@falekmiah

falekmiah.com

FalekMiah01

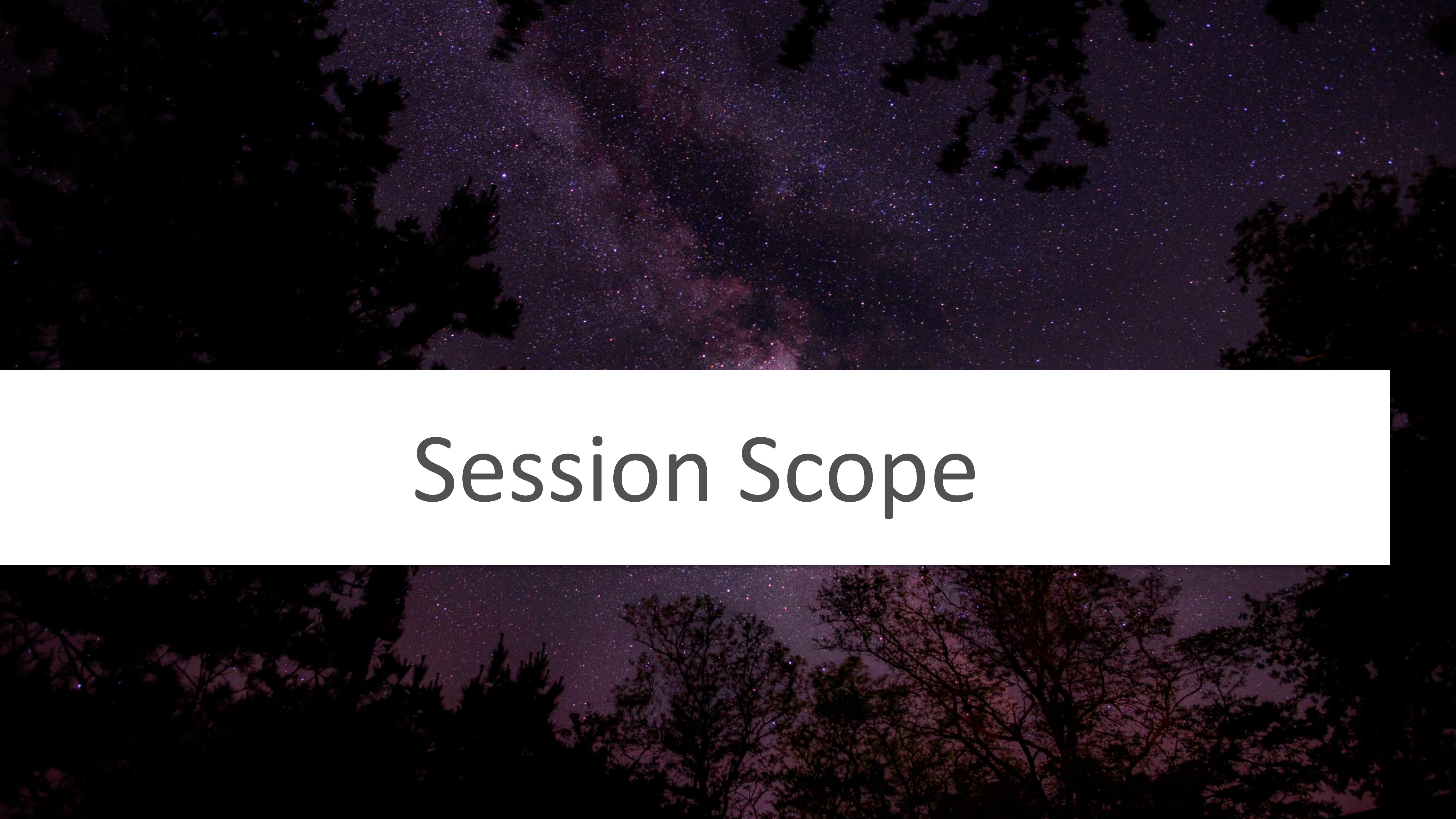
Quest to Delta Optimisation

Falek Miah

Principal Data Consultant

- 15+ Years Microsoft Data Analytics
- Intensive Data Engineering Experience
- Data, Cloud & DevOps Enthusiast
- Microsoft Azure, Databricks (Spark), Terraform (HashiCorp) certified

@falekmiah
falekmiah.com
FalekMiah01



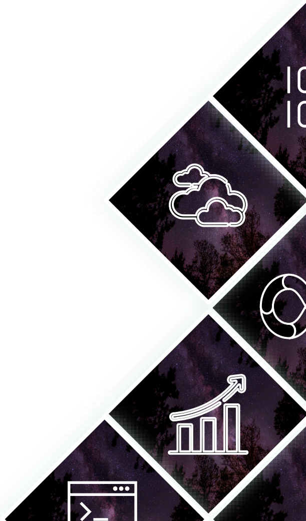
Session Scope

Session Scope

- Optimizing delta files and tables can be challenging and even a daunting task.

- Techniques like partitioning and z-ordering can be limited and inflexible
  - Partition – Difficult & Complex
  - Z-Order – Expensive Operation

- In this session
  - Small File Problem
  - Different Optimizing Approaches
  - Liquid Clustering



Small File Problem

Small File Problem

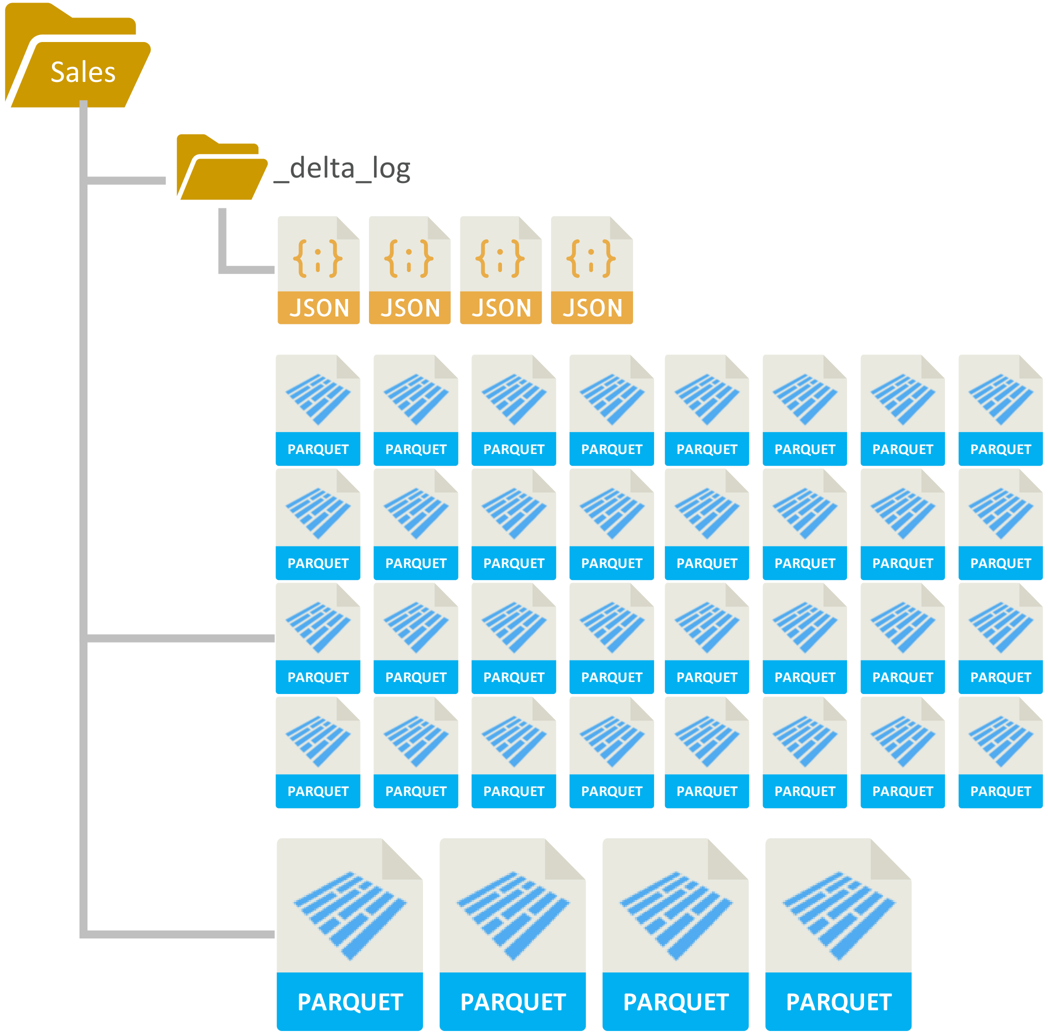
- Parquet files have an **optimal size**

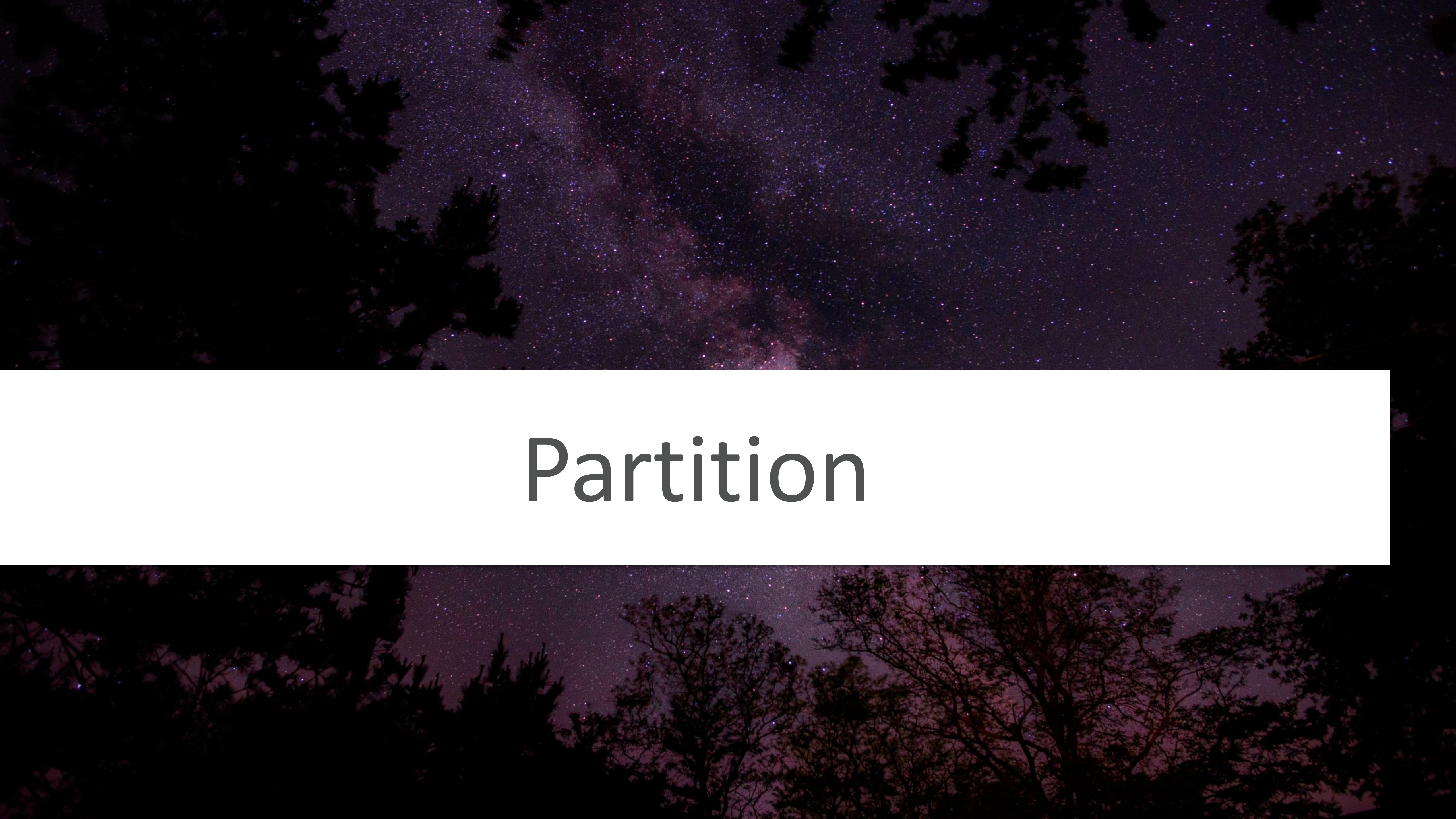
- Updates **small and inefficient files**

- \`**OPTIMIZE**\` command
  - Compacts small files into larger
  - Performed on entire table

- Files are **NOT** deleted and add to the JSON transaction log

- To remove obsolete history files using **VACUUM** command

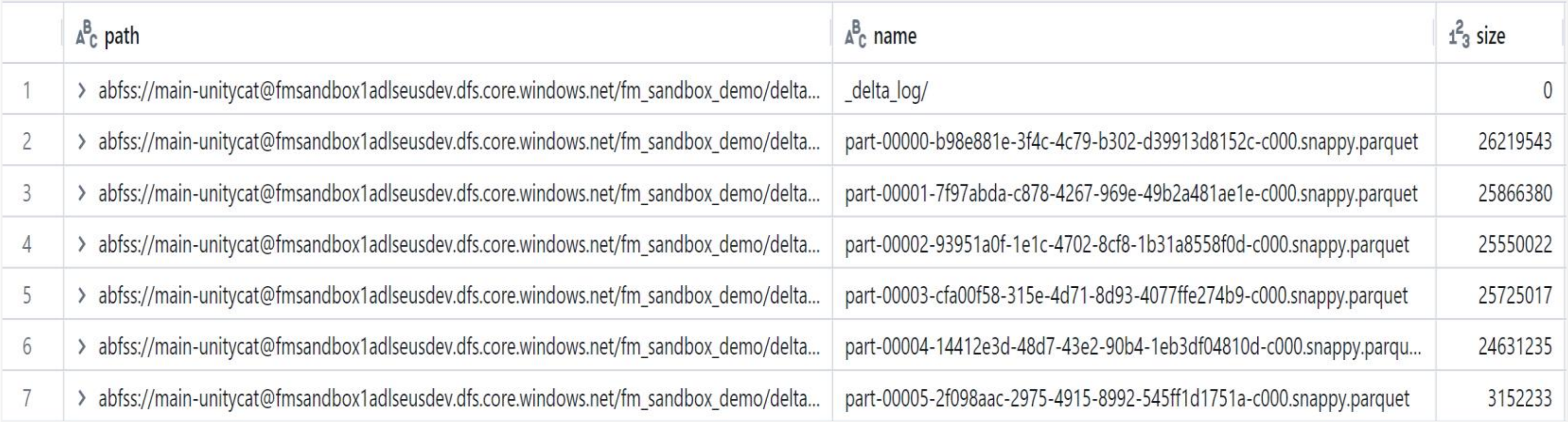




Partition

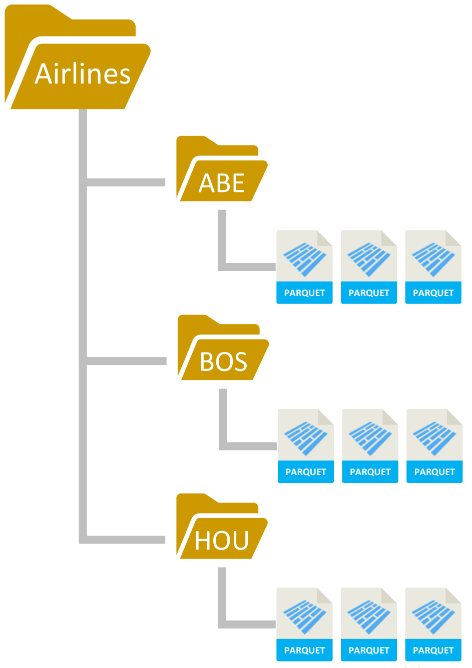
Partition

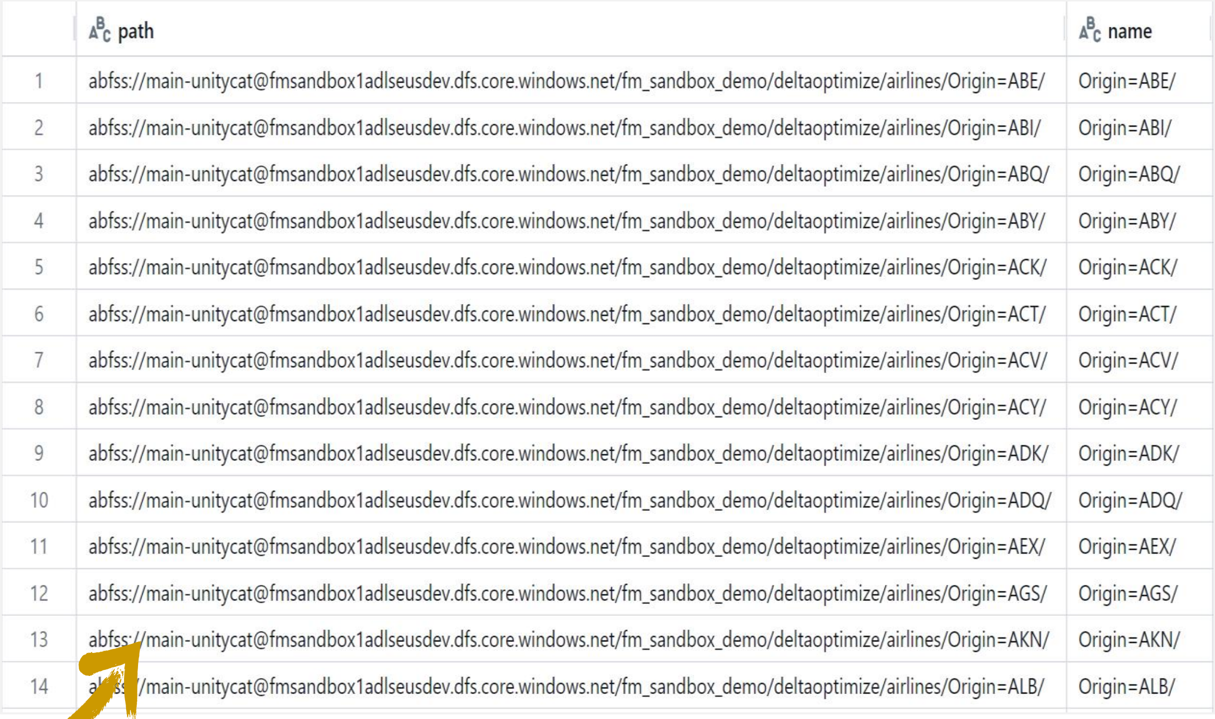
- Data is stored as multiple small files in folder



	path	name	size
1	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	_delta_log/	0
2	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	part-00000-b98e881e-3f4c-4c79-b302-d39913d8152c-c000.snappy.parquet	26219543
3	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	part-00001-7f97abda-c878-4267-969e-49b2a481ae1e-c000.snappy.parquet	25866380
4	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	part-00002-93951a0f-1e1c-4702-8cf8-1b31a8558f0d-c000.snappy.parquet	25550022
5	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	part-00003-cfa00f58-315e-4d71-8d93-4077ffe274b9-c000.snappy.parquet	25725017
6	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	part-00004-14412e3d-48d7-43e2-90b4-1eb3df04810d-c000.snappy.parqu...	24631235
7	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/delta...	part-00005-2f098aac-2975-4915-8992-545ff1d1751a-c000.snappy.parquet	3152233

- Partition divides those data file into useful slices





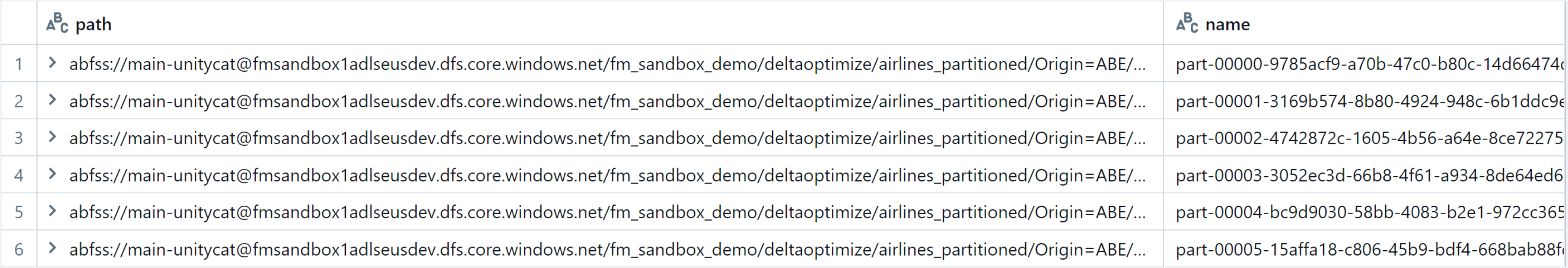
	path	name
1	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ABE/	Origin=ABE/
2	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ABI/	Origin=ABI/
3	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ABQ/	Origin=ABQ/
4	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ABY/	Origin=ABY/
5	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ACK/	Origin=ACK/
6	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ACT/	Origin=ACT/
7	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ACV/	Origin=ACV/
8	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ACY/	Origin=ACY/
9	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ADK/	Origin=ADK/
10	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ADQ/	Origin=ADQ/
11	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=AEX/	Origin=AEX/
12	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=AGS/	Origin=AGS/
13	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=AKN/	Origin=AKN/
14	abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines/Origin=ALB/	Origin=ALB/

```
-- Query
SELECT * FROM airline
WHERE Origin = “ABE”
```

Partition - Consideration

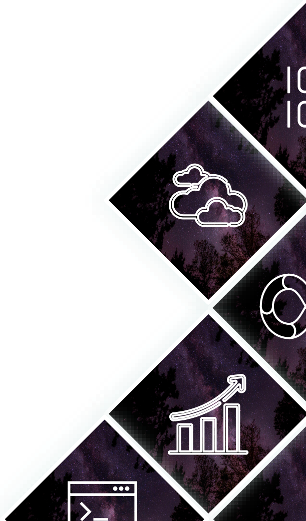
👎 High Cardinality

👎 Small Files Problem



	path	name
1	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines_partitioned/Origin=ABE/...	part-00000-9785acf9-a70b-47c0-b80c-14d66474
2	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines_partitioned/Origin=ABE/...	part-00001-3169b574-8b80-4924-948c-6b1ddc9
3	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines_partitioned/Origin=ABE/...	part-00002-4742872c-1605-4b56-a64e-8ce72275
4	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines_partitioned/Origin=ABE/...	part-00003-3052ec3d-66b8-4f61-a934-8de64ed6
5	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines_partitioned/Origin=ABE/...	part-00004-bc9d9030-58bb-4083-b2e1-972cc365
6	› abfss://main-unitycat@fmsandbox1adlseusdev.dfs.core.windows.net/fm_sandbox_demo/deltaoptimize/airlines_partitioned/Origin=ABE/...	part-00005-15affa18-c806-45b9-bdf4-668bab88f

👎 Not Flexible

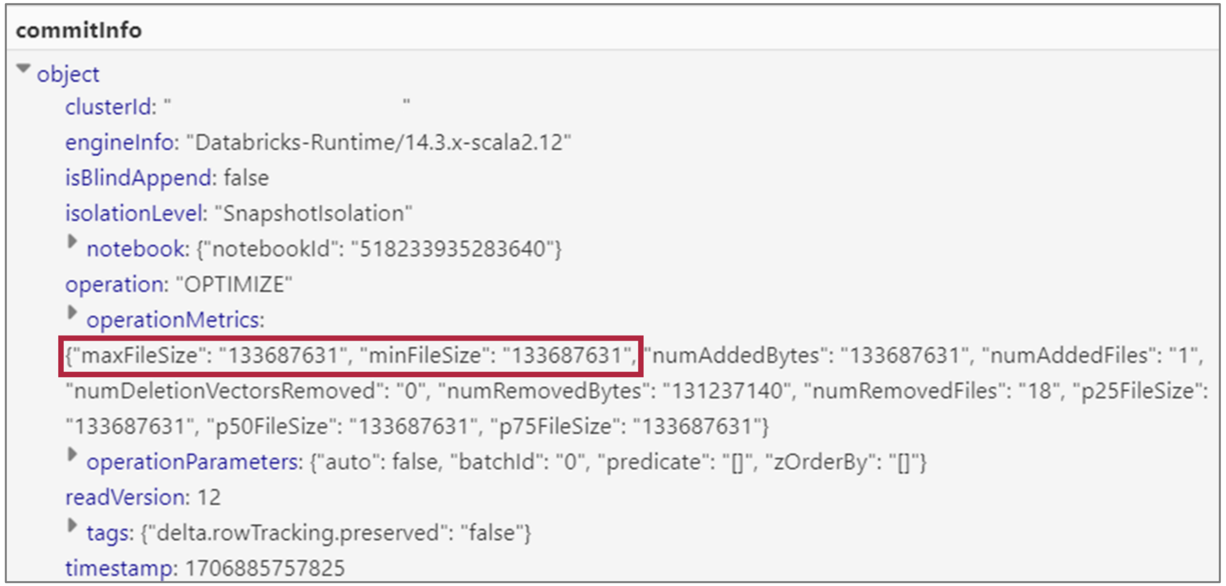


Demo

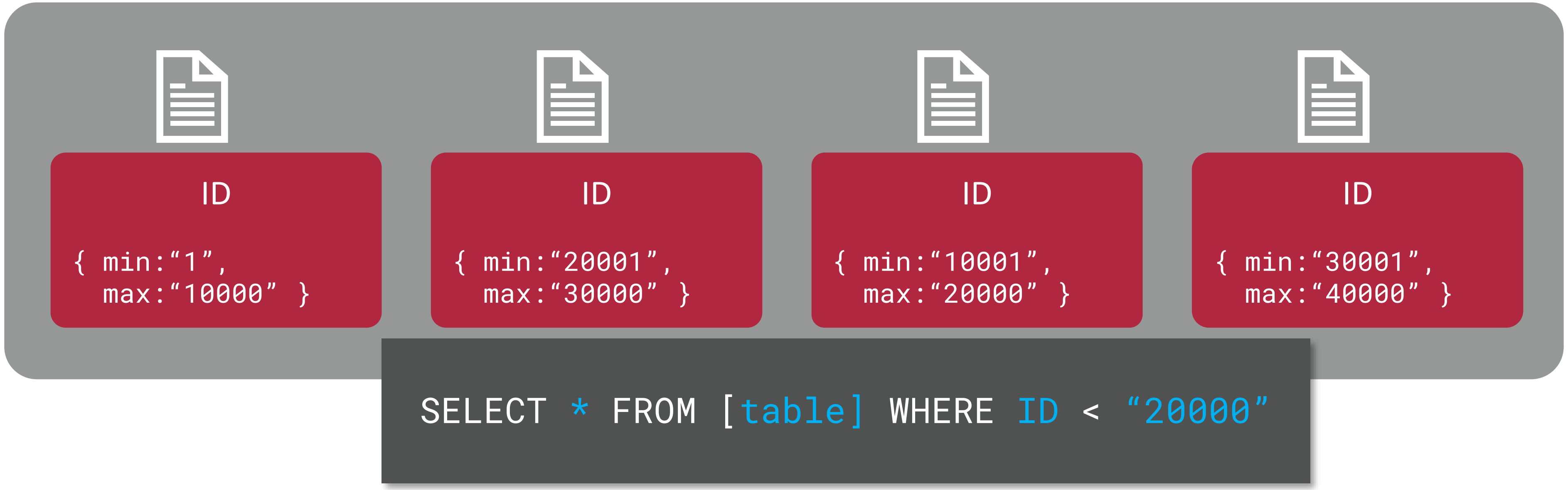
Data Skipping & Z-Ordering

Data Skipping

- Collect Statistics in Transaction Log
- Collects “min/max” Value For First 32 Columns
- Selectively Ignore Files



```
commitInfo
▾ object
    clusterId: "                    "
    engineInfo: "Databricks-Runtime/14.3.x-scala2.12"
    isBlindAppend: false
    isolationLevel: "SnapshotIsolation"
  ▸ notebook: {"notebookId": "518233935283640"}
    operation: "OPTIMIZE"
  ▸ operationMetrics:
    {"maxFileSize": "133687631", "minFileSize": "133687631", "numAddedBytes": "133687631", "numAddedFiles": "1",
    "numDeletionVectorsRemoved": "0", "numRemovedBytes": "131237140", "numRemovedFiles": "18", "p25FileSize":
    "133687631", "p50FileSize": "133687631", "p75FileSize": "133687631"}
  ▸ operationParameters: {"auto": false, "batchId": "0", "predicate": "[]", "zOrderBy": "[]"}
    readVersion: 12
  ▸ tags: {"delta.rowTracking.preserved": "false"}
    timestamp: 1706885757825
```



ID

{ min:“1”,
max:“10000” }

ID

{ min:“20001”,
max:“30000” }

ID

{ min:“10001”,
max:“20000” }

ID

{ min:“30001”,
max:“40000” }

```
SELECT * FROM [table] WHERE ID < “20000”
```

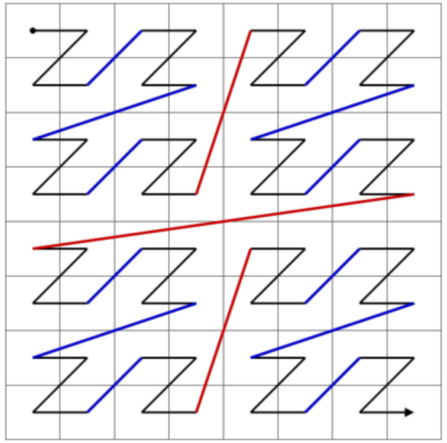
Z-Ordering

“Sort the data on specific columns before writing to files, to optimize data skipping”

```
--Optimize an entire table
OPTIMIZE [database].[table] ZORDER BY [columnName]
```

- Apply on Join Keys or Commonly Queried Columns
- Apply on Columns Statistics Are Collected

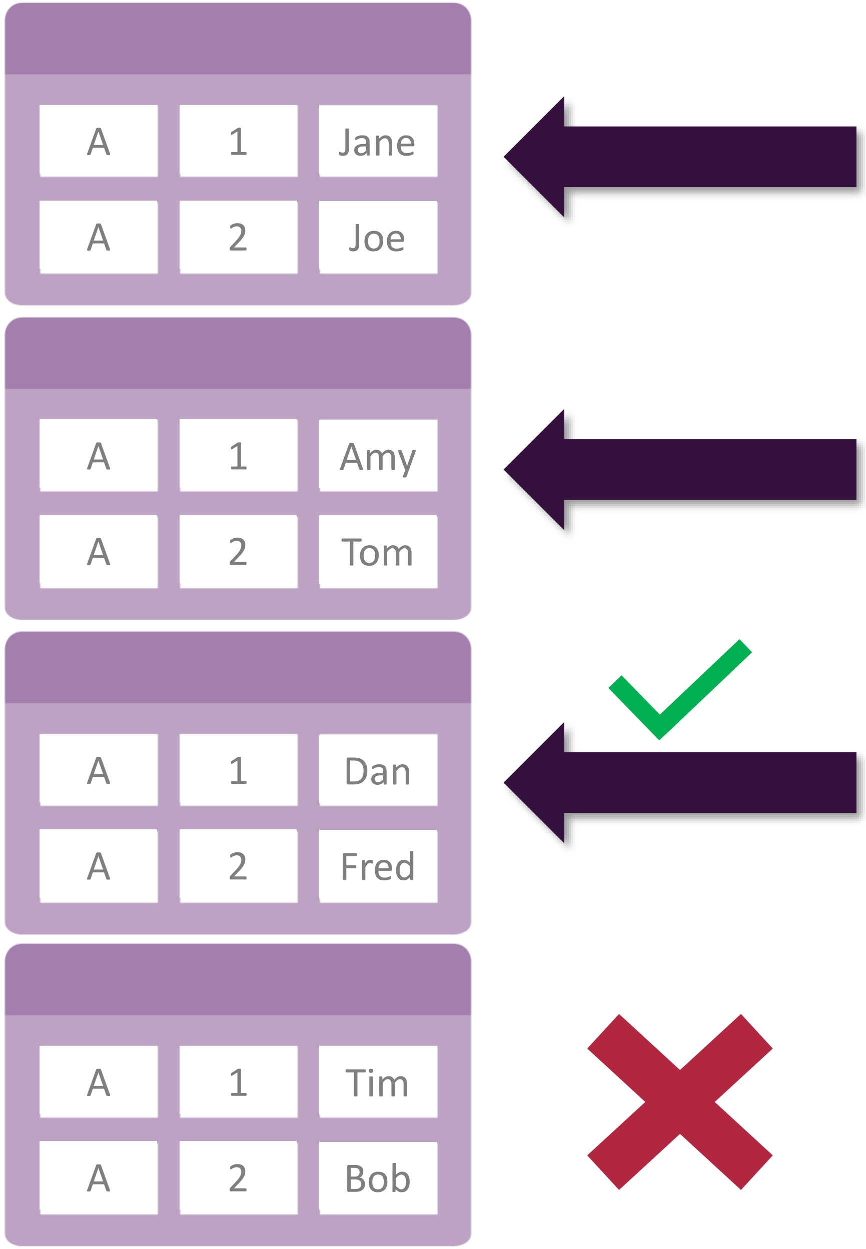
Z-Order Curve



Source: https://en.wikipedia.org/wiki/Z-order_curve

Z-Ordering & Data Skipping

```
SELECT count(*) FROM Employees
WHERE Name = ‘Dan’
```



A	1	Jane
A	2	Joe

A	1	Amy
A	2	Tom

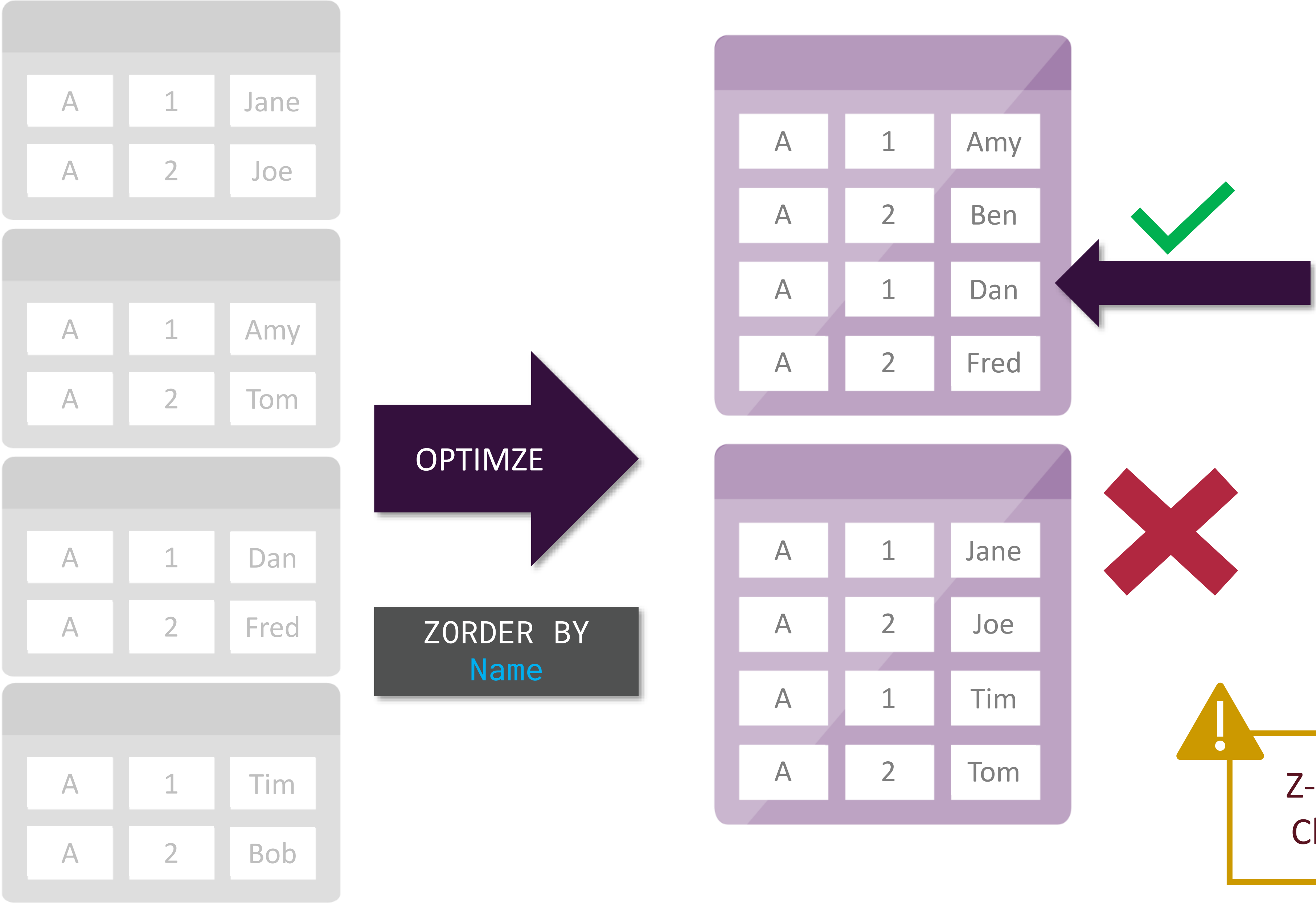
A	1	Dan
A	2	Fred

A	1	Tim
A	2	Bob

- The small files are not ordered
- SQL statement to query data

Z-Ordering & Data Skipping

```
SELECT count(*) FROM Employees
WHERE Name = ‘Dan’
```



A	1	Jane
A	2	Joe

A	1	Amy
A	2	Tom

A	1	Dan
A	2	Fred

A	1	Tim
A	2	Bob

```
ZORDER BY
Name
```

A	1	Amy
A	2	Ben
A	1	Dan
A	2	Fred

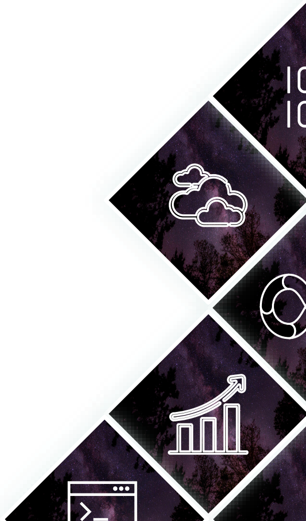
A	1	Jane
A	2	Joe
A	1	Tim
A	2	Tom

Z-Order as like a Clustered Index

Z-Ordering – Consideration

- 👎 Rewrite All The Data
- 👎 No Checkpointing
- 👎 No Table Level
- 👎 Not Flexible

⚠ Z-Order can be expense!
Best to perform as an out-of-hours
maintenance operation

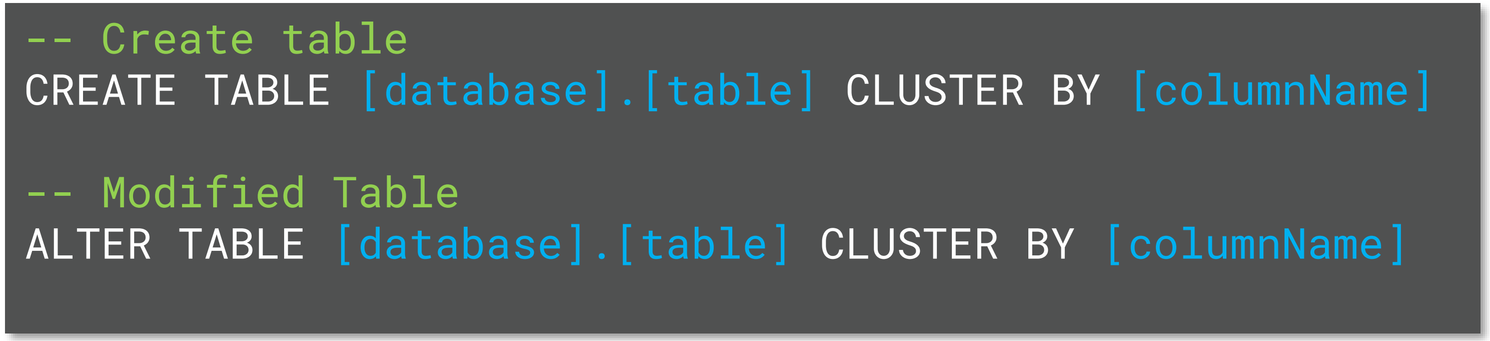


Demo

Liquid Clustering

Liquid Clustering

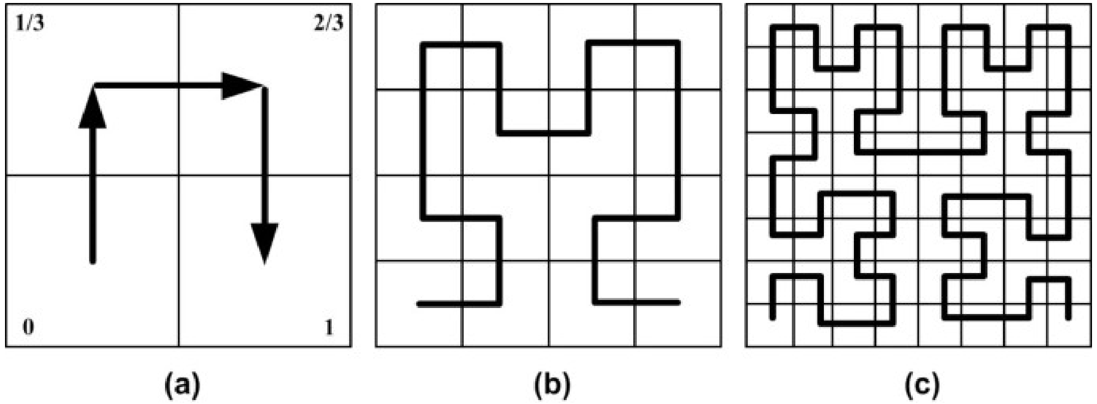
- Simplifies Data Layout changes
- More Adaptable & Flexible
- Table Level Clustering
- Streamlines Operations
- Simplifies Data Management
- Enhances query performance



```
-- Create table
CREATE TABLE [database].[table] CLUSTER BY [columnName]

-- Modified Table
ALTER TABLE [database].[table] CLUSTER BY [columnName]
```

Hilbert Curve Algorithm



Source: https://asp-eurasipjournals.springeropen.com/articles/10.1186/1687-6180-2012-181/figures/1

Liquid Clustering

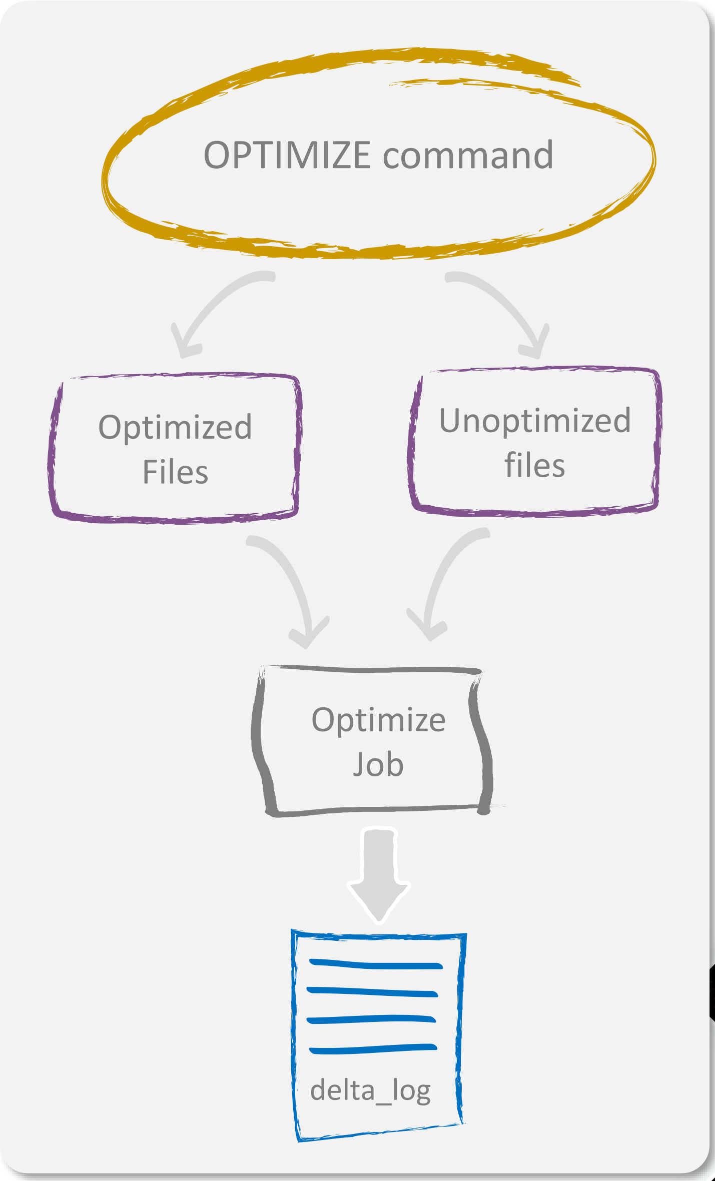
- Incremental Clustering

Unoptimized vs Optimized

- Optimize Flow

- Metadata Integration

_delta_log



Demo

Liquid Clustering – Consideration

- Default Choice

- Compatibility

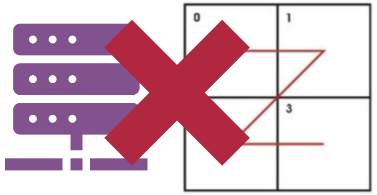
```
> Databricks Runtime 13.3 LTS
```

- Uses Specific Delta Versions

7 writer

3 reader

- Not Compatible With Previous Strategies

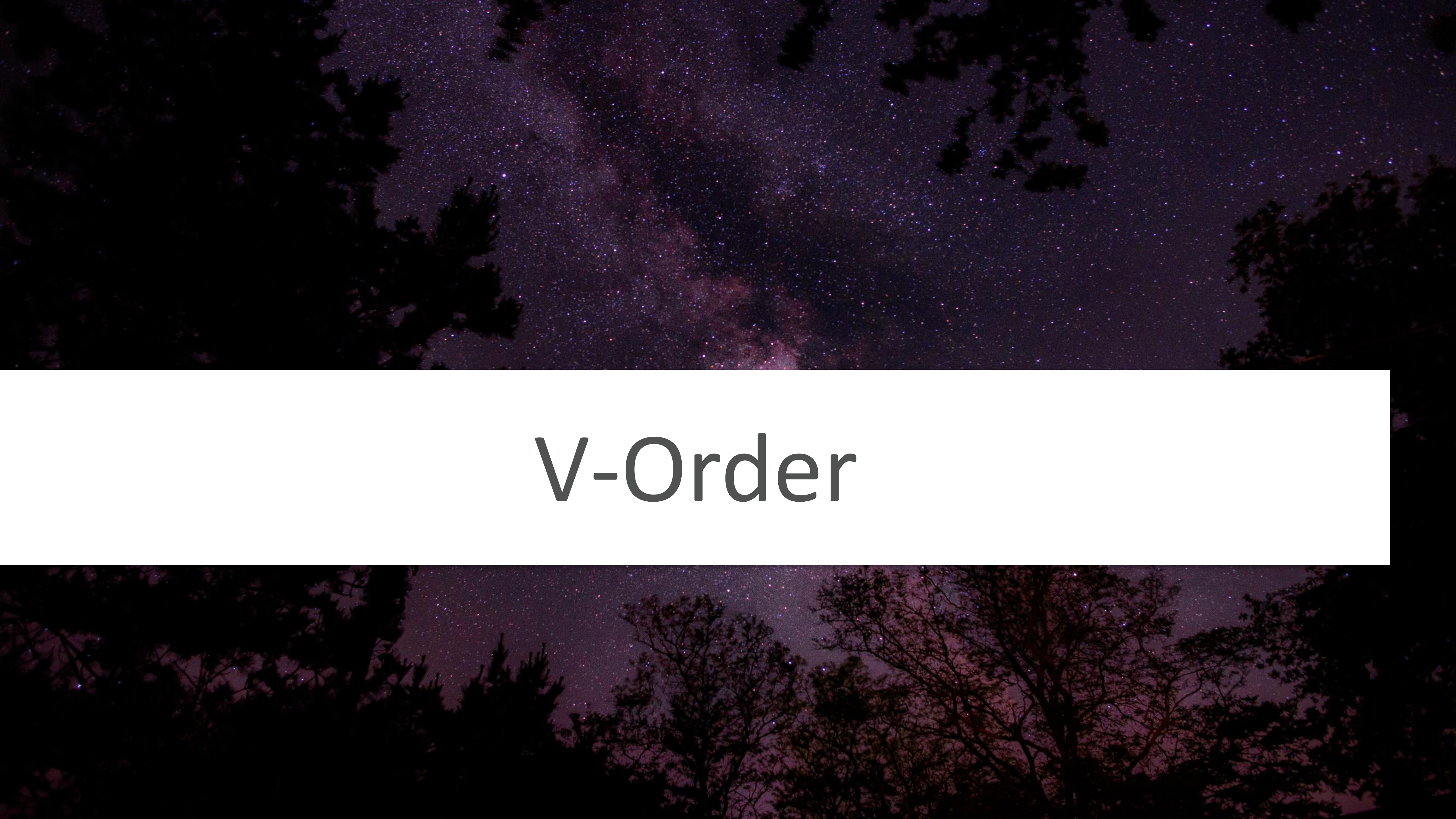


- Eliminates OPTIMIZE ZORDER BY

```
OPTIMIZE [da     e].[table]
ZORDER BY [c     Name]
```

Liquid Clustering – Limitations

- ⚠ 4 Columns
- ⚠ Columns With Statistics
- ⚠ Structured Streaming
- ⚠ Delta Sharing



V-Order

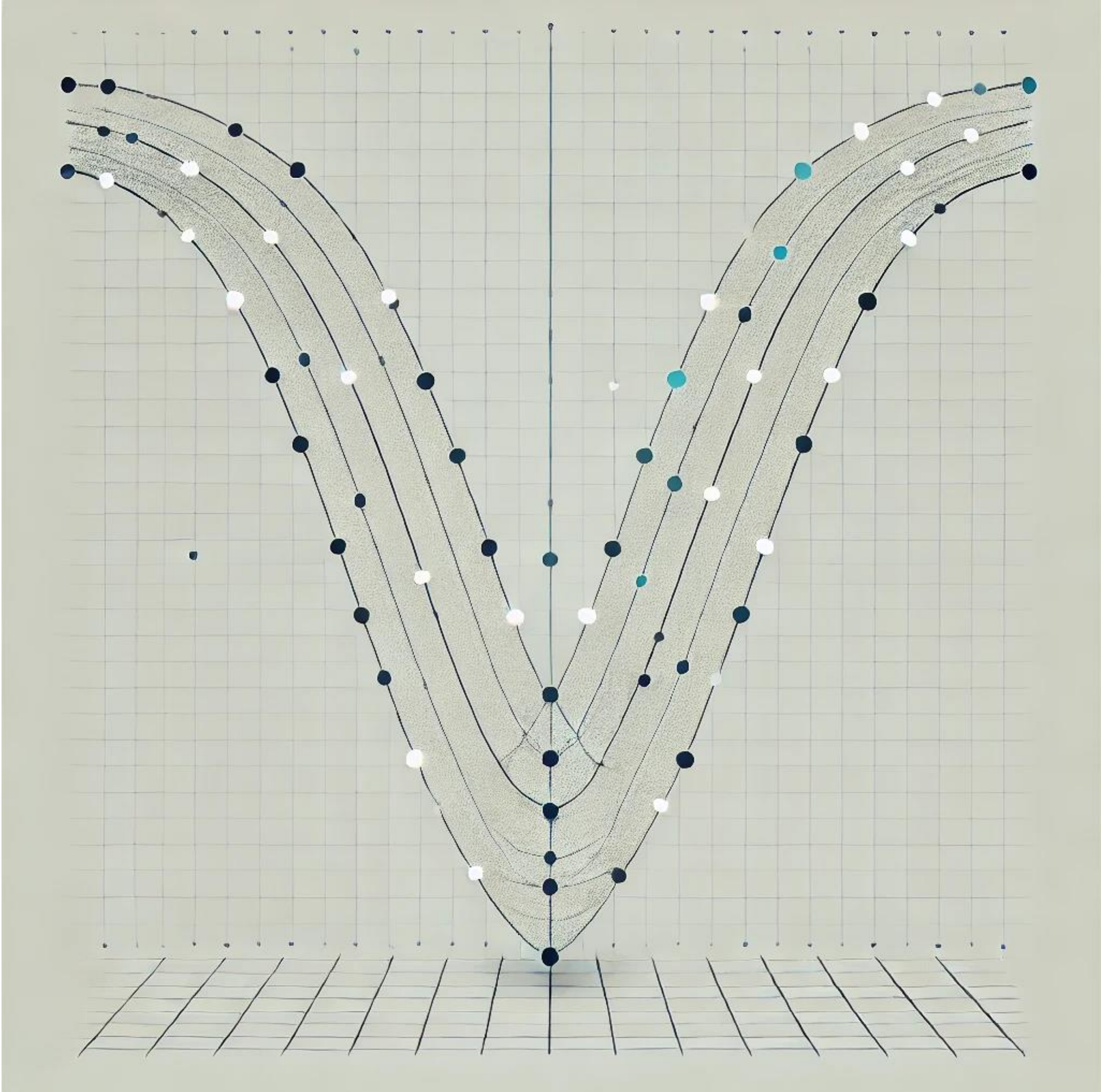
V-Order

- Write-Time Optimization
- Enhances Read Efficiency
- Optimizes Resource Utilization
- Enhances Performance
- Enabled by Default
- Fully Compatibility

```
ZORDER BY [columnName]

VACUUM [table]
```

- Considerations

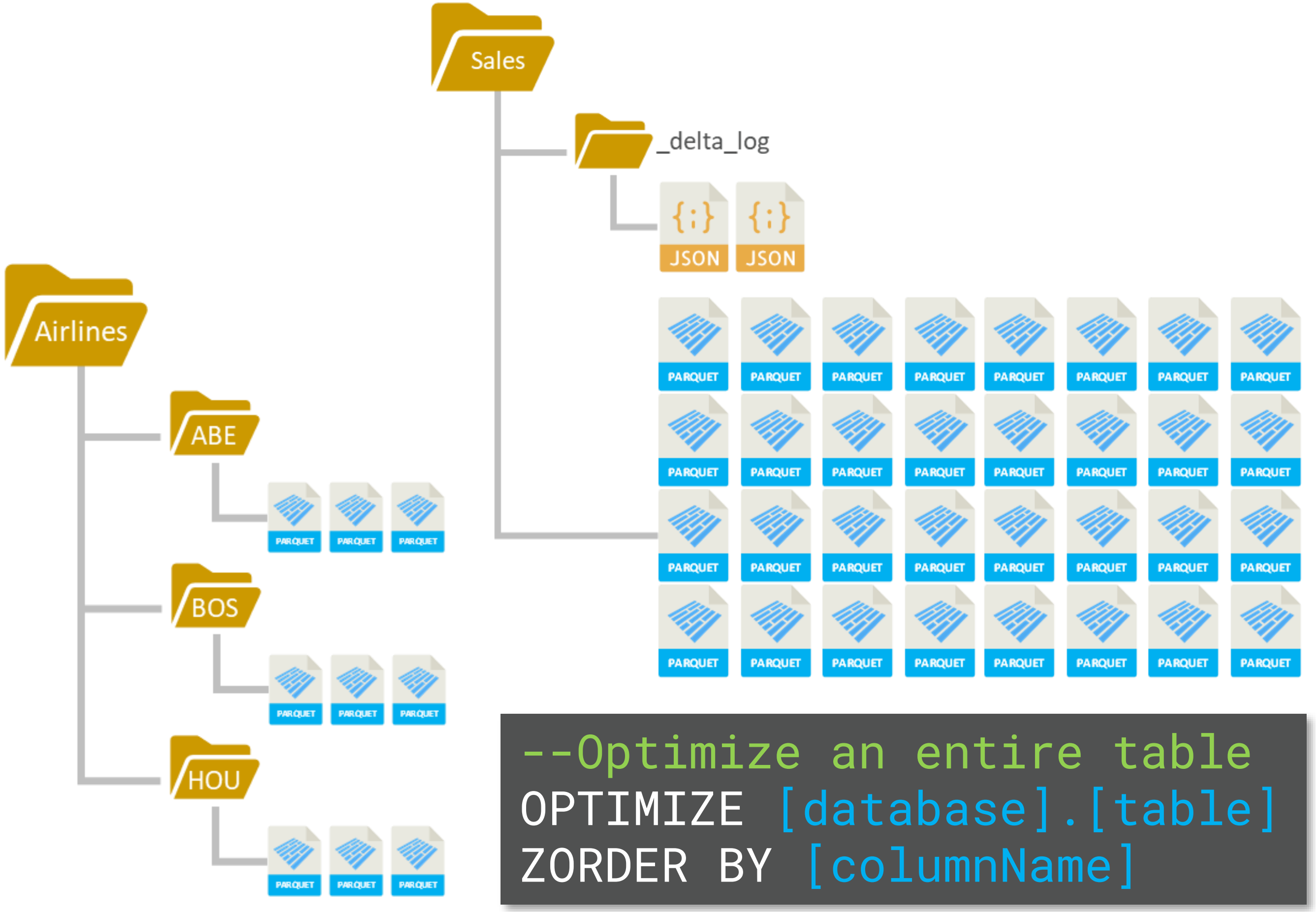


Demo

Wrap Up

Wrap Up

- Small File Problem
- Partition
- Z-Ordering



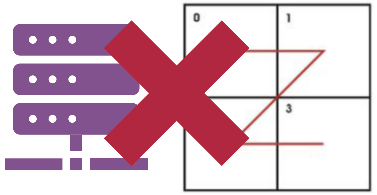
```
--Optimize an entire table
OPTIMIZE [database].[table]
ZORDER BY [columnName]
```

Wrap Up

```
-- Create table
CREATE TABLE [database].[table] CLUSTER BY [columnName]

-- Modified Table
ALTER TABLE [database].[table] CLUSTER BY [columnName]
```

- Small File Problem
- Partition
- Z-Ordering
- Liquid Clustering



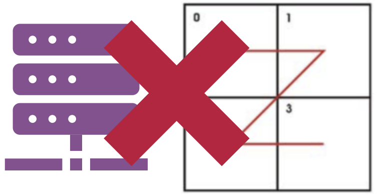
Wrap Up

- ↻ Small File Problem
- ↻ Partition
- ↻ Z-Ordering
- ↻ Liquid Clustering
- ↻ V-Order

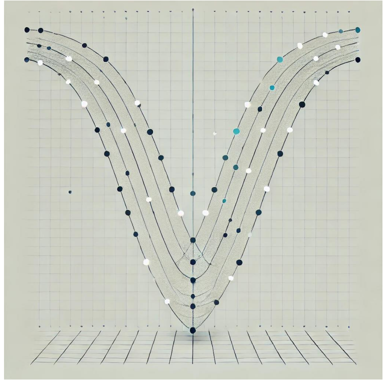
```
-- Create table
CREATE TABLE [database].[table] CLUSTER BY [columnName]

-- Modified Table
ALTER TABLE [database].[table] CLUSTER BY [columnName]
```

Unoptimized VS Optimized



Try it out



DATA:Scotland
2024

Session Feedback

Event Feedback

Thank You

✉ | falek@advancinganalytics.co.uk

@falekmiah

www | falekmiah.com

FalekMiah01

DATA:Scotland 2024

bridgeall

Tabular Editor

MySQL™

DATAmasterminds

30 Waterstons

ADVANCING ANALYTICS

quorum

hippo

aim