PGDM / PGDM (FINTECH) Program
TRIMESTER - VI (Batch: 2023-25)
END-TERM EXAMINATION, APRIL 2025

Course Name	Big Data Analytics	Course Code	
Duration	Three Hours	Max. Marks	60

Instructions:

1. All the questions are compulsory

2. **Read all questions carefully**: Ensure that you fully understand each question before attempting to answer. Provide clear and concise responses, adhering to the specific instructions given for each section.

3. **No unauthorized materials**: The use of books, notes, or electronic devices (unless explicitly permitted) is prohibited.

Problem Statement: Big Data Lifecycle in a FinTech Lending Platform – "LoanSnap"

Background:
LoanSnap is a fast-growing digital lending startup that provides instant personal loans via a mobile app. To maintain its competitive edge, LoanSnap uses Big Data technologies to streamline operations and enhance customer experience. The platform processes thousands of applications daily, relying on automated systems for credit scoring, fraud detection, and real-time decision-making.

The company uses various sources of data for decision-making:

- Customer KYC documents (PAN, Aadhaar, salary slips)

- Bank transaction history via UPI and account aggregators

- Mobile usage patterns (with consent)

- Credit bureau reports

- In-app behavior and feedback

LoanSnap's data passes through various stages—from ingestion and cleaning to analysis, modeling, and personalized product delivery. The firm also uses predictive analytics to suggest loan top-ups and optimize repayment schedules based on user behavior.

Despite these capabilities, LoanSnap faces challenges:

- Managing inconsistent data from multiple sources

- Ensuring real-time processing without delays

- Balancing user privacy with data collection

- Integrating structured and unstructured data effectively

Q.1 Based on the above case study, explain the Big Data lifecycle in the context of LoanSnap's digital lending platform. Your answer should cover:

- Key stages from data ingestion to product delivery

- Tools or technologies that could be used at each stage

- How Big Data helps improve customer experience and risk management

Marks: 20 (CO: 01)

Q.2 Describe the role of SQL in Big Data integration with python.
How do tools like Apache Hive or Spark SQL allow querying of large datasets? Illustrate with a
sample use case from a FinTech product.   Marks:15 (CO:02)

Q.3 Big Data is often described using the 5Vs. Explain these 5Vs with examples relevant to the FinTech industry. Also, define what is meant by "scale in" and "scale out" in Big Data systems.
Lastly, explain the Master-Slave architecture used in Hadoop, including the roles of NameNode (Master) and DataNode (Slaves) in managing data storage.

Marks :15 (CO: 03 &04)   **Q.4**

Case Study:
Unified Payments Interface (UPI) processes millions of financial transactions daily across India. Explain how Big Data technologies are used to support, manage, and secure the UPI ecosystem.
Your answer should cover aspects such as:

- Real-time transaction processing

- Fraud detection and risk analysis

- User behavior and personalization

- System performance and scalability

- Integration with banking and analytics systems

Marks: 10 (CO:05)