

# Data Loading Using Azure Synapse Analytics

## Introduction:

This document provides a step-by-step guide on how to load data into Azure Synapse Analytics, focusing on using serverless SQL databases to create views for data reporting. Azure Synapse Analytics is a powerful tool for data integration, analytics, and reporting.

## Prerequisites:

Before you begin, ensure that you have the following:

Azure Synapse Analytics workspace set up.

Data stored in Azure Data Lake Storage Gen2, organized in containers and folders.

Access to Power BI Desktop for data reporting.

## Steps:

### Accessing Azure Synapse Analytics:

Log in to the Azure portal.

Navigate to your Azure Synapse Analytics workspace.

### Opening Synapse Studio:

Within the Synapse workspace, click on "Open Synapse Studio." This takes you to the Synapse workspace, which is used for data loading and processing.

### Workspace Layout:

The Synapse Analytics workspace layout is like Azure Data Factory Studio.

It consists of tabs such as "Integrate," "Monitor," "Manage," "Data," and "Develop."

### Creating a Serverless SQL Database:

Go to the "Data" tab and click the plus icon.

Choose "SQL database" and select the "Serverless" option for smaller workloads.

Enter a database name (e.g., "gold\_DB") and create the database.

### Differences between Serverless and Dedicated SQL Database:

Serverless SQL Database is suitable for smaller workloads and uses a built-in pool.

Dedicated SQL Database is for higher workloads and requires creating an external compute.

Serverless uses only compute, while Dedicated uses both storage and compute.

### Linking Azure Data Lake Storage:

Azure Synapse Analytics is linked to Azure Data Lake Storage Gen2 by default.

Data Lake containers and folders are accessible within the Synapse workspace.

#### Creating Views Dynamically:

To create views dynamically, use a stored procedure with parameters.

The stored procedure retrieves table names from Data Lake and creates views accordingly.

#### Creating a Link Service Connection:

Go to the "Manage" tab and create a new link service connection for Serverless SQL Database.

Configure it with the database name, authentication type, and managed identity.

#### Creating a Pipeline:

In the "Integrate" tab, create a new pipeline.

Add a "Get Metadata" activity to retrieve table names from Data Lake.

Use a "For Each" activity to iterate through the table names.

Add a "Stored Procedure" activity to execute the dynamic view creation script.

#### Testing and Running the Pipeline:

Test the pipeline to ensure it retrieves table names and creates views.

Run the pipeline to create views for all tables in the Gold container.

#### Data Reporting with Power BI:

##### Open Power BI Desktop.

Connect Power BI to the Serverless SQL Database using the link service connection.

Fetch all the views created in the database.

Create interactive dashboards and reports for data analysis.

#### Conclusion:

This document has outlined the process of data loading using Azure Synapse Analytics, focusing on the creation of serverless SQL databases and dynamic view generation. It also touches on using Power BI for data reporting and analysis.