Vidya Pratishthan's
Kamalnayan Bajaj Institute of Engineering and Technology

Vidyanagari, Baramati, Dist. – Pune 413133
An Autonomous Institute Approved by AICTE and affiliated to SPPU,Pune

Department of Computer Engineering

VPKBIET

Curriculum Structure and Syllabus of Honors in Cloud Computing and Virtualization Computer Engineering (Course 2023)

With effective from Academic Year 2024-25

Vidya Pratishthan's
Kamalnayan Bajaj Institute of Engineering andTechnology
Faculty of Science and Technology
Board of Studies: Computer Engineering
Syllabus Honors(Cloud Computing and Virtualization)
w.e.f. AY:2024-2025

Course Code	SEM	Courses Name	Teaching Scheme			Examination Scheme and Marks							Credits				
			TH	PR	TUT	Activity	ISE	ESE	TW	PR	OR	Total	T H	PR	OR	TUT	Total
CO23281	III	Cloud Computing Foundation	2	2		10	20	50	20	20		120	2	1			3
CO23291	IV	Cloud Computing Architecture and Design	2	2		20	20	50	20	20		130	2	1			3
CO23381	V	Automation Tools for Cloud Deployment	3	2		20	20	70	20	20		150	3	1			4
CO23391	VI	Advanced Virtualization Techniques for High Performance Cloud Infrastructure	3	2		20	20	70	20	20		150	3	1			4
CO23481	VII	Big Data Technology and Cloud Management	3	2		20	20	70	20	20		150	3	1			4
Total			13	10	0	90	100	310	100	100	0	700	13	5	0	0	18

Total	23	500		200
Internal	290	41.4	%	
External	410	58.6	%	

Academic Coordinator
Dr. P. M. Paithane

Department Autonomy Coordinator
Mr. M. D. Shelar

Head of Department
Dr. G. J. Chhajed

Dean Autonomy
Dr. C. B. Nayak

Dean Academic
Dr. S Bhosale

Principal
Dr. R. S. Bichkar

Principal
Vidya Pratishthan's
Kamalnayan Bajaj Institute of
Engineering & Technology
Vidyanagari, Baramati

CO23281 : Foundation of Cloud Computing

Teaching Scheme: TH: 02 Hrs/Week PR: 02 Hrs/Week	Credit: 03	Examination Scheme: Course Activity: 10 Mark In-Semester: 20 Mark
	Theory =2 Practical=1	End-Semester: 50 Mark Term work: 20 Mark Practical: 20 Mark

Prerequisite: Computer Network, Database management System, Computer Organization

Course Objective:

- To study fundamental concepts of cloud computing
- To understand Virtualization in Cloud Computing
- To Learn Administrations in cloud computing

Course Outcomes: Articulate

1. Summarise fundamental concepts of Cloud Computing.
2. Explain the concepts Virtualization in Cloud Computing
3. Explain the concepts Administrations in cloud computing

Course Activity : The course coordinator should identify relative and innovative activities for course activity. Below are some suggested course activity for course coordinator

1. Poster Presentation
2. Seminar Presentations
3. Survey on various cloud computing making tools
4. Industry Visit
5. Group Discussion

Course Contents	
Mapping of Course Outcomes for Unit I	**CO1**

UNIT I	Introduction to Cloud Computing	07 Hrs

Origins and Influences, Brief History, Definition, Characteristics, Business Drivers, Capacity Planning, Cost Reduction Organizational Agility, Technology Innovations, Technology Innovations vs. Enabling Technologies

Mapping of Course Outcomes for Unit II	**CO1**

UNIT II	Cloud Computing Scaling & Services	07 Hours

Cloud, IT Resource, On-Premise, Cloud Consumers and Cloud Providers, Scaling, Horizontal Scaling, Vertical Scaling, Cloud Service, Cloud Service Consumer, Goals and Benefits, Reduced Investments and Proportional Costs, Increased Scalability, Increased Availability and Reliability, Risks and Challenges,

Increased Security Vulnerabilities, Reduced Operational Governance Control, Limited Portability Between Cloud Providers, Multi-Regional Compliance and Legal Issues

Mapping of Course Outcomes for Unit III	CO2	
UNIT III	Cloud Computing Virtualization	07 Hours

Abstraction and Virtualization, Introduction to Virtualization Technologies, Load Balancing and Virtualization, Understanding Hyper visors, Virtual Machines Provisioning and Manageability Virtual Machine Migration Services, Provisioning in the Cloud Context Virtualization of CPU, Memory, I/O Devices, Virtual Clusters and Resource management

Mapping of Course Outcomes for Unit IV	CO3	
	Cloud Computing Administrations	

Roles and Boundaries, Cloud Provider, Cloud Consumer, Cloud Service Owner, Cloud Resource Administrator, Additional Roles, Organizational Boundary, Trust Boundary, Cloud Characteristics On-Demand Usage, Ubiquitous Access, Multitenancy (and Resource Pooling), Elasticity, Measured Usage, Resiliency

## Books and Other Resources

Text Books:

1. "Cloud Computing Concepts, Technology & Architecture", Thomas Erl, Zaigham Mahmood, and Ricardo Puttini, PRENTICE HALL

2. "Cloud Security A Comprehensive Guide to secure Cloud Computing", Ronald L. Krutz, Russell Dean Vines, Wile

Reference Books:

1. " Cloud Computing: A Practical Approach for Learning and Implementation", A. Srinivasan, J. Suresh, Pearson, ISBN: 978-81-317-7651-3

2. " Mastering Cloud Computing" Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi, McGraw Hill Education, ISBN-13:978-1-25-902995-

Practical Assignments

1. To study cloud service providers (aws, google & Microsoft-azure).
2. To demonstrate Infrastructure as a service using a cloud service provider.
3. To demonstrate Software as a service using a cloud service provider.
4. To demonstrate Storage as a service using a cloud service provider.
5. To make spreadsheets and notes using Google Drive.
6. Installation of VMWARE workstation & access the tools.
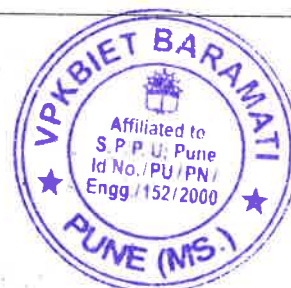
CO23291: Cloud Computing Architecture and Design

Teaching Scheme: TH: 02 Hrs/Week PR: 02 Hrs/Week	**Credit: 03**	**Examination Scheme:** Course Activity: 20 Mark
	Theory =2 **Practical=1**	In-Semester: 20 Mark End-Semester: 50 Mark Term work : 20 Mark Practical: 20 Mark

Prerequisite:

Operating System

Course Objective:

- To study fundamental concepts of cloud computing

- To learn various data storage methods on cloud

- To understand the implementation of Virtualization in Cloud Computing

- To learn the application and security on cloud computing

Course Outcomes:

CO1: Understand the different Cloud Computing environment

CO2: Use appropriate data storage technique on Cloud, based on Cloud application

CO3: Analyze virtualization technology and install virtualization software

CO4: Develop and deploy applications on Cloud

Course Activity : The course coordinator should identify relative and innovative activities for course activity. Below are some suggested course activity for course coordinator
1. Poster Presentation

2. Seminar Presentations

3. Survey on various cloud computing making tools

4. Industry Visit

5. Group Discussion

| Course Contents	
Mapping of Course Outcomes for Unit I	**CO1**

UNIT I	Introduction to Cloud Computing	07 Hrs

Importance of Cloud Computing, Characteristics, Pros and Cons of Cloud Computing, Migrating into the Cloud, Seven-step model of migration into a Cloud, Trends in Computing. **Cloud Service Models**: SaaS, PaaS, IaaS, Storage. **Cloud Architecture**: Cloud Computing Logical Architecture, Developing Holistic Cloud Computing Reference Model, Cloud System Architecture, Cloud Deployment Models.

Mapping of Course Outcomes for Unit II	CO2	
UNIT II	**Data Storage and Cloud Computing**	**07 Hours**

Data Storage: Introduction to Enterprise Data Storage, Direct Attached Storage, Storage Area Network, Network Attached Storage, Data Storage Management, File System, Cloud Data Stores, Using Grids for Data Storage. **Cloud Storage**: Data Management, Provisioning Cloud storage, Data Intensive Technologies for Cloud Computing. **Cloud Storage from LANs to WANs**: Cloud Characteristics, Distributed Data Storage.

Mapping of Course Outcomes for Unit III	CO3	
UNIT III	**Virtualization in Cloud Computing**	**07 Hours**

Introduction: Definition of Virtualization, Adopting Virtualization, Types of Virtualization, Virtualization Architecture and Software, Virtual Clustering, Virtualization Application, Pitfalls of Virtualization. **Grid, Cloud and Virtualization**: Virtualization in Grid, Virtualization in Cloud, Virtualization and Cloud Security. **Virtualization and Cloud Computing**: Anatomy of Cloud Infrastructure, Virtual infrastructures, CPU Virtualization, Network and Storage Virtualization

Mapping of Course Outcomes for Unit IV	CO4	
UNIT IV	**Cloud Architectures and Cloud Applications**	**07 Hours**

Amazon Web Services (AWS): Amazon Web Services and Components, Amazon Simple DB, Elastic Cloud Computing (EC2), Amazon Storage System, Amazon Database services (Dynamo DB). **Microsoft Cloud Services**: Azure core concepts, SQL Azure, Windows Azure Platform Appliance. **Cloud Computing Applications**: Healthcare: ECG Analysis in the Cloud, Biology: Protein Structure Prediction, Geosciences: Satellite Image Processing, Business and Consumer Applications: CRM and ERP, Social Networking, Google Cloud Application: Google App Engine. Overview of Open Stack architecture.

Books and Other Resources

Text Books:

3. "Cloud Computing: A Practical Approach for Learning and Implementation", A. Srinivasan, J. Suresh, Pearson, ISBN: 978-81-317-7651-3
4. "Mastering Cloud Computing" Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi, McGraw Hill Education, ISBN-13:978-1-25-902995-

Practical Assignments

1. Case study on Microsoft azure to learn about Microsoft Azure is a cloud computing platform and infrastructure, created by Microsoft, for building, deploying and managing applications and services through a global network of Microsoft-managed data centers.
2. Installation and configure Google App Engine.
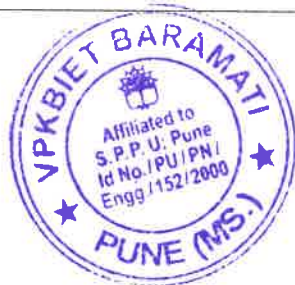
3. Creating an Application in SalesForce.com using Apex programming Language.

4. **Mini Project:** Setup your own cloud for Software as a Service (SaaS) over the existing LAN in your laboratory. In this assignment you have to write your own code for cloud controller using open-source technologies to implement with HDFS. Implement the basic operations may be like to divide the file in segments/blocks and upload/ download file on/from cloud in encrypted form. Case Study: Single Bus Organization and Microinstructions

CO23381: Automation Tools for Cloud Deployment

Teaching Scheme: TH: 03 Hrs/Week PR: 02 Hrs/Week	**Credit: 04**	**Examination Scheme:** Course Activity: 20 Mark In-Semester: 20 Mark End-Semester: 70 Mark Term work: 20 Mark Practical: 20 Mark
	Theory =3 **Practical=1**	

Prerequisite:

Basic of database ,Basic of web technology

Course Objective:

- To understand the cloud computing technology

- To understand the basic of cloud storage

- To understand the developing the application of cloud

- To understand VMWare Simulator

- To learn Configuration Management Tools and Deployment tools

Course Outcomes:

1. Use of Cloud computing technology

2. understand the basic of cloud storage

3. use of the developing the application of cloud

4. Understand the VMWare Simulator

5. Use of the Configuration Management Tools

6. Use of cloud Deployment tools

Course Activity : The course coordinator should identify relative and innovative activities for course activity. Below are some suggested course activity for course coordinator

1. Poster Presentation

2. Seminar Presentations

3. Survey on various cloud computing making tools

4. Industry Visit

5. Group Discussion

Course Contents		
Mapping of Course Outcomes for Unit I		**CO1**
UNIT I	**Introduction to cloud Computing Technology**	**07 Hours**

Hardware and Infrastructure : Clients: Mobile, Thin, Thik Security: Data Leakage, Offloading Work, Logging, Forensics, Development, Auditing Network: Basic Public Internet, Cloud Providers,   Cloud

Consumers Services: Identity, Integration, Mapping, Payments. **Accessing the Cloud** : Platforms :Web Application Framework, Web Hosting Service ,Proprietary Methods ,Web Applications : Sample Applications Web APIs: What Are APIs? ,How APIs Work ,API Creators

Mapping of Course Outcomes for Unit II		CO2
UNIT II	Cloud Storage	07 Hours

Overview: The Basics, Storage as a Service, Providers, Security, Reliability, Advantages, Cautions, Outages, Theft **Cloud Storage Providers** : Amazon Simple Storage Service (S3), Nirvanix, Google Bigtable Datastore, MobileMe, Live Mesh

Mapping of Course Outcomes for Unit III		CO3
UNIT III	Developing Applications	07 Hours

Google : Payment,Force.com and Google, Google Gears **Microsoft** : Live Services, Microsoft SQL Services, Microsoft .NET Services, Microsoft SharePoint Services and Dynamics CRM Services. Design Cast Iron Cloud **Development:** Google App Engine ,Salesforce.com, Microsoft Windows Azure, Troubleshooting, Application Management.

Mapping of Course Outcomes for Unit IV		CO4
UNIT IV	Introduction to VMWare Simulator	07 Hours

Basics of VMWare, advantages of VMware virtualization, using Vmware workstation, creating virtual machines-understanding virtual machines, create a new virtual machine on local host, cloning virtual machines, virtualize a physical machine, starting and stopping a virtual machine.

Mapping of Course Outcomes for Unit V		CO5
UNIT V	Configuration Management Tools	07 Hours

Introduction to Configuration Management Tools.
Configuration Management Tools : Ansible , Chef , Puppet, AWS Cloud Formation ,IBM Cloud Schematics

Mapping of Course Outcomes for Unit VI		CO6
UNIT VI	Continuous Integration/Continuous Deployment tools	07 Hours

Introduction to cloud deployment, Key cloud infrastructure automation features.
Deployment Tools : Jenkins , CircleCI, DuploCloud, Microsoft Azure Automation

Books and Other Resources

Text Books:

1. Cloud computing a practical approach - Anthony T.Velte , Toby J. Velte Robert Elsenpeter, TATA McGraw-Hill , New Delhi – 2010

2. Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online -

Michael Miller - Que 2008

Reference Books:

1. https://www.nops.io/blog/cloud-automation-tools/
2. https://www.shiksha.com/online-courses/cloud-computing-by-nptel-course-nptel18

Practical Assignments

1. Installation and configuration of own Cloud

2. Write a Program to Create, Manage and groups User accounts in ownCloud by Installing Administrative Features.

3. Case study on Amazon EC2 to learn about Amazon EC2,Amazon Elastic Compute Cloud is a central part of Amazon. com's cloud computing platform, Amazon Web Services. How EC2 allows users torrent virtual computers on which to run their own computer applications.

4. Case study on Microsoft azure to learn about Microsoft Azure is a cloud computing platform and infrastructure, created by Microsoft, for building, deploying and managing applications and services through a global network of Microsoft-managed data centers. How it work, different services provided by it.

5. Assignment to install and configure Google App Engine.

CO23391 : Advanced Virtualization Techniques for High Performance Cloud Infrastructure

Teaching Scheme: TH: 03 Hrs/Week PR: 02 Hrs/Week	Credit: 04	Examination Scheme: Course Activity: 20 Mark In-Semester: 20 Mark
	Theory =3 Practical=1	End-Semester: 70 Mark Term work: 20 Mark Practical : 20 Mark

Prerequisite:

Automation Tools for Cloud Development

Course Objective:

- To understand the introduction of Virtualization
- To learn the implementation of virtual machines
- To analyse the advance virtualization solutions for high performance computing
- To describe migration of cloud model
- To learn the advanced cloud simulator
- To understand the VMWare simulator

Course Outcomes:

1. Demonstrate the virtualization using CloudSim simulator.
2. Implementation of virtual machines
3. Demonstrate cloud based scenario for virtualization solutions.
4. Understand the Migrating into a Cloud
5. Demonstrate of advanced virtualization technique for high performance computing
6. Create the virtual machine using VMWare simulator

Course Activity : The course coordinator should identify relative and innovative activities for course activity. Below are some suggested course activity for course coordinator

1. Poster Presentation
2. Seminar Presentations
3. Survey on various cloud computing making tools
4. Industry Visit
5. Group Discussion

Course Contents	
Mapping of Course Outcomes for Unit I	CO1

UNIT I	Introduction to Virtualization	08 Hours

Virtualization and cloud computing - Need of virtualization – cost, administration, fast deployment, reduce infrastructure cost – limitations, Types of hardware virtualization: Full virtualization - partial virtualization - para virtualization, Desktop virtualization: Software virtualization – Memory virtualization - Storage virtualization – Data virtualization – Network virtualization

Mapping of Course Outcomes for Unit II		CO2	
UNIT II	Hypervisors and Virtual machines		06 Hours

Server Virtualization: Understanding Server Virtualization, types of server virtualization, Virtual machine basics, types of virtual machines, hypervisor concepts and types

Mapping of Course Outcomes for Unit III		CO3	
UNIT III	Virtualization Solutions		06 Hours

Understanding Microsoft's Virtualization solutions: Microsoft's Infrastructure Optimization Model, Virtualization and the Infrastructure Optimization Model, Benefits of Virtualization, Achieving the Benefits of Datacenter Virtualization, Achieving the Benefits of Client Virtualization, Achieving the Benefits of Cloud Virtualization

Mapping of Course Outcomes for Unit IV		CO4	
UNIT IV	Migrating into a Cloud		08 Hours

Introduction, Challenges while migrating to Cloud, Broad approaches to migrating into the cloud why migrate -deciding on cloud migration, the Seven-step model of migration into a cloud, Migration Risks and Mitigation, Enterprise cloud computing paradigm, relevant Deployment Models for Enterprise Cloud Computing, Adoption and Consumption Strategies, issues for enterprise applications on the cloud

Mapping of Course Outcomes for Unit V		CO5	
UNIT V	Cloud Simulators- CloudSim and GreenCloud		08 Hours

Introduction to Simulator, understanding CloudSim simulator, CloudSim Architecture(User code, CloudSim, GridSim, SimJava) Understanding Working platform for CloudSim, Introduction to GreenCloud

| Mapping of Course Outcomes for Unit VI		CO6	
UNIT VI	Introduction to VMWare Simulator		08 Hours

Basics of VMWare, advantages of VMware virtualization, using Vmware workstation, creating virtual machines-understanding virtual machines, create a new virtual machine on local host, cloning virtual machines, virtualize a physical machine, starting and stopping a virtual machine.

Books and Other Resources

Text Books:

3. David Marshall, Wade A. Reynolds, Advanced Server Virtualization: VMware and Microsoft Platform in the Virtual Data Center, Auerbach
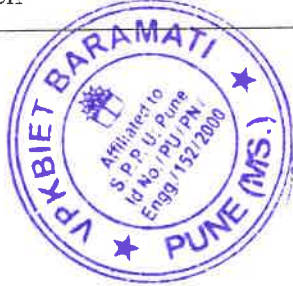
4. Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online - Michael Miller - Que 2008

5. Cloud computing a practical approach - Anthony T.Velte , Toby J. Velte Robert Elsenpeter, TATA McGraw- Hill , New Delhi – 2010

6. Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online - Michael Miller - Que 2008

Reference Books:

3. Publications, 2006. Cloud Computing (Principles and Paradigms), Edited by Rajkumar Buyya, James Broberg, Andrzej Goscinski, John Wiley & Sons, Inc. 2011

4. Cloud computing a practical approach - Anthony T.Velte , Toby J. Velte Robert Elsenpeter, TATA McGraw- Hill , New Delhi – 2010

5. Cloud computing for dummies- Judith Hurwitz , Robin Bloor , Marcia Kaufman, Fern Halper, Wiley Publishing, Inc, 2010

6. Cloud Computing (Principles and Paradigms), Edited by Rajkumar Buyya, James Broberg, Andrzej Goscinski, John Wiley & Sons, Inc. 2011

Practical Assignments

5. Installation and Simulation of Cloud Computing in CloudSim.

6. Installation and Simulation of Cloud Computing in GreenCloud

7. Simulate a cloud scenario using CloudSim and run a scheduling algorithm that is not in CloudSim

8. Find a procedure to launch virtual machine using trystack (online Openstack demo version).

9. Find a procedure to transfer the files from one virtual machine to another virtual machine.

10. Demonstrate how to simulate a Data Center with 1 Host and run one cloudlet on it using Cloudsim

11. How to create a virtual machine (VM) on google cloud platform.

CO23481 : Big Data Technology and Cloud Management

Teaching Scheme: TH: 03 Hrs/Week PR: 02 Hrs/Week	Credit: 04	Examination Scheme: Course Activity: 20 Mark In-Semester: 20 Mark
	Theory =3 Practical=1	End-Semester: 70 Mark Term work: 20 Mark Practical : 20 Mark

Prerequisite:

Basic understanding of cloud computing concepts

Familiarity with programming (e.g., Python, Java)

Knowledge of databases and SQL

Course Objective:

- To Understand the architecture and components of Big Data systems on GCP, AWS, and Azure
- To utilize cloud-based storage solutions for managing large datasets
- To implement data processing workflows using distributed computing frameworks
- To perform data analysis and visualization at scale
- To evaluate and compare Big Data solutions across different cloud platforms

Course Outcomes:

1. Understanding of the architecture and components of Big Data systems on GCP, AWS, and Azure
2. Hands-on Experience with Cloud-Based Tools and Services
3. Ability to Evaluate and Compare Cloud Solutions

Course Activity : The course coordinator should identify relative and innovative activities for course activity. Below are some suggested course activity for course coordinator

1. Poster Presentation
2. Seminar Presentations
3. Survey on various cloud computing making tools
4. Industry Visit
5. Group Discussion

Course Contents

Mapping of Course Outcomes for Unit I	CO1	
UNIT I **Introduction to Big Data Processing**		**08 Hours**

Introduction to Big Data Analytics, what is Big Data? What are the challenges? Introduction to Apache Hadoop and Map Reduce Apache Spark, Spark programming. (Python and Spark),Spark - Resilient

-Distributed Dataset (RDDs).

Mapping of Course Outcomes for Unit II		CO2	
UNIT II	**Storage Solutions**		**06 Hours**

Google Cloud Storage (GCS): Features, pricing, and use cases, Amazon Simple Storage Service (S3): Bucket creation, object operations, and storage classes, Google Cloud Dataflow: Streaming and batch processing with Apache Beam, Amazon EMR (Elastic MapReduce): Cluster creation, Hadoop, and Spark integration

Mapping of Course Outcomes for Unit III		CO3	
UNIT III	**NoSQL Databases**		**06 Hours**

Google Cloud Bigtable: Key concepts, schema design, and integration with other GCP services, Amazon DynamoDB: Key-value and document database capabilities, partition keys, and indexes, Azure Cosmos DB: Global distribution, consistency models, and SQL API.

Mapping of Course Outcomes for Unit IV		CO4	
UNIT IV	**Data Warehousing**		**08 Hours**

Google BigQuery: Data ingestion, querying, and optimization techniques, Amazon Redshift: Data warehousing concepts, cluster management, and query optimization, Azure Synapse Analytics: Data integration, querying with SQL pools, and serverless SQL queries

Mapping of Course Outcomes for Unit V		CO5	
UNIT V	**Real-time Analytics**		**08 Hours**

Google Cloud Pub/Sub and Dataflow for streaming data processing, Amazon Kinesis: Streams, firehose, and analytics with Kinesis Data Analytics, Azure Stream Analytics: Real-time data processing, event hubs, and data integration

Mapping of Course Outcomes for Unit VI		CO5	
UNIT VI	**Machine Learning and AI Integration**		**08 Hours**

Using Google AI Platform for machine learning model training and deployment, Amazon SageMaker: Model building, training, and deployment, Azure Machine Learning: Model training, deployment pipelines, and integration with Azure services.

Books and Other Resources

Text Books:
1. "Designing Data-Intensive Applications" by Martin Kleppmann
2. "Cloud Computing: Concepts, Technology & Architecture" by Thomas Erl, Ricardo Puttini, and Zaigham Mahmood
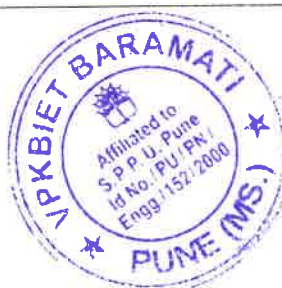
Reference Books:

1. Platform-specific documentation and online resources (GCP, AWS, Azure)

1. Setting up a Data Lake on GCP:

   - Create a storage bucket on Google Cloud Storage (GCS).
   - Ingest data into GCS using various methods (e.g., Cloud Console, gsutil, Cloud Storage Transfer Service).
   - Implement access controls and permissions for the data lake.

2. BigQuery Data Warehouse:

   - Load data into BigQuery from GCS or other sources (e.g., Cloud Storage, Cloud SQL).
   - Write SQL queries to perform data analysis and aggregation.
   - Explore partitioning and clustering strategies for performance optimization.

3. Data Processing with Dataflow:

   - Develop a Dataflow pipeline to process streaming or batch data.
   - Use Apache Beam SDK for Dataflow to transform and enrich data.
   - Monitor and troubleshoot pipeline execution using Stackdriver.

4. Building a Data Lake on AWS:

   - Set up an S3 bucket to store raw data.
   - Ingest data into S3 using AWS CLI, SDKs, or AWS Transfer Family.
   - Secure the data lake using IAM policies and bucket policies.

5. Building a Data Lake on AWS:

   - Set up an S3 bucket to store raw data.
   - Ingest data into S3 using AWS CLI, SDKs, or AWS Transfer Family.
   - Secure the data lake using IAM policies and bucket policies.