JINDAL GLOBAL BUSINESS SCHOOL
INDIA'S FIRST MULTI-DISCIPLINARY GLOBAL BUSINESS SCHOOL

O.P. Jindal Global University
A Private University Promoting Public Service
NAAC Accreditation - 'A' Grade

Jindal Global Business School
Course Outline

Course Title	Big Data Analytics - BS-ISA-6016
Core or Elective	Elective
Program and Batch	MBA-2025, IBM-2022, IBM-2023
Semester & Academic Year	Fall 2026
Credits	1.5
Discipline/Area	IS & Analytics
Name of the Faculty Member/Course Instructor	Prof. Mohit Bhatnagar
Contact Details of the Faculty Member	mohit.bhatnagar@jgu.edu.in
Contact Details of Support Staff	jgbs-eo@jgu.edu.in
Faculty Member's Open Office Day/s & Time	TBA

Introduction to the Course

Be it social media platforms, Internet of Things (IoT) Sensors, fitness applications, server log's one thing is common, that they are all generating data with great volume, velocity and variety. This type of data is ubiquitous and is what is referred as Big Data. This is not an overhead anymore but is what companies are mining for insights and creating competitive advantage. For managers and entrepreneurs, it is now required and essential to be aware of tools and technologies that can help harness Big Data.

In this hands-on course we will gain an understanding of what the emerging big data applications are and the nature of analytical insights that can be drawn from them. This course will also provide basic grounding in use of Big Data tools and techniques, including Hadoop and other ecosystem tools, as well as in the ways of storing information that allow for efficient processing. Hand on experience is provided on both a simulated single node installation and a cloud based multi node HADOOP cluster setup. Programming skills are not a prerequisite but interest in analytics and technology is an essential. The course is also an introductory course for picking up basic hands-on experience in Linux, SQL & Python.

Course Learning Objectives

At the end of the course, students should be able to

1. CLO1- Understand Big Data application and its importance and need in business context

2. CLO2- Understand the basic concepts of distributed storage and computing using HADOOP

3. CLO3- Familiarization with the tools for querying & drawing analytical insights from Big Data

4. CLO4- Understand the design and working of streaming Big Data applications

Programme Competency Goals

MBA Programme Competency Goals (PCGs)		**MBA Programme Learning Objectives (PLOs)**
		Students will be able to
1	**Technological Agility:** Ability to adopt relevant technologies for better business decision making.	1. Understand relevant business technologies
		2. Understand future technologies in business domain
2	**Responsible Global Citizenship:** Ability to understand the interplay between local and global issues and to act with sensitivity towards ethical and social issues	3. Understand the interplay between local and global business issues
		4. Demonstrate sensitivity towards ethical issues
		5. Demonstrate sensitivity towards social issues
		6. Address societal issues
3	**Effective communication:** Ability to effectively exchange ideas and information	7. Present their ideas with clarity
		8. Prepare an organized and logical business document
		9. Use technology for effective communication
4	**Critical Thinking:** Ability to identify, analyze business problems and propose effective solutions	10. Identify main issues of business problems
		11. Examine information from different sources
		12. Draw inferences from analysis
		13. Evaluate alternatives
		14. Summarize and conclude
5	**Leadership:** Ability to take initiative, inspire and collaborate with others	15. Take initiative
		16. Contribute effectively in groups

PLO-PCG Assessments Mapping Matrix

Program Learning Objectives (PLOs)	Program Competency Goals (PCGs)	Course Assessment Item
This course helps you to develop the following Program Learning Outcomes:	This course helps you to develop the following Program Competency Goals:	This learning outcome will be assessed in the following items
PLO1, PLO2, PLO3, PLO4, PLO5, PLO6, PLO7, PLO8, PLO9	PCG1, PCG2, PCG3	A1, A2, A4,A5
PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14	PCG1, PCG4	A1, A2, A4
PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16	PCG1, PCG4, PCG5	A1, A2, A4
PLO1, PLO2, PLO10, PLO11, PLO12, PLO15, PLO16	PCG1, PCG4, PCG5	A1, A2, A3, A4
PLO1, PLO2, PLO3, PLO4, PLO5, PLO6, PLO7, PLO8, PLO9	PCG1, PCG2, PCG3	A1, A2, A4

Evaluation Schema

The course grade will be determined based on:

Assessment Task	Weightage (Percentage)	Nature (Individual/Group)	Week of Assessment	PLOs to be Assessed
A1: Quiz 1, 2	20%	Individual	Ongoing	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14
A2: Individual Presentations	20%	Individual	7th Session	PLO1, PLO2, PLO3, PLO4, PLO5, PLO6, PLO7, PLO10, PLO18, PLO19
A3: Project Presentation	20%	Group	15th Session	PLO1, PLO2, PLO7, PLO8, PLO9, PLO10, PLO11,

				PLO12, PLO13, PLO14, PLO15, PLO16, PLO17, PLO18, PLO19
A4: End term Examination	30%	Individual	In the JGU Examination period/week	PLO1, PLO2, PLO7, PLO8, PLO10, PLO11, PLO12, PLO13, PLO14, PLO17, PLO18, PLO19
A5: Class Participation	10%	Individual	Ongoing	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14

Description of Assessments:

A1 – Quiz – These will be multi choice questions and used to assess students' ability to understand conceptually and syntactically the critical concepts discussed in the class.

A2 – Individual Presentation – This will be done on an ongoing basis where each student will be assigned a case which he /she has to present to the class. The evaluation will be done on the coverage , research and presentation

A3 - Project Presentation – Identify and evaluate a business analytics issue and implement your learning. The presentation must carry a problem identification, analysis and a demo of the code implemented. A detailed note on the format for the presentation & submission would be provided separately.

A4 - End term examination – The end term examination will be of **30 marks of 1.5 hours duration**. This will be an invigilated exam according to the mode, modalities and process as decided by CoE.

A5 – Class Participation – The purpose of the component is to encourage students to take advantage of the industry interface being provided by the GLs, participate in the case discussions, and delve deeper into the topic by going beyond what is covered in the session. The idea is to have a 'practice' focused assessment in addition to the 'theoretical' assessments. Class participation, in general, has a 7.5% weightage, and participation in the GLs has a 2.5% weightage.

Rewards attendance, active discussion, and the "Troubleshooting Culture" (helping peers debug their broken flows in class).

Rubrics for Assessment

A2 Individual Presentation Rubrics

Criteria/Level	Poor (0)	Fair (1-2)	Good (3-4)	Excellent (5)
Coverage & Logical Flow	Did not cover the entire material and the flow was not haphazard and	Coverage was somewhat incomplete and organisation of	Coverage was mostly complete with organisation of ideas was mostly	Coverage was complete with no mistakes and organisation of

	disconnected	ideas was sometimes clear	clear	ideas was very clear
Analysis & Examples	Gave no examples and no analysis of concepts	Very few examples were given and mostly reading from slides with minimal analysis	Some examples were given and sometimes spoke outside of slides with analysis of concepts	Many examples given and used slides only for cues linked examples with concepts with good analysis
Communication & Presentation skills	Poor communication skills could not engage the audience	Fair communication skills could somewhat engage the audience	Good communication skills could engage audience most of the time	Excellent communication skills could engage audience all throughout the presentation

A3 Project Presentation Rubrics

Criteria/Level	Poor (0)	Fair (1-2)	Good (3-4)	Excellent (5)
Identification of the right dataset	The dataset identified has little relevance to the learning objectives or project guidelines	The dataset identified has some relevance to the learning objectives or project guidelines	The dataset identified is relevant and topical and aligns to the learning objectives and project guidelines	The dataset identified is relevant and topical and aligns to the learning objectives and project guidelines
Understanding of problem(s)	Demonstrates superficial understanding of problem(s)	Demonstrates limited understanding of problem(s)	Demonstrates deep understanding of problem(s)	Demonstrates a clear and deep understanding of an issue/problem
Quality of Analysis	No attempt to draw linkages between topic and research	Some connections drawn between topic and research with basic insights	Frequent connections drawn between topic and research showing some detail and deeper insights	Consistent insightful connections drawn between topic and research with adequate detail/ clearly explained and strong insights.
Logical Flow	No coherence or organisation of ideas	Little coherence and organisation of ideas generally clear.	Some coherence and organisation of ideas generally clear.	Coherent and clear organisations of ideas
Understanding and linking theory and concepts	Incomplete links between the problems and the	Limited links between the problems and the	Good documented links between the problems and the	Excellent documented links to the theory and

	theory.	theory.	theory.	possibly additional material read and used.

Teaching Method

The course will have a judicious mix of lectures, experiential exercises, and cases. The students will be exposed to the HADOOP ecosystem and basic concepts in Linux, SQL & Python programming. Case studies and examples during the lecture will help the students understand the concepts and how they are applicable in real world settings.

Textbook / Other Readings

Textbook:

1. Hadoop: The definitive guide, 4th Edition - Tom White, O'Reilly Publication
2. Spark in Action, Second Edition, Manning Publications.

Session Plan

Session Details	**Topics**	**PLOs Covered**
Session 1	Introduction to Big Data	PLO1, PLO2, PLO3, PLO4, PLO5, PLO6
Objective of the sessions	Learn about the Big Data characteristics	
Subtopics to be covered	• Characteristics of Big Data • Ethics of Big Data	
Readings	N/A	
Case Title & Number	N/A	
Pedagogy	Lecture	
Session 2	Big Data Applications	PLO1,PLO2, PLO3, PLO4, PLO5, PLO6
Objective of the session	Introduction to the Big Data Applications across Industry	
Subtopics to be covered	• Discussion on Big Data Applications	
Readings	Blog https://www.analyticsvidhya.com/blog/2021/05/what-is-big-data-introduction-uses-and-applications/	
Case Title & Number	N/A	
Pedagogy	Lecture	
Session 3	The HADOOP ecosystem	PLO1, PLO2
Objective of the session	Introduction to the HADOOP ecosystem	
Subtopics to be covered	• Different technologies in HADOOP • The HADOOP core stack	
Readings	Book : Hadoop: The definitive guide, Chapter 1 Refer https://hadoopecosystemtable.github.io/	
Case Title & Number	N/A	

Pedagogy	Lecture	
Session 4	Hadoop Distributed File System (HDFS)	PLO1, PLO2
Objective of the session	Introduction to HDFS	
Subtopics to be covered	• Distributed Storage • YARN Overview	
Readings	Hadoop: The definitive guide, Chapter 3	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 5	Distributed Computing using MapReduce	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	Working with MapReduce	
Subtopics to be covered	• Distributed Computing • MapReduce Paradigm	
Readings	Book : Hadoop: The definitive guide, Chapter 2	
Case Title & Number	N/A	
Pedagogy	Lecture	
Session 6	SQL Re-Cap	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	SQL Queries	
Subtopics to be covered	• Recap of SQL Key Concepts	
Readings	Tutorial https://www.w3schools.com/sql/trysql.asp?filename=trysql_op_in	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 7	Introduction to Hive	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	Querying Big Data using Hive	
Subtopics to be covered	• Introduction to Hive	
Readings	Book : Hadoop: The definitive guide, Chapter 16, 17	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 8	Introduction to PIG	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13,
Objective of the session	Scripting using PIG	
Subtopics to be covered	• Introduction to PIG	

		PLO14, PLO15, PLO16
Readings	Book : Hadoop: The definitive guide, Chapter 16	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 9	NOSQL Databases	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	Introduction to NOSQL Databases	
Subtopics to be covered	• CAP Theorem • Deciding on the Database technology • Introduction to HBASE	
Readings	Book : Hadoop: The definitive guide, Chapter 20	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 10	Spark	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	Introduction to Spark	
Subtopics to be covered	• In memory processing Spark Introduction	
Readings	Book: Spark in Action, Chapter 1, 2	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 11	Spark Streaming & Machine Learning	PLO1, PLO2, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	Introduction to Spark Streaming & Machine Learning	
Subtopics to be covered	• Spark Streaming • Spark Machine Learning	
Readings	N/A	
Case Title & Number	N/A	
Pedagogy	Lecture and Hands On	
Session 12	Guest Lecture	PLO1, PLO2
Objective of the session	Guest Lecture on Big Data	
Subtopics to be covered	• Big Data Applications in Industry – Akshay Mathur (Data Scientist , StatusNeo)	
Readings	None	
Case Title & Number	N/A	

Session 13	Project Presentations	PLO7, PLO8, PLO9, PLO10, PLO11, PLO12, PLO13, PLO14, PLO15, PLO16
Objective of the session	Applications of Learned Concepts	
Subtopics to be covered	• Group wise presentations	
Readings	None	
Case Title & Number	N/A	
Pedagogy	Class Discussion & Presentation	
Session 14	Reading & Revision Week/ Examination Week*	
Objective of the session	NA	
Subtopics to be covered	NA	
Readings	NA	
Case Title & Number	NA	
Pedagogy	NA	
Session 15	Reading & Revision Week/ Examination Week*	
Objective of the session	NA	
Subtopics to be covered	NA	
Readings	NA	
Case Title & Number	NA	
Pedagogy	NA	

*Elective Endterm Examinations may take place in the last week of classes.

Disability Support

JGU endeavours to make all its courses accessible to students. The Disability Support Committee (DSC) has identified conditions that could hinder a student's overall wellbeing. These include physical and mobility-related difficulties, visual impairment, hearing impairment, mental health conditions, and intellectual/learning difficulties, e.g., dyslexia and dyscalculia. Students with any known disability needing academic and other support are required to register with the Disability Support Committee (DSC) by following the procedure specified at **https://jgu.edu.in/disability-support-committee/**

Students who need support may register any time during the semester up until a month before the end semester exam begins. Those students who wish to continue receiving support from the previous semester, must re-register within the first month of a semester. Last-minute registrations and support might not be possible as sufficient time is required to make the arrangements for support.

The DSC maintains strict confidentiality about the identity of the student and the nature of their disability and the same is requested from faculty members and staff as well. The DSC takes a strong stance against in-class and out-of-class references made about a student's disability without their consent and disrespectful comments referring to a student's disability.

All general queries are to be addressed to **disabilitysupportcommittee@jgu.edu.in**

Disclaimer: This course outline including assessments, mode, nature and weightage of assessments, sessions, sequence of sessions and/or readings may be revised during the semester if such need arises.