Course Specification
(Bachelor)

Course Title:	BIG DATA ANALYTICS
Course Code:	ITEC 4**(416) - Elective-3
Program:	Bachelor in Information Technology (BIT)
Department:	: Computer Science
College:	College Of Engineering & Computer Science
Institution:	Jazan University
Version:	1
Last Revision Date:	Aug. 18, 2024

Table of Contents

A. General information about the course:

1. Course Identification

1. Credit hours: (3)

2. Course type

A.	☐University	☐College	☒ Department	☐Track	☐Others
B.	☐Required		☒ Elective		

3. Level/year at which this course is offered: (Level 8 / 4th Year)

4. Course General Description:

Big Data is the term for a collection of datasets so large and complex that they become difficult to process using on-hand database management tools or traditional data processing applications. The challenges include capture, storage, search, sharing, transfer, analysis, and visualization. This Big Data Analytics course will first introduce the overview applications, market trend, and the things to learn. Then, it will introduce the fundamental platforms, such as Hadoop, Spark, and other tools, such as IBM System G for Linked Big Data. Afterwards, the course will introduce several data storage methods and how to upload, distribute, and process them. This will include HDFS, HBase, KV stores, document database, and graph database. The course will go on to introduce different ways of handling analytics algorithms on different platforms. Then, it will introduce visualization issues and mobile issues on Big Data Analytics. Students will then have fundamental knowledge on big data analytics core concepts, technologies and techniques so they can gain the necessary skills needed to design highly scalable data analytics systems and to handle various real-world challenges.

5. Pre-requirements for this course (if any):

ITEC-313, ITEC-314

6. Co-requisites for this course (if any):

7. Course Main Objective(s):

This course will develop the students' ability to learn:

- Describe developed knowledge, skills and understanding around a range of capabilities and benefits of big data.
- Handle analytics algorithms on different platforms to big data.
- Learn techniques about uploading, distributing, storing and processing large amount of big data (structured and unstructured data)
- Understand big data management and their technologies to design and build big data applications through highly scalable systems including visualization issues and mobile issues.
- Apply concepts and principles to handle, analyze and interpret various real-world challenges on Big Data Analytics using analytics and computer-based techniques.

- Demonstrate how to effectively interpret and communicate their ideas through reports/presentations to exhibit the ability to function in a group as a team member.

2. Teaching mode (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	Traditional classroom	64	100
2	E-learning	--	--
3	Hybrid • Traditional classroom • E-learning	--	--
4	Distance learning	--	--

3. Contact Hours (based on the academic semester)

No	Activity	Contact Hours
1.	**Lectures**	30
2.	**Laboratory/Studio**	30
3.	**Field**	--
4.	**Tutorial**	--
5.	**Others (Revision + Final Lab or Case-Study Exam)**	04
Total		64

B. Course Learning Outcomes (CLOs), Teaching Strategies and Assessment Methods

Code	Course Learning Outcomes	Code of PLOs aligned with the program	Teaching Strategies	Assessment Methods
1.0	**Knowledge and understanding**			
1.1	**Define** the basic concepts and terminologies of Big data analytics process	**K1**	Visual & Verbal [Lectures / Presentations]	Mid Term, Final Exam, Assignment
1.2	**Explain** the critical methods, techniques and algorithms	**K1**	Visual & Verbal [Lectures / Presentations]	Mid Term, Final Exam, Assignment

Code	Course Learning Outcomes	Code of PLOs aligned with the program	Teaching Strategies	Assessment Methods
	commonly used in different platforms of big data			
2.0	**Skills**			
2.1	**Demonstrate** proficiency with the methods and techniques for uploading, distributing, storing and processing large amount of big data	**S1**	Visual & Verbal [Lectures / Presentations]	Mid Term, Assignment, Final Exam
2.2	**Construct** and **build** big data applications through highly scalable systems including visualization and statistical modeling tools	**S3**	Visual & Verbal [Lectures / Lab Practical]	Mini Project, Lab Exam, Final Exam
2.3	**Apply** concepts and principles to handle, analyze and interpret various real-world challenges on Big Data Analytics using analytics and statistical modelling	**S3**	Visual & Verbal [Lectures / Lab Practical]	Mini Project, Lab Exam, Final Exam
3.0	**Values, autonomy, and responsibility**			
3.1	**Evaluate** and **Determine** appropriate solutions for a problem and communicate through reports/presentations to **Demonstrate** the ability to function in a group and **Elaborate** as a team member to attain a common assignment.	**V3**	Self-study, Visual & Verbal [Lectures / Lab Practical]	Lab Exam, Mini Project

C. Course Content

No	List of Topics	Contact Hours
1.	***What is Big data?*** ***Why is big data analytics important?*** ***Concepts*** ***Data Analysis*** ***Data analytics*** ***General categories of analytics:*** ***Business intelligence(BI)*** ***Big Data Characteristics: 5 V's-*** ***Different types of data:***	4T + 4P

	Metadata *Summary* *Sample Questions*	
2.	**Chapter 2: *Business Motivations & Drivers for Big Data Adoption*** **Business Architecture (BA)** **Drivers of Big Data** **Data evolution and the rise of Big Data sources** **Multiple Sources of Big data** **Big Data adoption in businesses:** **Information and Communications Technology(ICT)** **Factors responsible for Big Data adoption in businesses** **Big Data Adoption and Planning Considerations** **Big Data Analytics Lifecycle** ***Summary*** ***Sample Questions*** **Mid Term Exam (ch-1 & ch-2):15 Marks**	4T + 4P
3.	**Chapter 3 :Enterprise Technologies and Big Data Business Intelligence :: Technology Foundations for Big Data** **Online Transaction Processing (OLTP)** **Online Analytical Processing (OLAP)** **Extract Transform Load (ETL)** **Data Warehouses** **Data Marts** **Traditional BI** **Big Data BI** **Data Visualization** **Types of Visual Analysis** **Summary** **Sample Questions**	4T + 4P
4.	**Chapter 4: BIG DATA STORAGE CONCEPTS** **Clusters** **File Systems and Distributed File Systems** **NoSQL** **Sharding** **Replication** **Sharding and Replication** **CAP Theorem** **ACID** **BASE** **Summary** ***Sample Questions***	4T + 4P
5.	**Chapter 5 :*Big Data Processing Concepts*** **PART-1:** **Parallel Data Processing**	4T + 4P

	Distributed Data Processing **Hadoop** **Processing Workloads** **Cluster** **Processing in Batch Mode** **Processing in Realtime Mode** **PART-2:** **Batch Processing:** **Hadoop and MapReduce** **Hadoop - MapReduce Examples** **Pig** **Summary** ***Sample Questions***	
6.	***Chapter-6:: Big Data storage Technology -Real Time Analysis :: In-Memory Processing*** ➢ **Apache spark -Spark architecture and Spark operations** ➢ **No SQL- Non-relational databases** ➢ **Key-value databases** ➢ **Document databases** ➢ **Column family databases-HBASE** ➢ **HBASE:** ▪ **Data Model, Table Structure, Key Format, Features, Architecture,** ▪ **HBASE Storage Structure** ➢ **Graph databases** ➢ **Summary** ***Sample Questions***	4T+4P
	Revision/Lab	2T + 2P
Total		30T + 30P

D. Students Assessment Activities

No	Assessment Activities *	Assessment timing (in week no)	Percentage of Total Assessment Score
1.	Mini Project	3rd Week	15%
2.	Midterm Exam (Chapter-1,2) **[As scheduled]**	6th / 7th Week	15%
3.	Assignment	8th Week	10%
4.	Final Lab Exam (Internal assessment:10Marks +Final lab:10 Marks)	13th/14th Week	20%
5.	Final Exam-All Chapters **[As scheduled]**	15th Week	40%

*Assessment Activities (i.e., Written test, oral test, oral presentation, group project, essay, etc.).

E. Learning Resources and Facilities

1. References and Learning Resources

Essential References	• Th. Erl, W. Khattak and P. Buhler," Big Data Fundamentals: Concepts, Drivers & Techniques", Pearson Publisher, ISBN: 9780134291079, 2016. • Big Data Analytics: A Hands-On Approach 2019 by Arshdeep Bahga & Vijay Madisetti, Published by Arshdeep Bahga & Vijay Madisetti,ISBN: 978-1-949978-00-1 • Big Data For Dummies® Published by John Wiley & Sons, Inc. 111 River Street Hoboken, NJ 07030-577
Supportive References	• In addition, material from "Data Science and Big Data Analytics Student Guide" distributed by EMC Education Services will be provided to the students. • Ivanka Menken, "Big Data Complete Certification Kit", Core Series for IT, 2013. • Foster Provost and Tom Fawcett, "Data Science for Business: What You Need to Know about Data Mining and Data-analytic Thinking" ©2013 | ISBN 1449361323. • Demirbaga, Ümit, Gagangeet Singh Aujla, Anish Jindal, and Oğuzhan Kalyon. Big data analytics: theory, techniques, platforms, and applications. Springer Nature, 2023. • Big-data-analytics-for-beginners-Mastering the Art of data driven decision making, sam-campbell.
Electronic Materials	https://www.datacamp.com/courses/ https://lms.jazanu.edu.sa/webapps/login
Other Learning Materials	**Text books:** • An introduction to Data Science by *Jeffrey Stanton* • The Elements of Data Analytic Style by *Jeff Leek* • Exploratory Data Analysis with R, by *Roger Peng* • OpenIntro Statistics, by *Diez, Barr*, and *Centinkaya-Rundel* • R Programming for Data Science, by *Roger Peng* • Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem, Douglas Eadline, Addison-Wesley, Pearson Education India; First edition (1 March 2016) • https://intro2r.com/r-style-guide.html **Data Resources:** • Variety of consumer datasets https://www.kaggle.com/datasets • World Bank https://data.worldbank.org/data-catalog/ • US Government Data https://www.data.gov/

2. Required Facilities and equipment

Items	Resources
facilities (Classrooms, laboratories, exhibition rooms, simulation rooms, etc.)	30 working PC with internet connection Working Projector, Smart Board Teacher's chair and desktop PC with internet connection
Technology equipment (projector, smart board, software)	IDEs: Google Colab / Pycharm/Jupiter notebook IDE: R Studio The Comprehensive R Archive Network (r-project.org) Libraries: DataMaid, DataExplorer, and SmartEDA tidyverse etc.
Other equipment (depending on the nature of the specialty)	--

F. Assessment of Course Quality

Assessment Areas/Issues	Assessor	Assessment Methods
Effectiveness of teaching	Students	Indirect(Course Evaluation Form)
Effectiveness of Students assessment	PL/QAU/HOD	**Direct(Course Report/Result Analysis)**
Quality of learning resources	CEO / Track Leaders/PL	Indirect(Review, meetings and star rating with suggestions for further modifications and improvements)
The extent to which CLOs have been achieved	PL/HOD / QAU	Direct(CLo assessment template further verified at course coordinator, track leader and QAU level)
Other		

Assessors (Students, Faculty, Program Leaders, Peer Reviewers, Others (specify)

Assessment Methods (Direct, Indirect)

G. Specification Approval

COUNCIL /COMMITTEE	**DEPARTMENT COUNCIL**
REFERENCE NO.	**ENGCSCS2406**
DATE	**19/09/2024**