammad Shahid, Mohammad Zunnun Khan, and Tamanna Siddiqui, Auerbach Publications. 2023. **3.** Big data: concepts, warehousing, and analytics. Santos, M.Y. and Costa, C., CRC Press, 2022. **4.** Data Science from Scratch: First Principles with Python. By Joel Grus. O'REILLY, 2nd Edition, 2019.			
Supporting websites				
The physical environment for	☐ Class	☐ labs	☐ Virtual	☐ Others

QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department

teaching	room		educational platform	
Necessary equipment and software				
Supporting people with special needs				
For technical support				

Course learning outcomes (S= *Skills, C= Competences K= Knowledge,)*

No.	Course learning outcomes	The associated program learning output code
	Knowledge	
K1	To show excellent knowledge in the basic data warehouse and big data topics	**MK2**
K2	To be acquainted with the basics of various advanced data warehouse and big data topics.	MK2
	Skills	
S1	To be able to apply data warehouse concepts on a real case scenario using SQL	**MS1**
S2	To be able to use Spark for machine learning of a big data	**MS1**
	Competences	
C1	To apply the various concepts of data warehouse and big data in solving real life problems	**MC1**

Mechanisms for direct evaluation of learning outcomes

Type of assessment / learning style	Fully electronic learning	Blended learning	Traditional Learning (Theory Learning)	Traditional Learning (Practical Learning)
Project based learning (PBL)	0	0	5%	**10%**
Second / midterm exam	%30	%30	30%	**30%**
Participation / practical applications/Assignments	0	0	25%	**20%**
Asynchronous interactive activities	%30	%30	0	**0**
final exam	%40	%40	40%	**40%**

Note: Asynchronous interactive activities are activities, tasks, projects, assignments, research, studies, projects, work within student groups ... etc, which the student carries out on his own, through the virtual platform without a direct encounter with the subject teacher.

Schedule of simultaneous / face-to-face encounters and their topics

Week	Subject	learning style*	Reference **
1	Introduction to Data Warehouse	Lecture	**T:1-13**
2	Data Warehouse Types	Lecture	**T:111-128**
3	ETL	Lecture	**R1: 84-95**
4	Data Warehouse Architectures	Lecture	**R1: 95-110**

5	Implementing Data Warehouse	Lecture	**R1: 111-117**
6	**Mid Exam Estimated + Revision**	learning through problem solving	
7	Case Study 1: Data Warehouse implementation using SQL	learning through problem solving	**Handouts**
8	Introduction to Big Data	Lecture	**T:43-52**
9	Map Reduce	Lecture	**R1: 125-149**
10	Hadoop and Spark	Lecture	**T:55-99**
11	Spark Machine Learning	Lecture	**Handouts**
12	NoSQL databases	Lecture	**Handouts**
13	Streamline Data Ingestion	Lecture	**Handouts**
14	Case Study 2: Spark Machine Learning using pyspark + Case Study 3: Querying Unstructured data using Hadoop Hive	learning through problem solving	**Handouts**
15	Project based learning (PBL): Projects discussions	learning through projects	**Handouts**
16	**Final Exam**		

*** Learning styles: Lecture, flipped learning, learning through projects, learning through problem solving, participatory learning ... etc.**
**** Reference: Pages in a book, database, recorded lecture, content on the e-learning platform, video, website ... etc.**

Schedule of asynchronous interactive activities (in the case of e-learning and blended learning)

Week	Task / activity	Reference	Expected results
1	Introduction to Data Warehouse		To show a good comprehension of the data warehousing basic concepts
2	Data Warehouse Types		To distinguish among different types of data warehouses
3	ETL		To be acquainted of basics of ETL concepts
4	Data Warehouse Architectures		To distinguish among different architectures of data warehouses
5	Implementing Data Warehouse		To implement a data warehouse using SQL
6	Mid Exam Estimated + Revision		
7	Case Study 1: Data Warehouse implementation using SQL		To apply data warehouse concepts
8	Introduction to Big Data		To show a good comprehension of the big data basic concepts
9	Map Reduce		To show a good comprehension of Map Reduce

10	Hadoop and Spark		To show a good comprehension of the Hadoop and Spark
11	Spark Machine Learning		To apply Spark to handle big data using python
12	NoSQL databases		To be acquainted with basics of NoSQL
13	Streamline Data Ingestion		To have basic understanding of data ingestion
14	Case Study 2: Spark Machine Learning using pyspark + Case Study 3: Querying Unstructured data using Hadoop Hive		To apply Spark to handle big data using python
15	Project based learning (PBL): Projects discussions		To apply Hive to query NoSql databases
16	**Final Exam**		