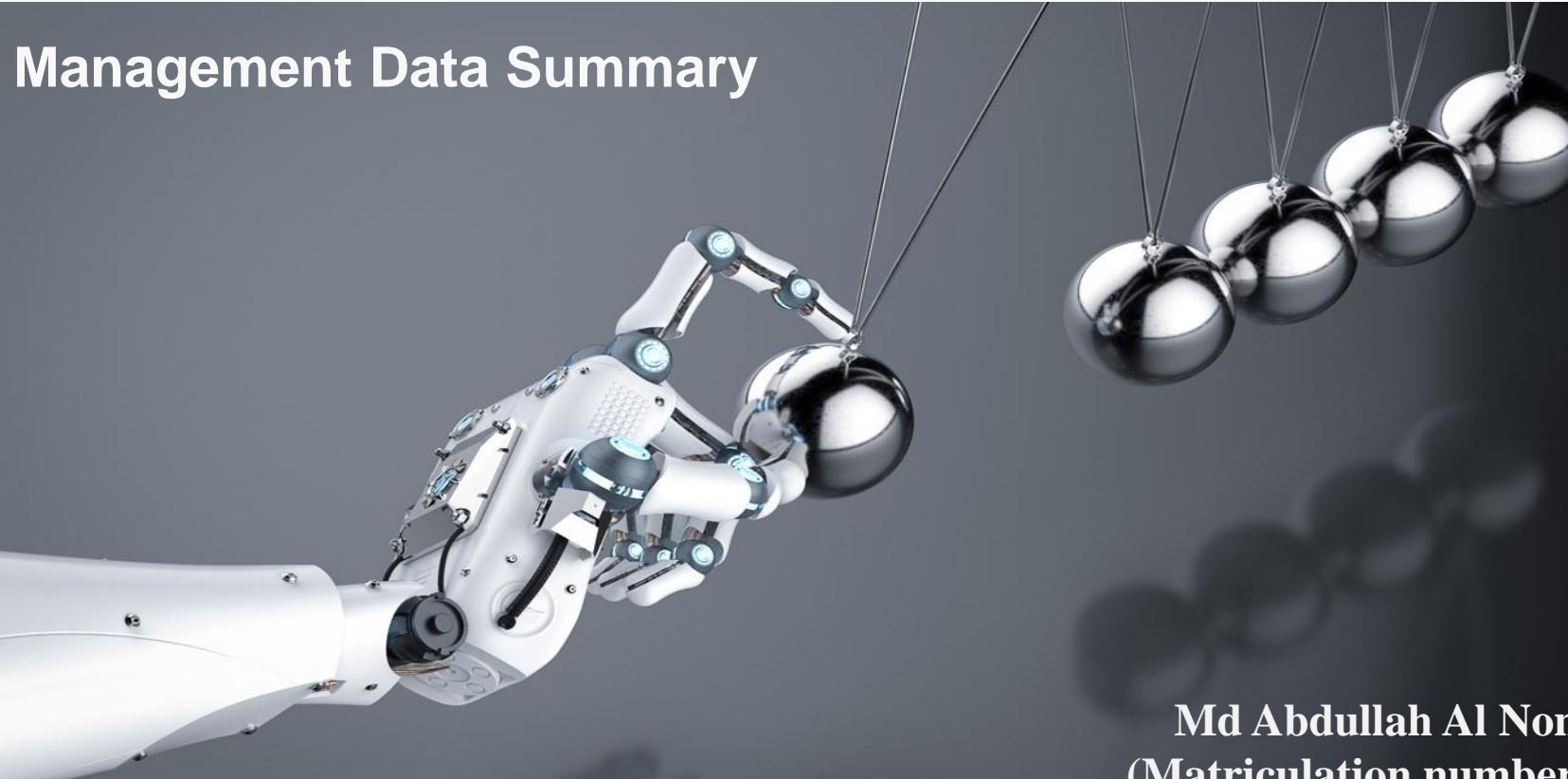


South Westphalia University of Applied Sciences

Data Science Project Seminar Driving Corporate Performance

Summer Semester 2025

Management Data Summary



Md Abdullah Al Noman Mridha
(Matriculation number: 30494592)



Cloudera Overview

Cloudera is a leading data company that provides a hybrid data platform designed for data engineering, data warehousing, machine learning, and analytics. Founded in 2008, Cloudera pioneered the development of open-source big data platforms based on Apache Hadoop



Cloudera: Enterprise Data Management



1. Cloudera offers a unified platform for large-scale data engineering, analytics, ML & Governance
2. Specializes in Hadoop-based solutions tailored for complex data workflows.
3. Cloudera Data Platform (CDP) centralizes the entire data lifecycle management
4. CDP manages data ingest, processing, analysis and governance seamlessly.



Cloudera's Data Management Features



Data Governance with SDX

SDX enforces consistent security, metadata and compliance policies for data governance.



Scalability Across Clouds

Cloudera supports petabytes of data across hybrid clouds including AWS, Azure and on-prem solutions.



Analytics and Machine Learning Tools

Tools like Cloudera Machine Learning (CML) enable self-service analytics and insights.



Robust Security Features

Cloudera offers role-based access control (RBAC), encryption and audit trails for security.

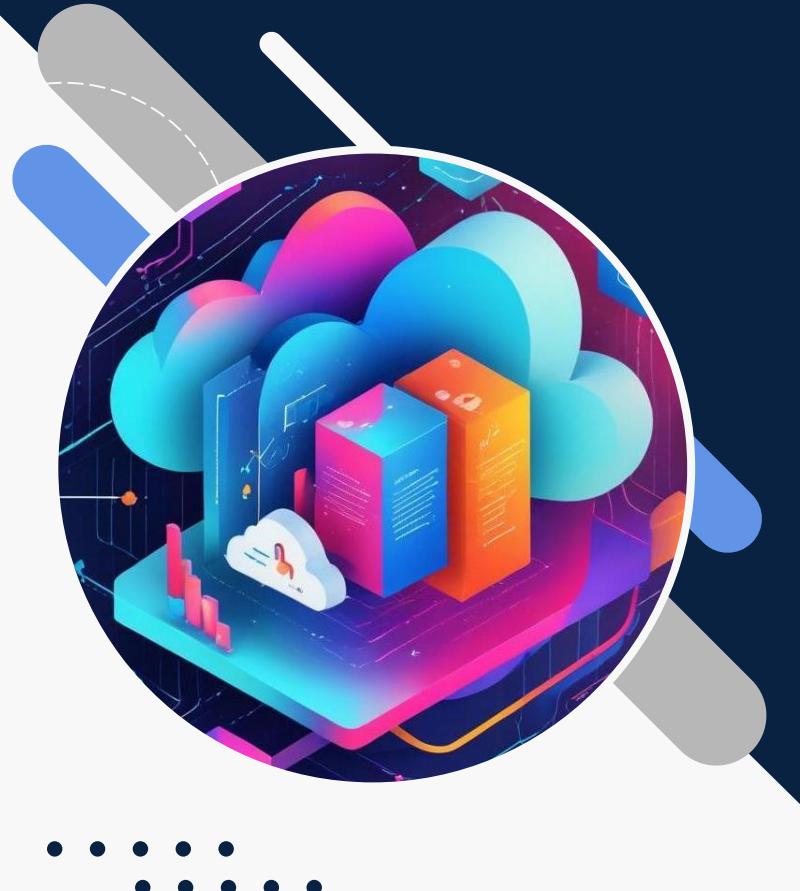




IBM Cloud

IBM Cloud Overview

IBM Cloud is a comprehensive cloud computing platform developed by IBM that offers both infrastructure-as-a-service (IaaS) and platform-as-a-service (PaaS) capabilities. It supports a wide range of services including compute, storage, artificial intelligence, analytics, DevOps, and security.



• • : : : . .

IBM Cloud: AI-Driven Data Management



Full-stack Cloud Platform

Offers a comprehensive cloud solution with AI/ML capabilities, enhancing data workflows.



AI/ML Integration

Utilizes IBM Watson for intelligent automation and data analysis, improving decision-making.



Hybrid Cloud Leadership

Leverages Red Hat OpenShift for seamless integration across multiple cloud environments.



IBM Cloud Pak for Data

A unified platform facilitating data integration, governance and AI automation for enterprises.



Target Users

Designed for enterprises seeking AI-driven insights and flexibility in hybrid/multi-cloud setups.



IBM Cloud – Key Features for Data Management

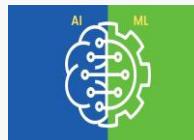
Data Integration with Watson Query

Enables federated queries across distributed data sources, enhancing data accessibility.



AI and ML Integration

Utilizes Auto AI for automated model building and Watson Studio for collaborative machine learning.



Hybrid Cloud Capabilities

Facilitates seamless data portability across public, private and on-premise environments.



Robust Security Features

Offers 'Keep Your Own Key' encryption and ensures GDPR Compliance for data protection.



IBM Cloud AI-Ready Infrastructure Foundations

IBM Cloud Foundational



Thank you!

Do you have any questions?

