

Name: Krushnakumar Patle

Email: krishnapatle128@gmail.com

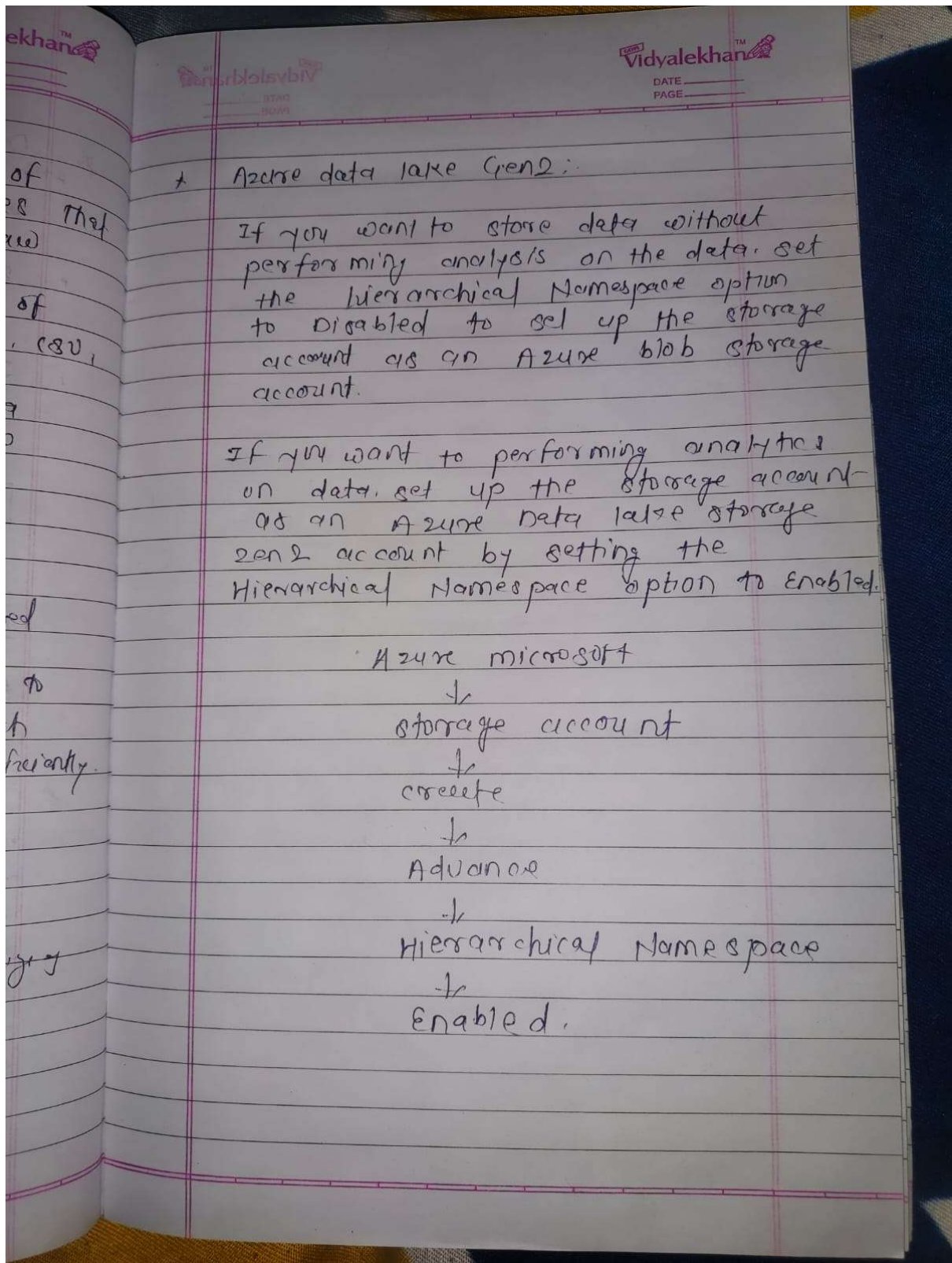
Batch: Data Engineering Batch-1

ADLS:

Azure Data Lake Storage Gen2 is a set of capabilities dedicated to big data analytics, built on Azure Blob Storage.

Data Lake Storage Gen2 converges the capabilities of Azure Data Lake Storage Gen1 with Azure Blob Storage. For example, Data Lake Storage Gen2 provides file system semantics, file-level security, and scale. Because these capabilities are built on Blob storage, you also get low-cost, tiered storage, with high availability/disaster recovery capabilities.

Data Lake Storage Gen2 makes Azure Storage the foundation for building enterprise data lakes on Azure. Designed from the start to service multiple petabytes of information while sustaining hundreds of gigabits of throughput, Data Lake Storage Gen2 allows you to easily manage massive amounts of data.



Creating data lake on Azure Storage :

The top screenshot shows the 'Create a storage account' page in the Microsoft Azure portal. The 'Advanced' tab is selected, and the 'Hierarchical Namespace' option is checked under 'Access protocols'. The bottom screenshot shows the 'workspace1064' storage account overview page. It displays details like Resource group (Test1064), Location (eastus), and various settings for Data Lake Storage and Security.

Create a storage account

Basics **Advanced** Networking Data protection Encryption Tags Review

Hierarchical Namespace

Hierarchical namespace, complemented by Data Lake Storage Gen2 endpoint, enables file and directory semantics, accelerates big data analytics workloads, and enables access control lists (ACLs) [Learn more](#)

Enable hierarchical namespace ☒

Access protocols

Blob and Data Lake Gen2 endpoints are provisioned by default [Learn more](#)

Enable SFTP ☐

Enable network file system v3 ☐

Blob storage

[Review](#) [< Previous](#) [Next: Networking >](#) [Give feedback](#)

workspace1064 Storage account

Search

Upload Open in Explorer Delete Move Refresh Open in mobile CLI / PS Feedback

Essentials

Resource group (move) : Test1064 Performance : Standard

Location : eastus Replication : Read-access geo-redundant storage (RA-GRS)

Primary/Secondary Location : Primary: East US, Secondary: West US Account kind : StorageV2 (general purpose v2)

Subscription (move) : Azure subscription 1 Provisioning state : Succeeded

Subscription ID : 984f097c-963c-4eb6-a20d-839457ae9f08 Created : 2/16/2024, 4:24:21 PM

Disk state : Primary: Available, Secondary: Available

Tags (edit) : Add tags

Properties Monitoring Capabilities (5) Recommendations (0) Tutorials Tools + SDKs

Data Lake Storage

Hierarchical namespace	Enabled
Default access tier	Hot
Blob anonymous access	Disabled
Blob soft delete	Enabled (7 days)
Container soft delete	Enabled (7 days)
Versioning	Disabled

Security

Require secure transfer for REST API operations	Enabled
Storage account key access	Enabled
Minimum TLS version	Version 1.2
Infrastructure encryption	Disabled

Networking

Connecting account to local :

