# Azure Discovery Days 2019

## Data Analytics & Near Real Time Intelligence with Azure - Hands-On Lab Guide

## Lab 2: Copy processed data from Azure Blob storage to Azure SQL DB

### Summary

In this hands-on lab, you will:

1. Set up Azure SQL Database
2. Set up Azure Data Factory Project
3. Configure Azure Data Factory project using Copy wizard/Pipelines
4. Create Power BI Reports using Power BI desktop and Azure SQL DB as data source
5. Deploy Power BI reports to Power BI Online and create Dashboards

### About this Lab

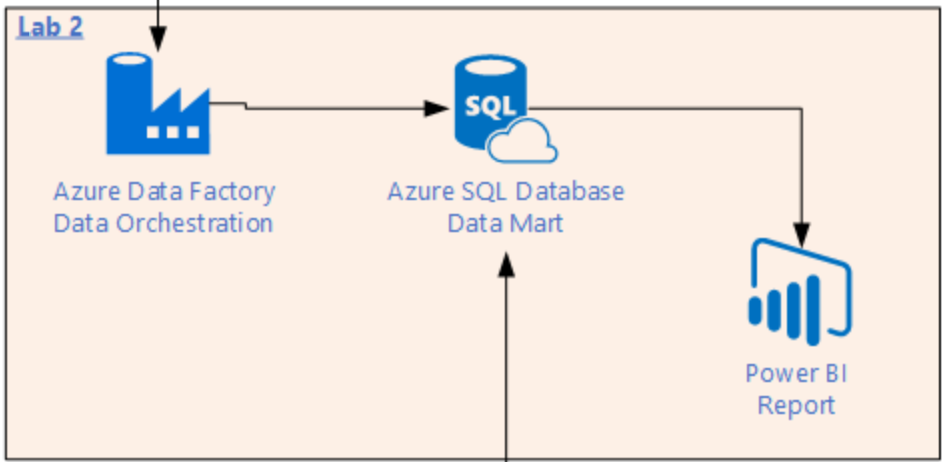
Copy data from Azure Blob storage (Parquet format) to Azure SQL Database using copy recursion option. And then create Power BI reports on top of Azure SQL Database.

### References

### General Notes

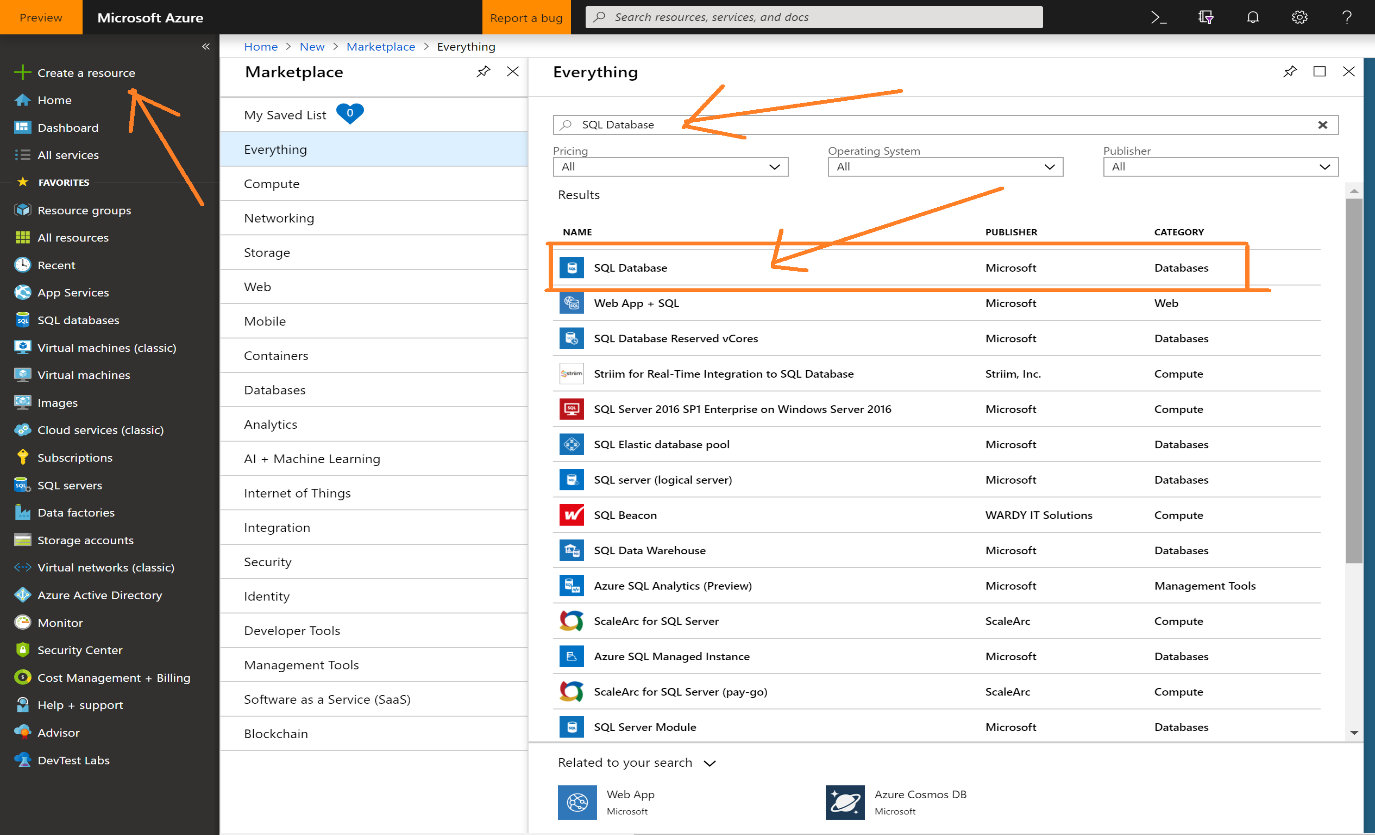
### Architecture for this Lab

The tasks in this lab cover the following components of the overall architecture.

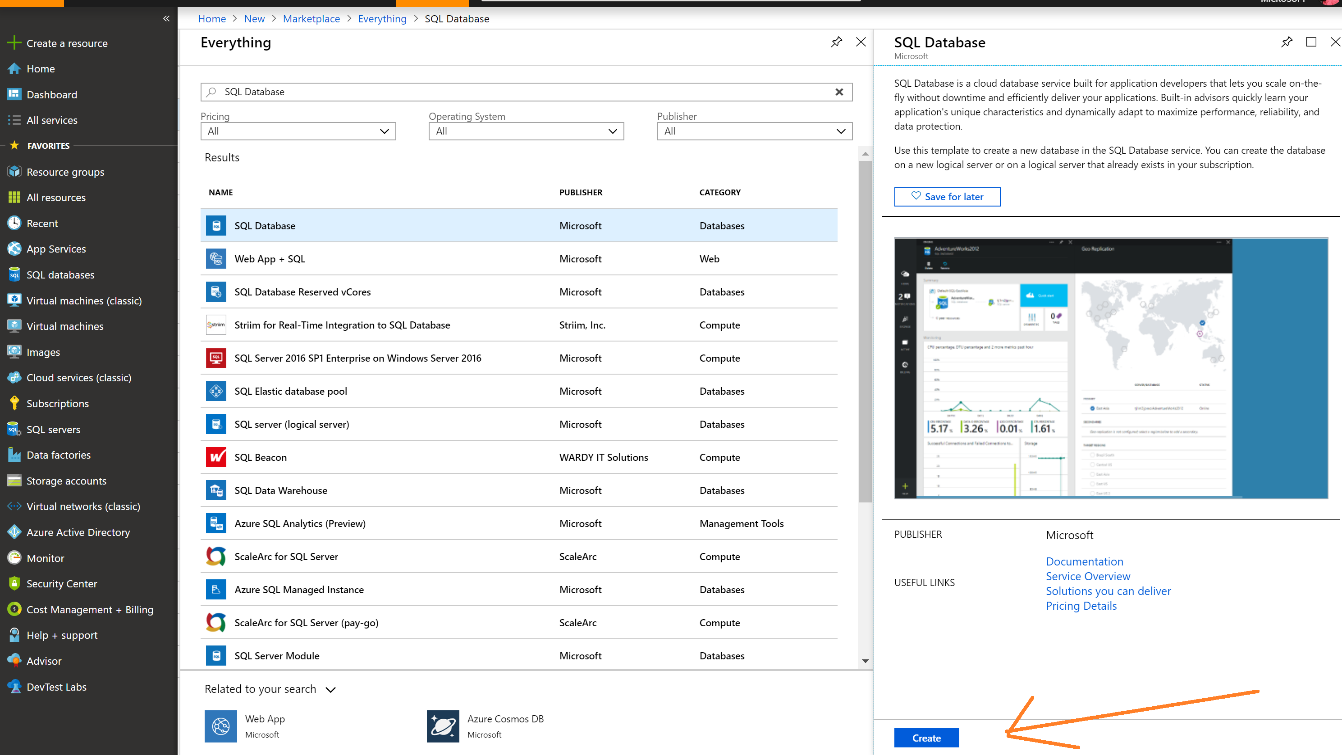


### Task 1 – Create Azure SQL Database

Go to Azure portal and then select ‘Create a resource’ 🡪 Search for SQL database 🡪 Select SQL Database as shown:



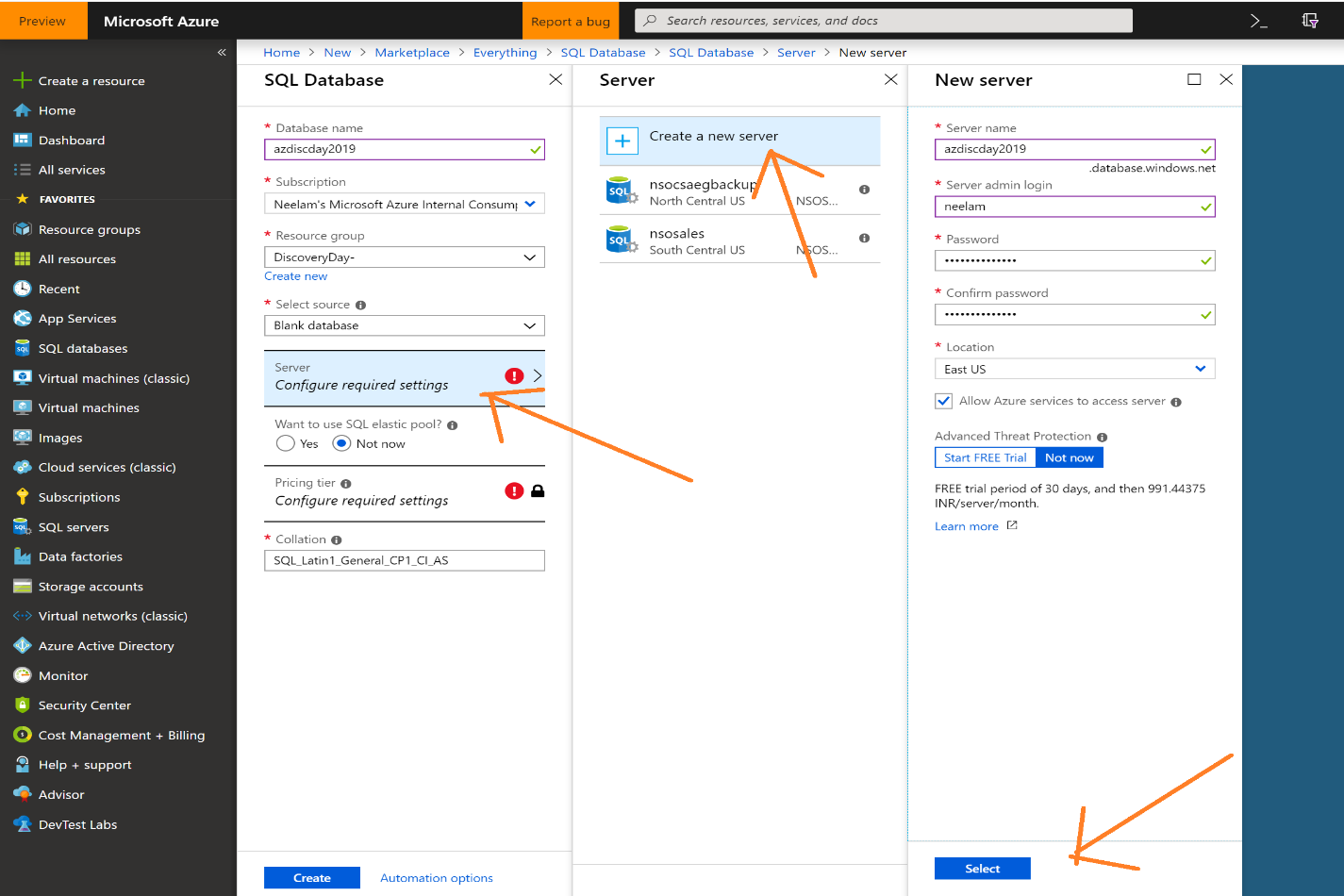
Please click create.



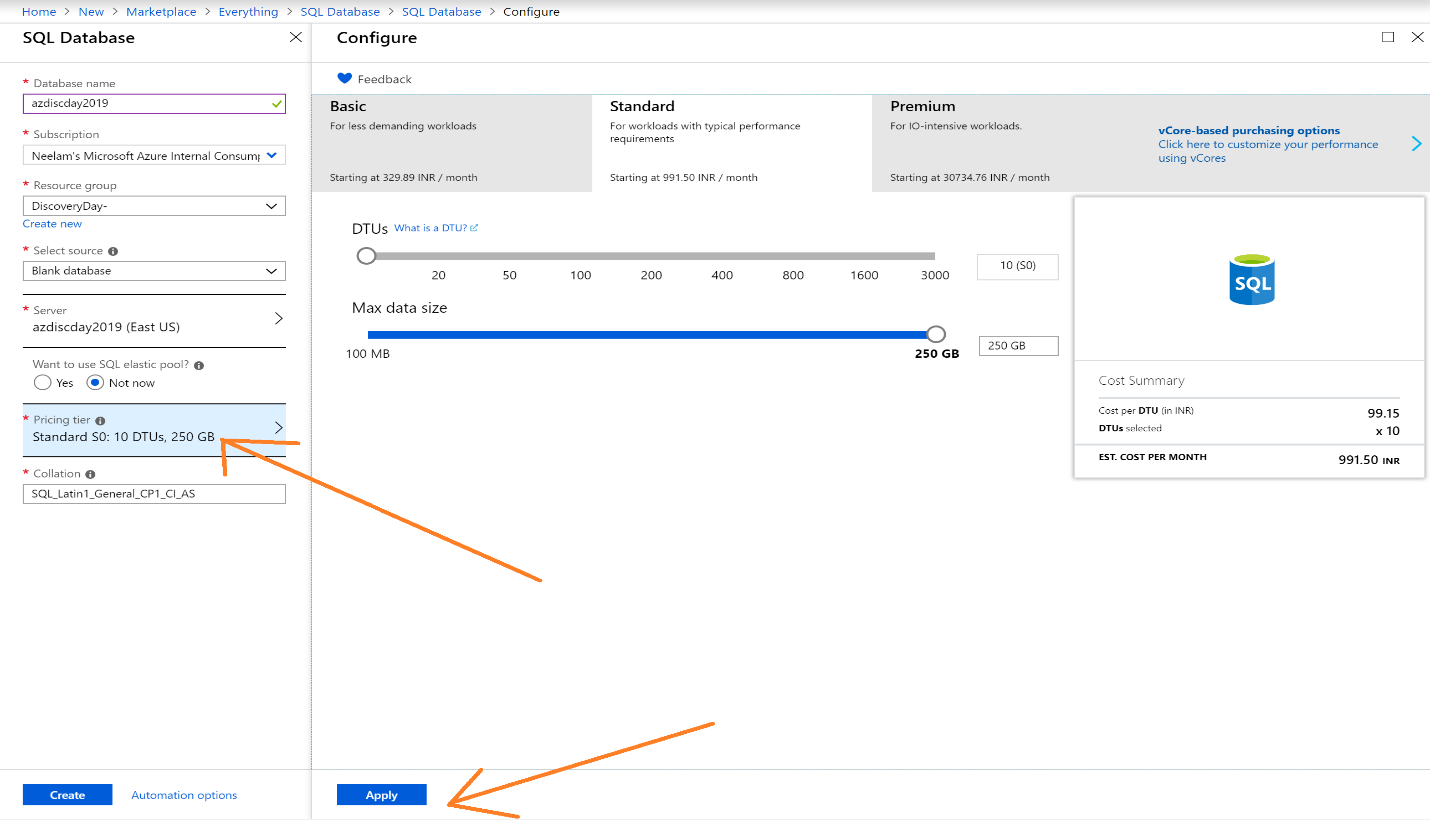
Enter Database name, select Azure subscription and select your existing resource group.

Select source as Blank Database. Then click on Server to create a new logical server.

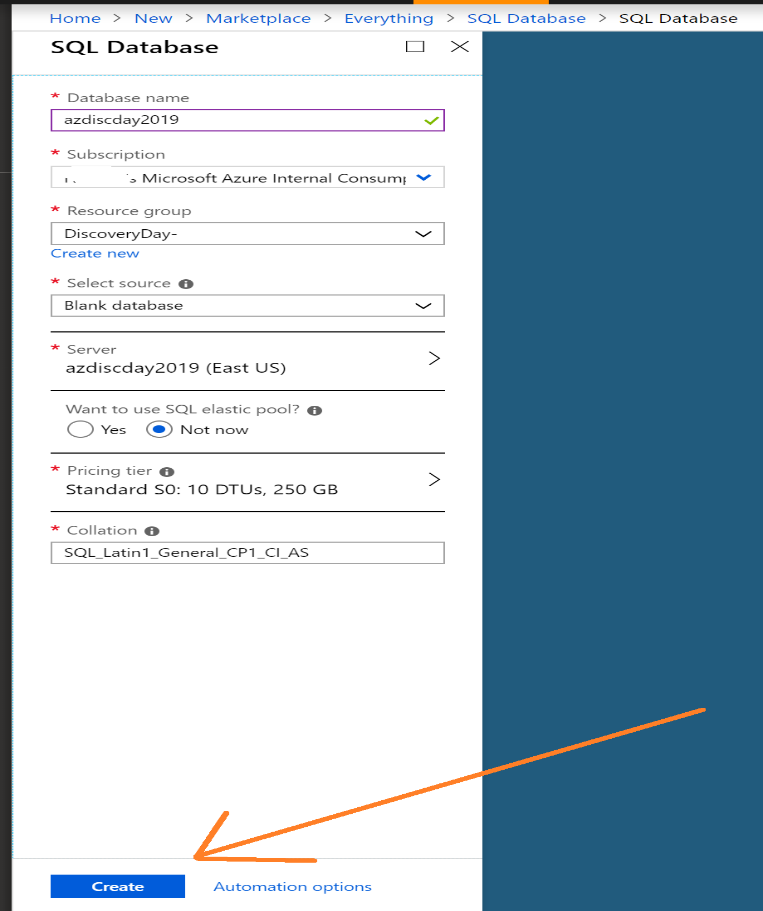
Enter details as shown and please take note of admin user name and password.



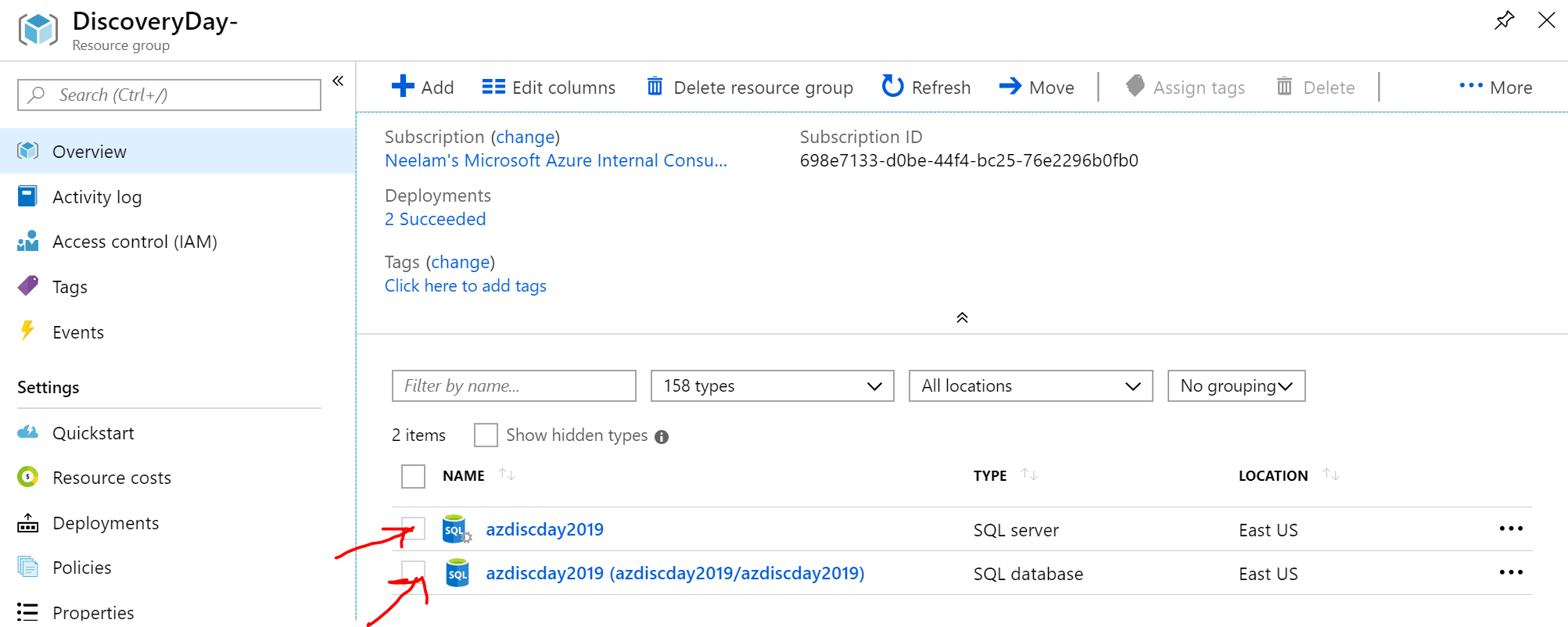
Click on the pricing tier and select Standard “S0” as default. Please click Apply.



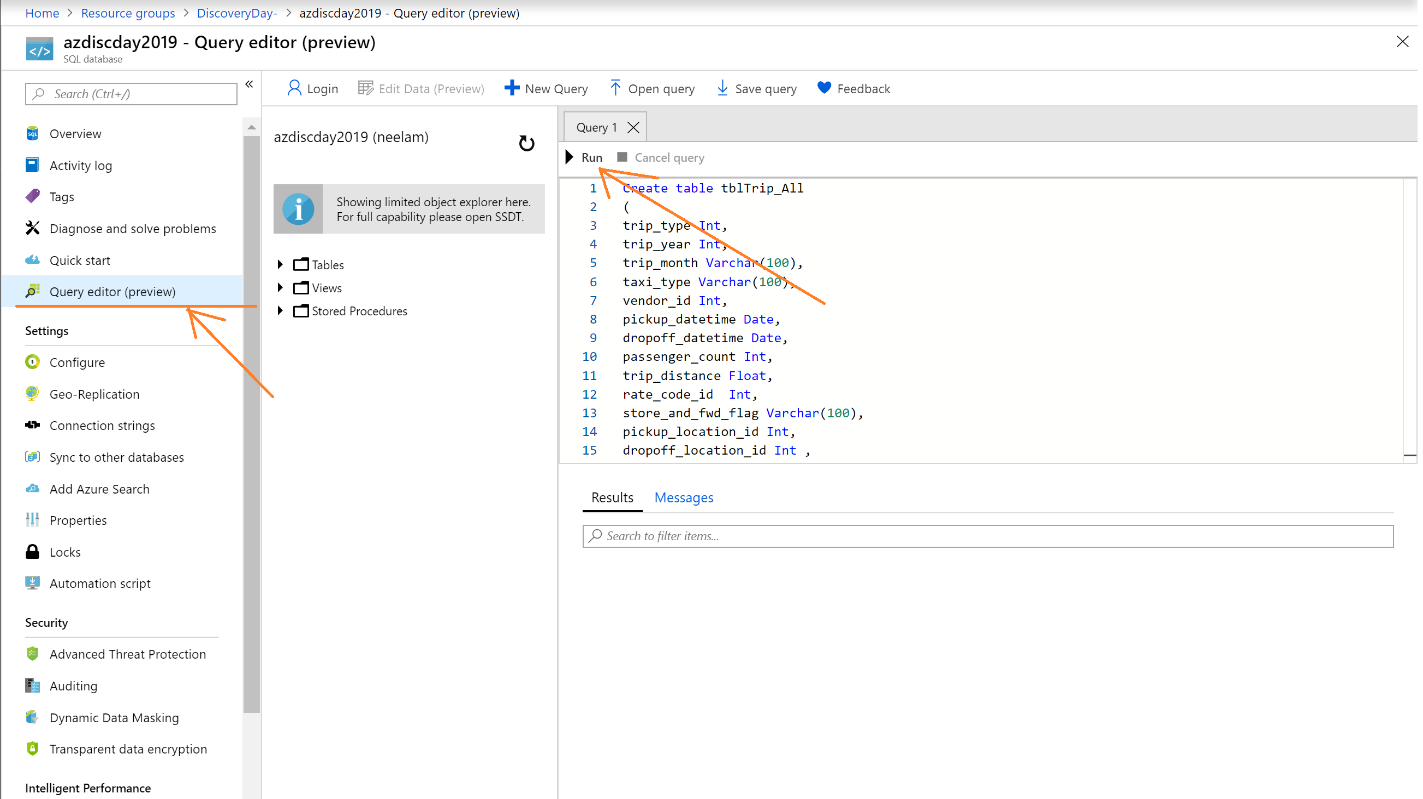
Please review the details and then click create as shown -



Go to resource group and you will see two new components added under your resource group – Azure SQL Database and logical Server.



Go to Azure SQL Database query editor and execute sql statement to create schema for the table.



Create table tblTrip\_All

(

trip\_type Int,

trip\_year Int,

trip\_month Varchar(100),

taxi\_type Varchar(100),

vendor\_id Int,

pickup\_datetime Date,

dropoff\_datetime Date,

passenger\_count Int,

trip\_distance Float,

rate\_code\_id Int,

store\_and\_fwd\_flag Varchar(100),

pickup\_location\_id Int,

dropoff\_location\_id Int ,

pickup\_longitude Varchar(100),

pickup\_latitude Varchar(100),

dropoff\_longitude Varchar(100),

dropoff\_latitude Varchar(100),

payment\_type Int,

fare\_amount Float,

extra Float ,

mta\_tax Float ,

tip\_amount Float ,

tolls\_amount Float ,

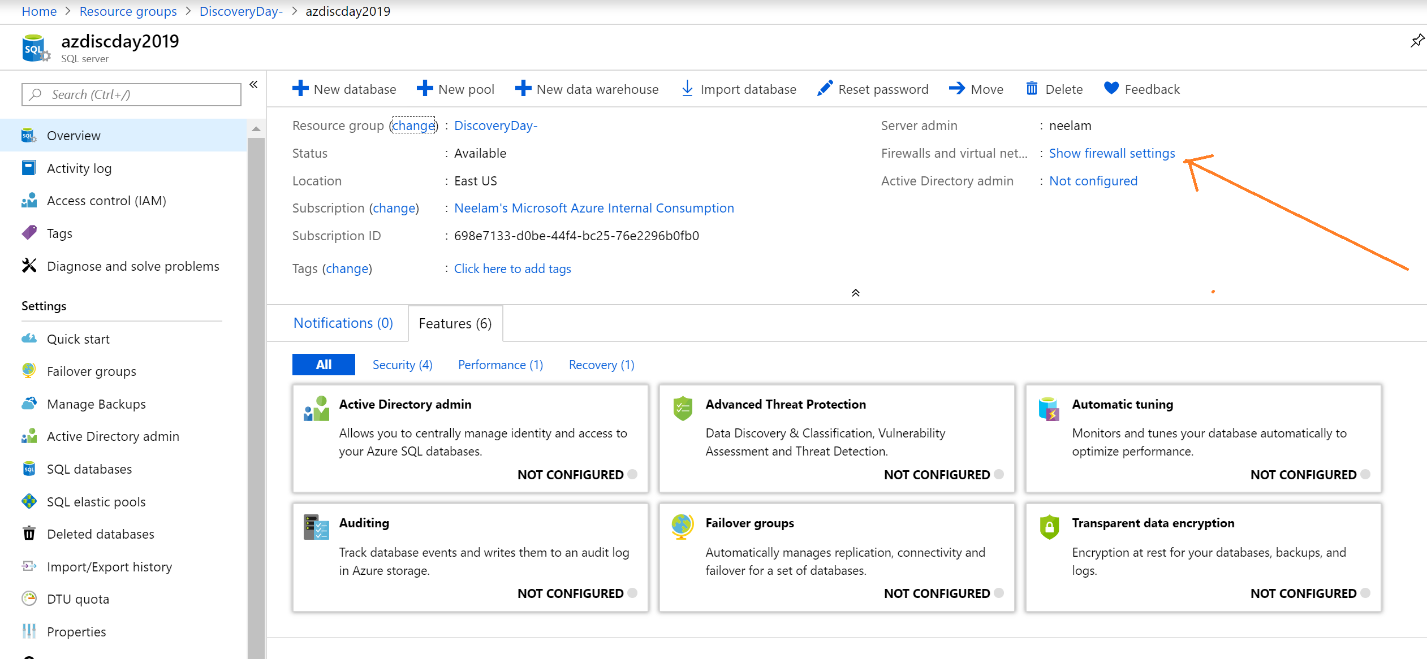
improvement\_surcharge Float ,

ehail\_fee Float ,

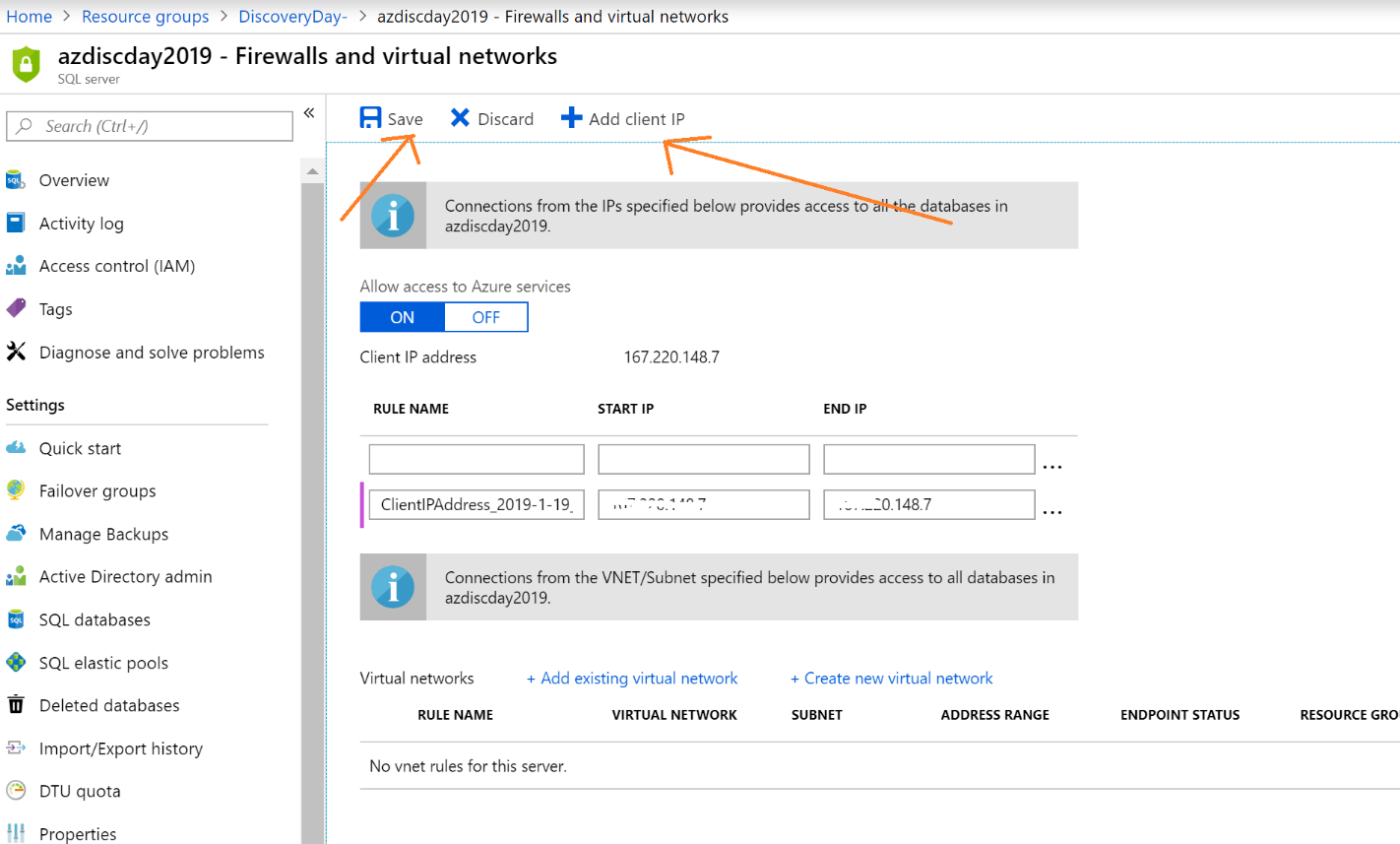
total\_amount Float

)

By Default, Azure services can connect to each other. If you want to connect SQL Server to on-premises SQL Server management studio or Power BI desktop then you will need to whitelist your IP address.

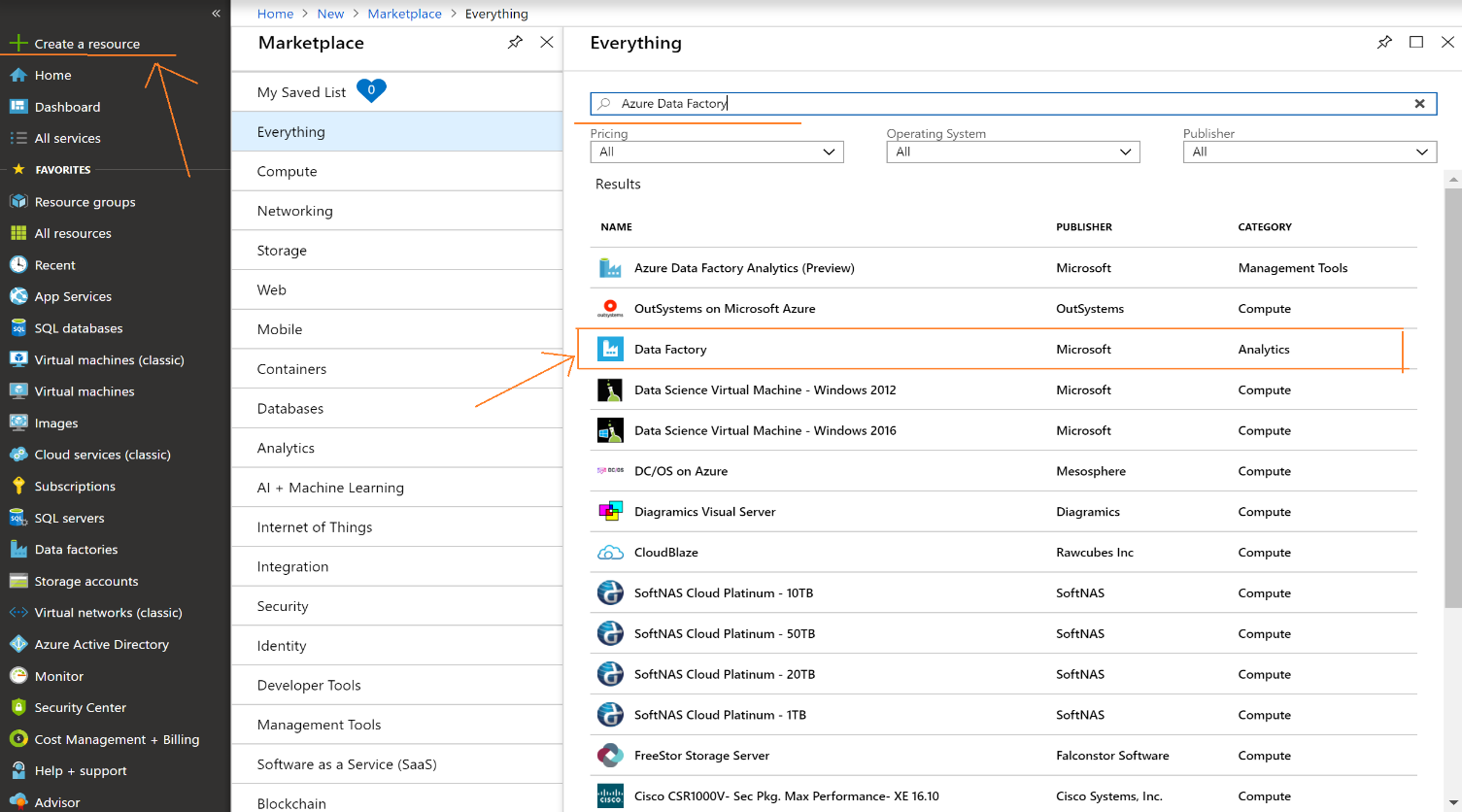


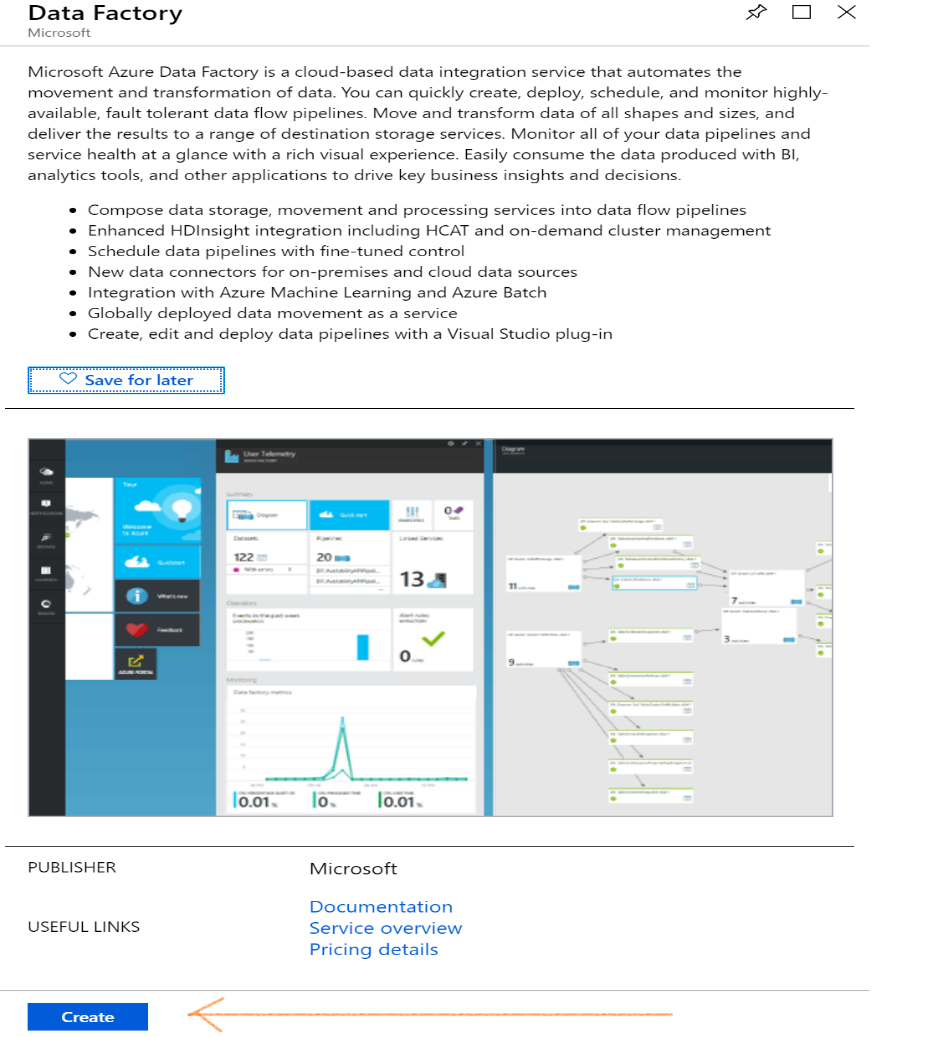
Click on Add Client IP and then click save. It will automatically whitelist your IP to use Azure Sql Database. It may take up to 5 minutes for this change to take effect.



### Task 2 – Create Azure Data Factory Project

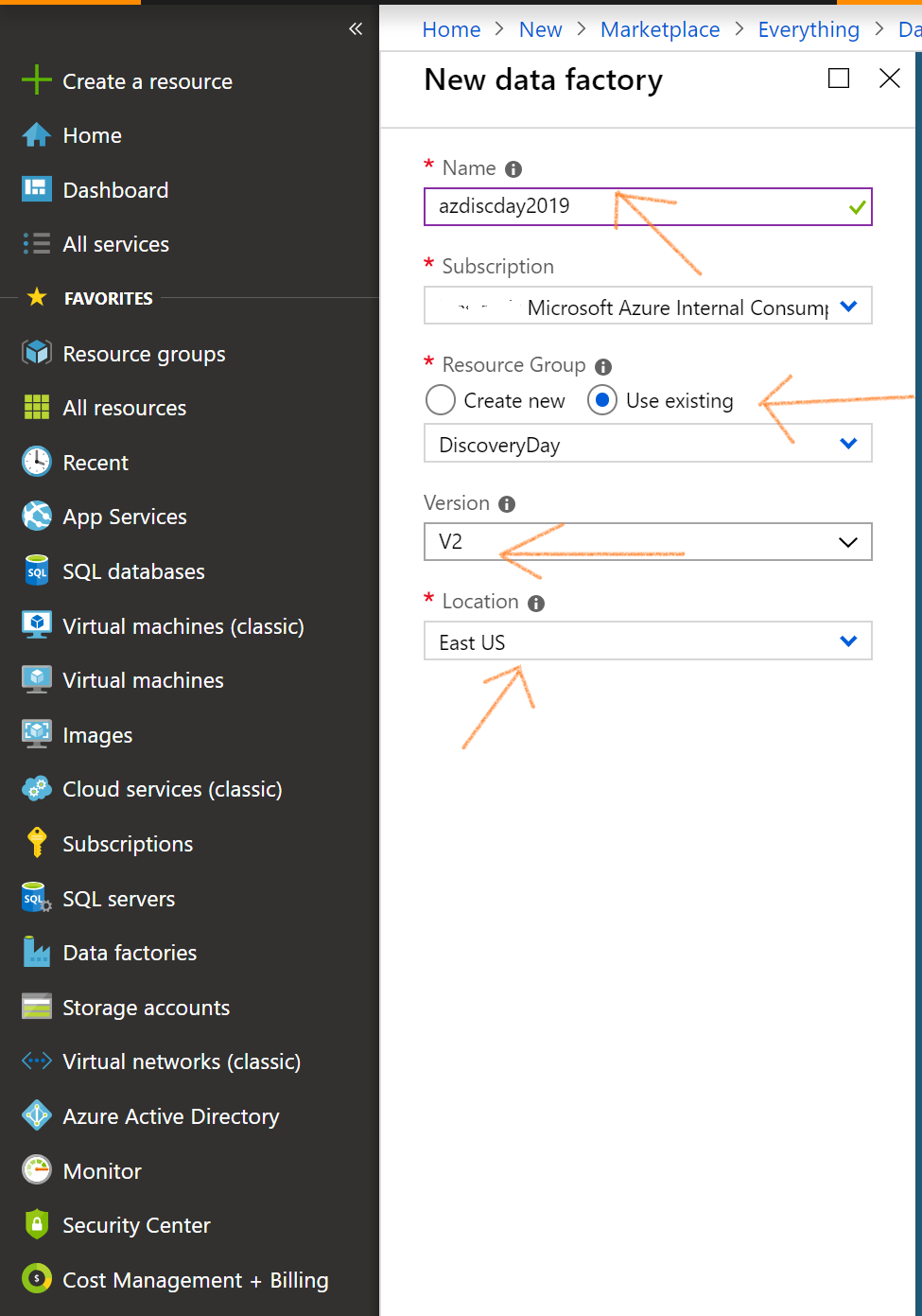
Go to Azure portal and search for Data Factory.

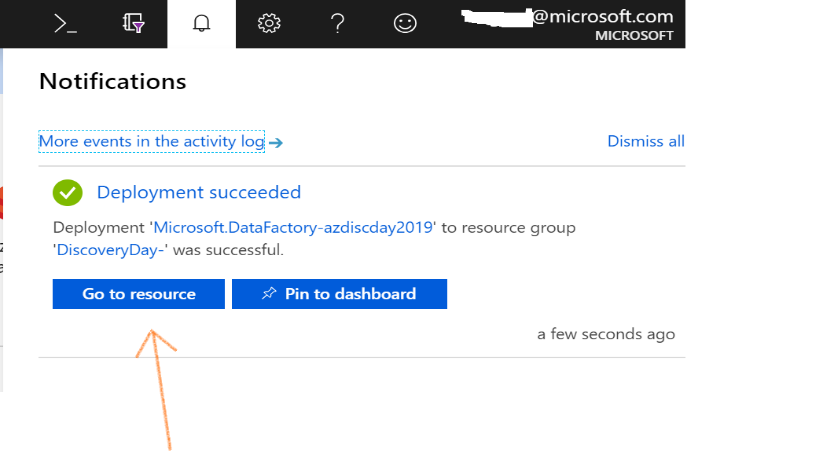




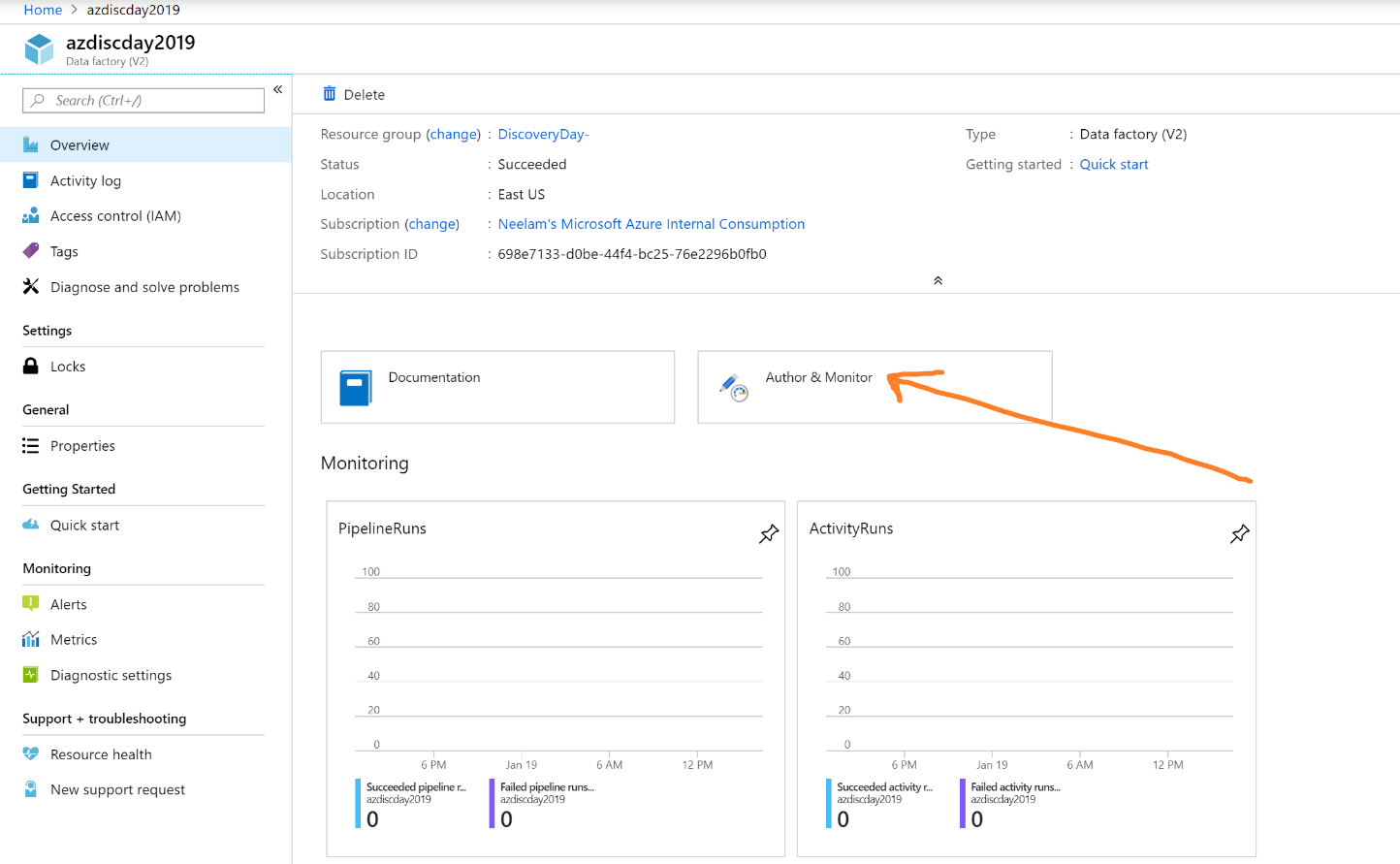
Please enter Data Factory Name, Select Azure subscription and then select existing resource group.

Select version as V2 and region as same what you selected for other resources.





When you click on go to resource, it will take you to the Data Factory project as shown – on this screen, click Author & Monitor.



Author & Monitor gives two options to create Azure Data Factory copy activity –

1. Using copy data wizard.
2. Using create pipeline option.

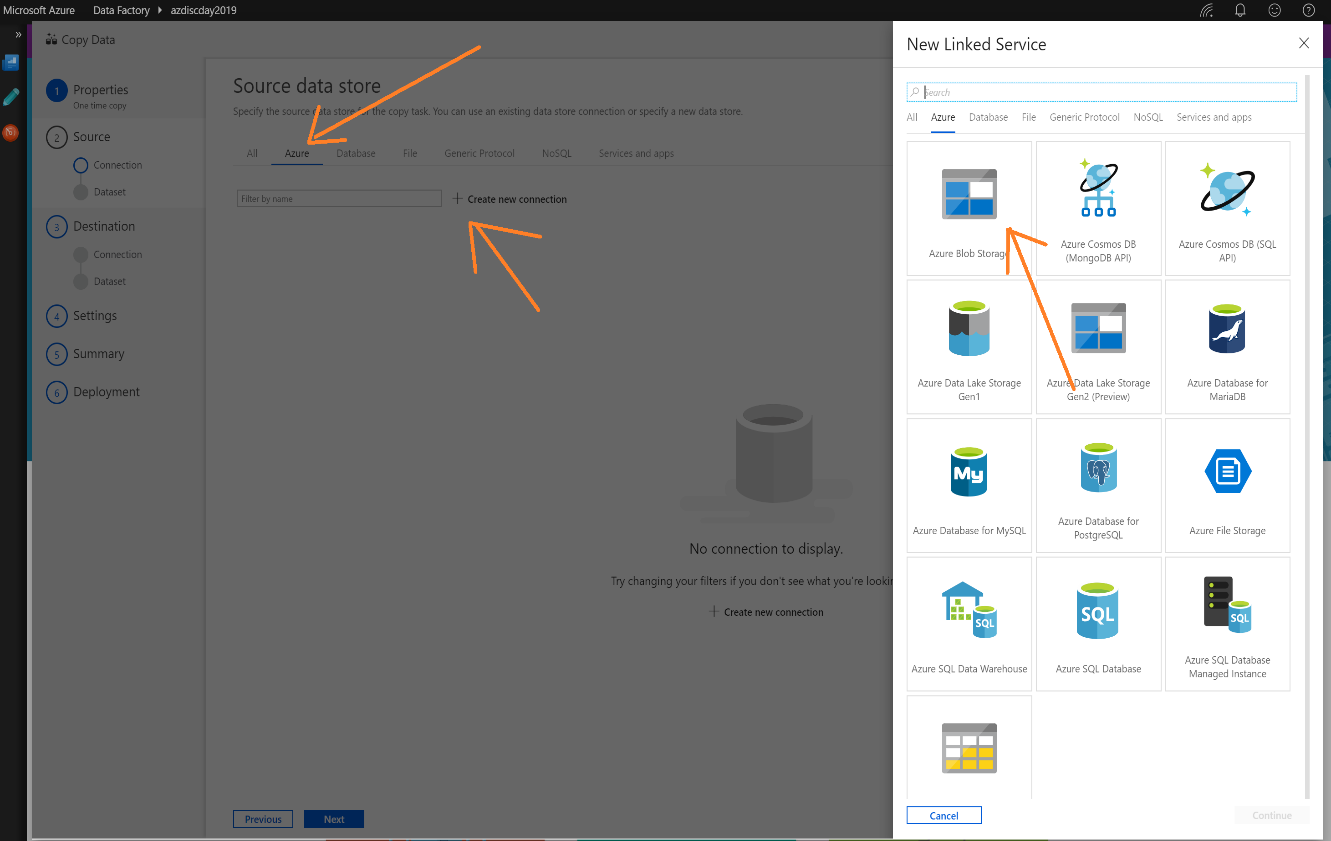
Option -1 – Copy Data (Wizard)



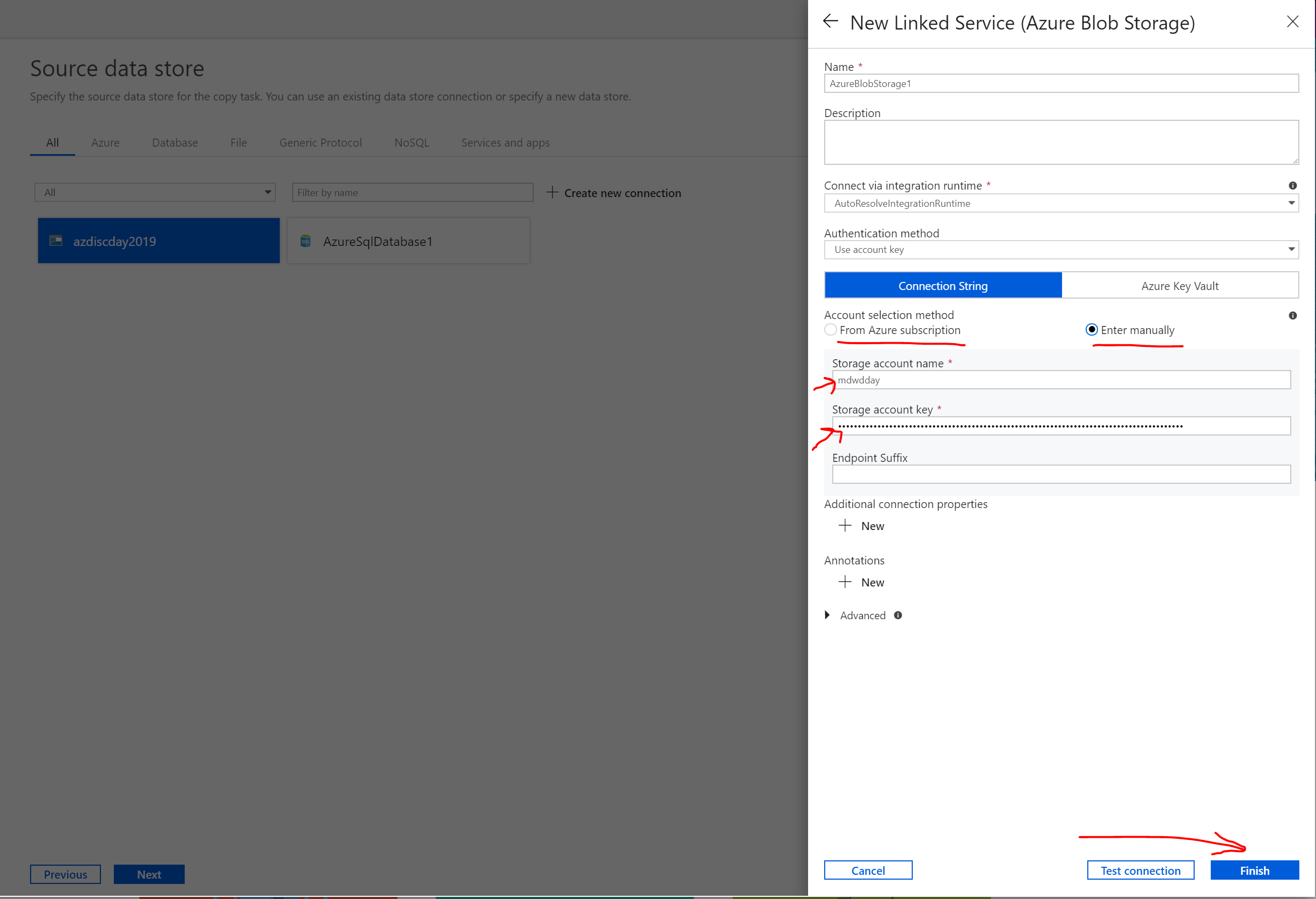
Enter Task Name and then select run once now for this lab. But if you want to schedule this activity then select Run regularly on schedule option and select the frequency and data & time to run.

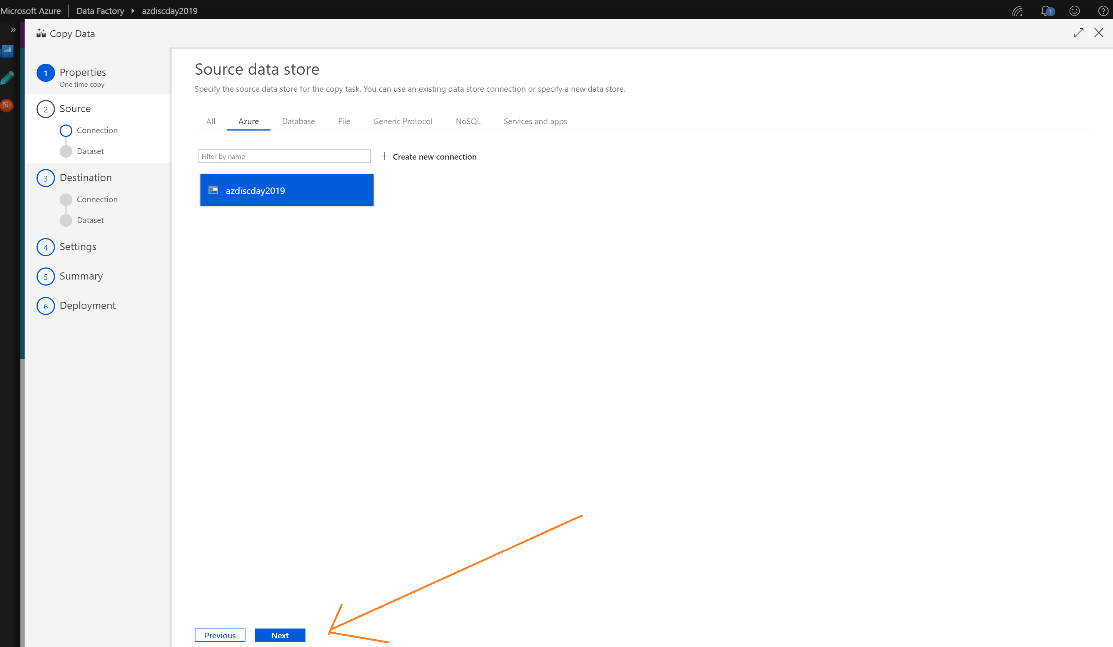


In this lab, we are going to copy data from Azure blob storage to Azure SQL DB. Select Source data source as Azure blob storage.



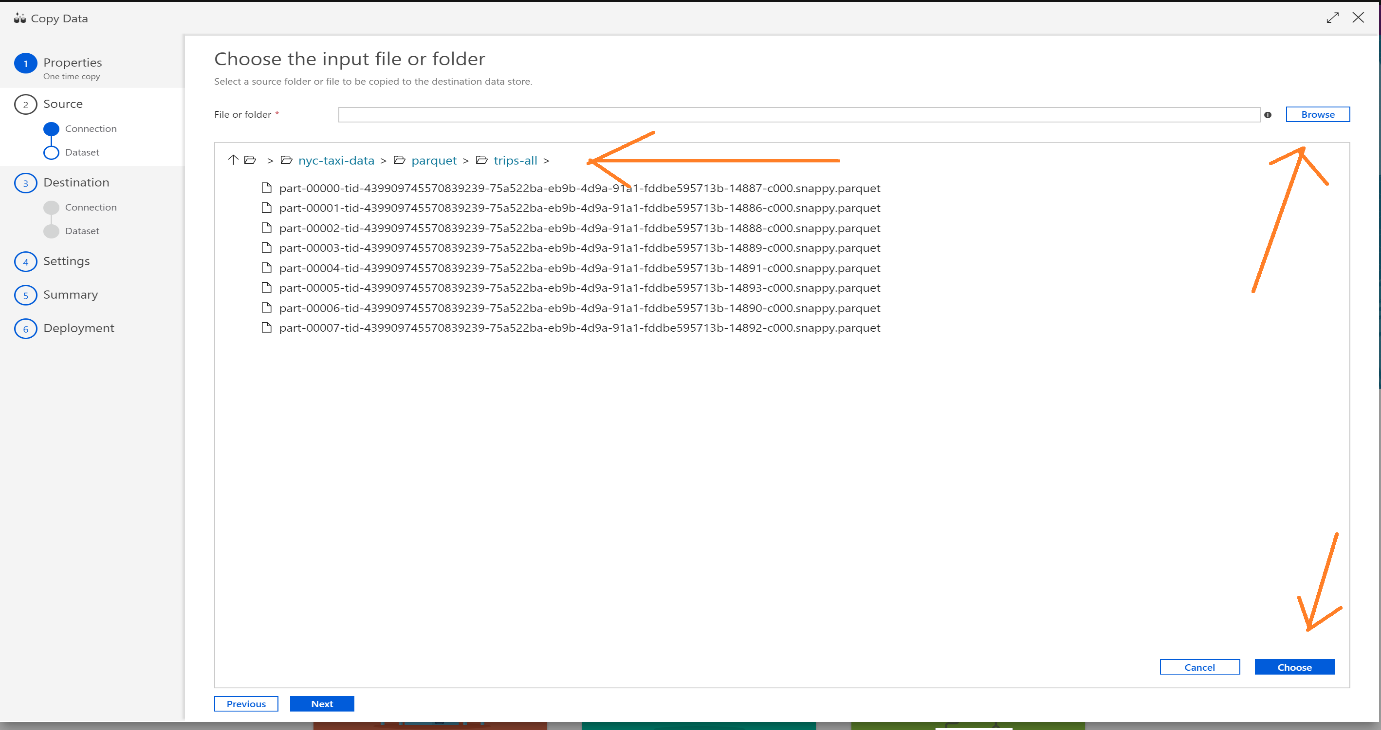
If the source blob storage is in the same subscription then select – ‘From Azure subscription’ else select ‘Enter manually’ and enter the details accordingly.



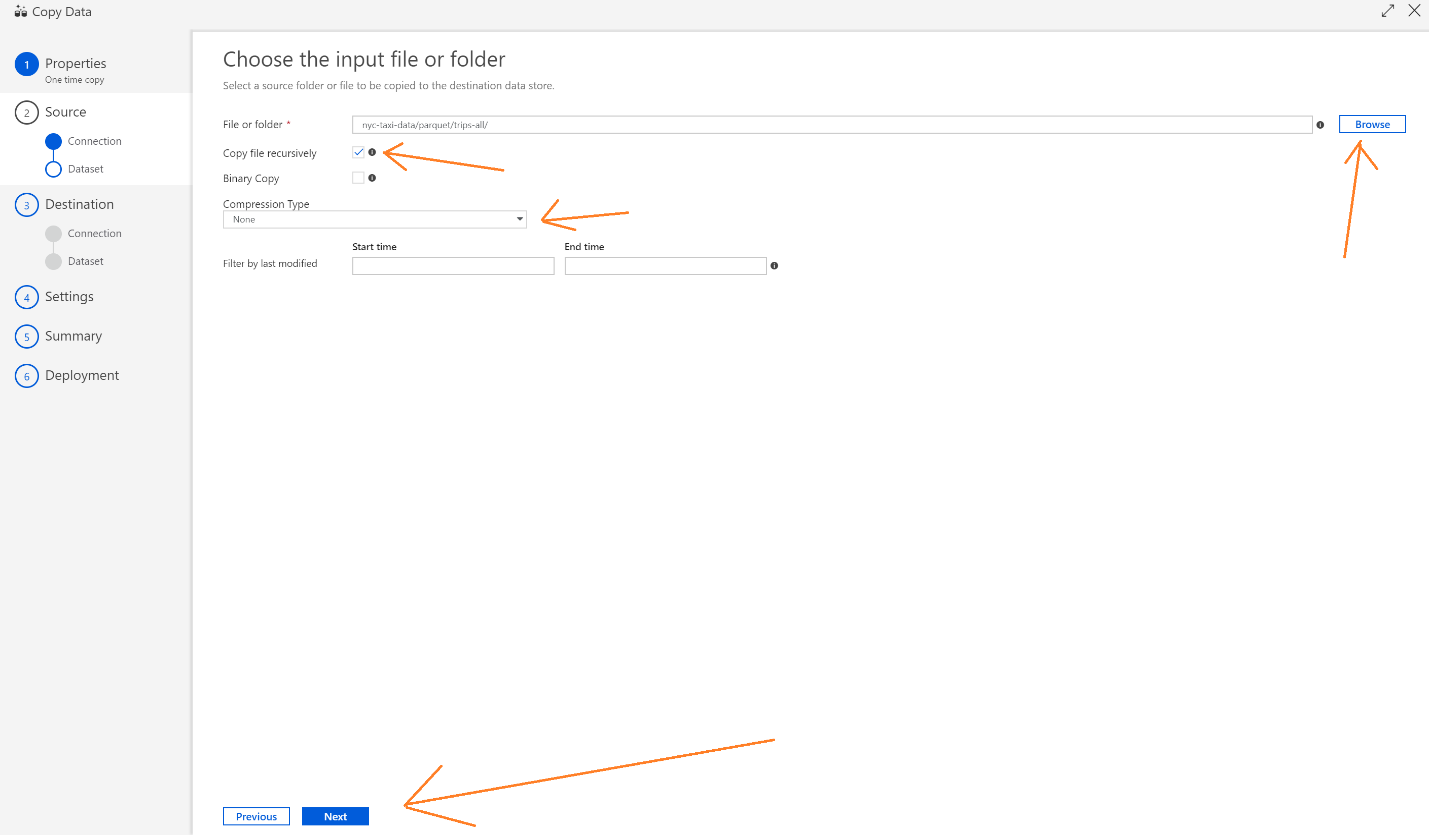


Once you configure blob storage, next step is to select container/folder where the parquet files are.

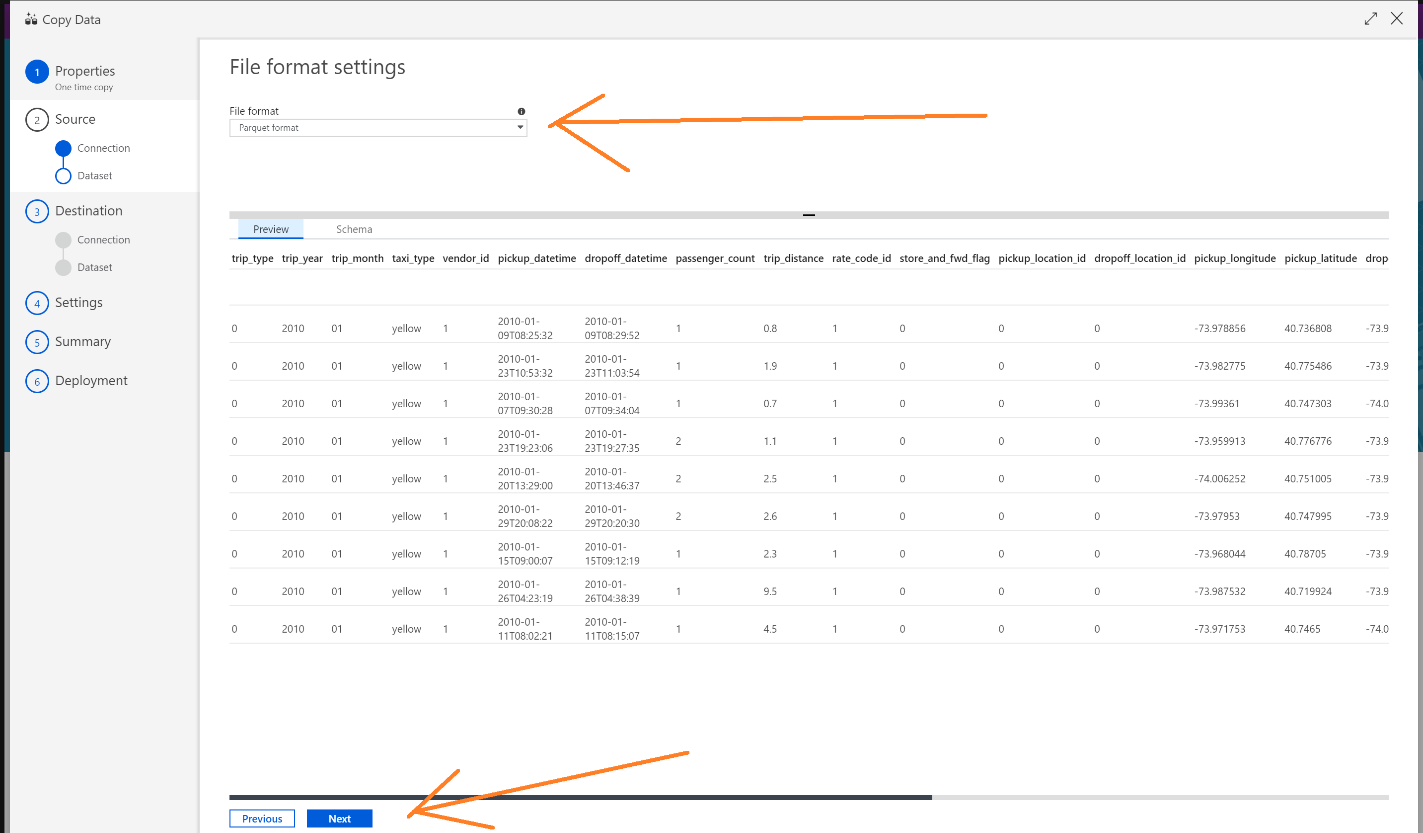
Click Browse and navigate to the folder where you have all files.



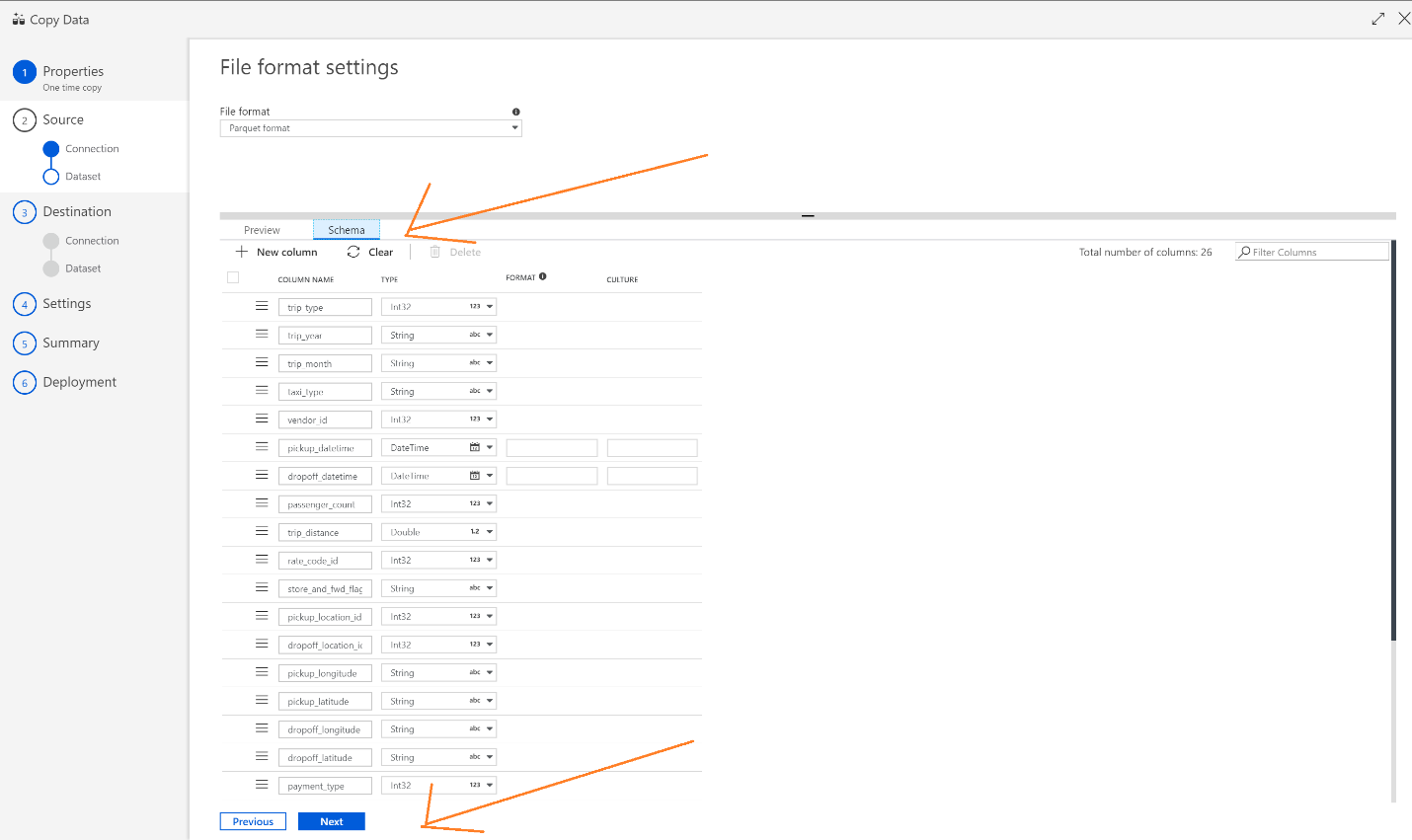
Select ‘Copy File Recursively’ and then next



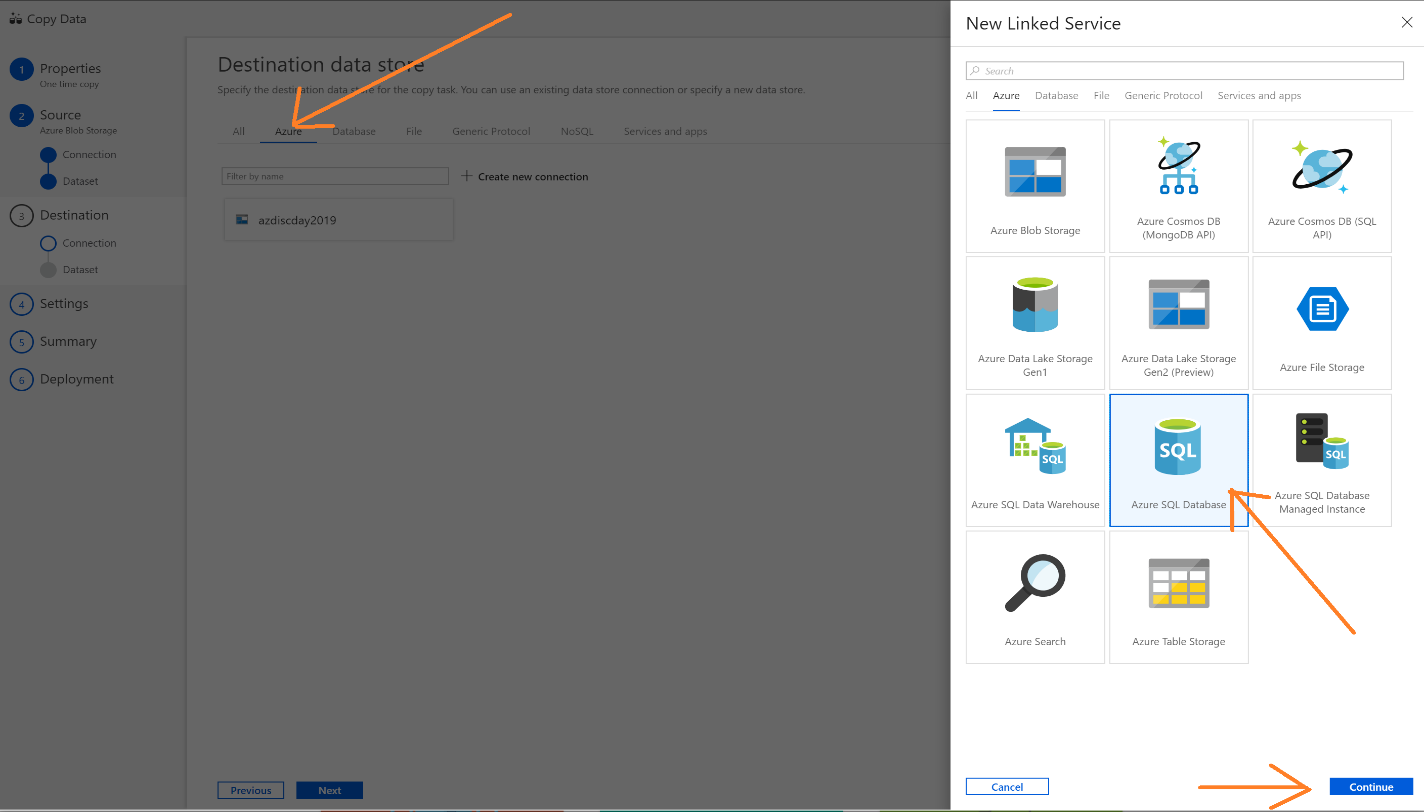
Select file format as parquet and review data in preview.



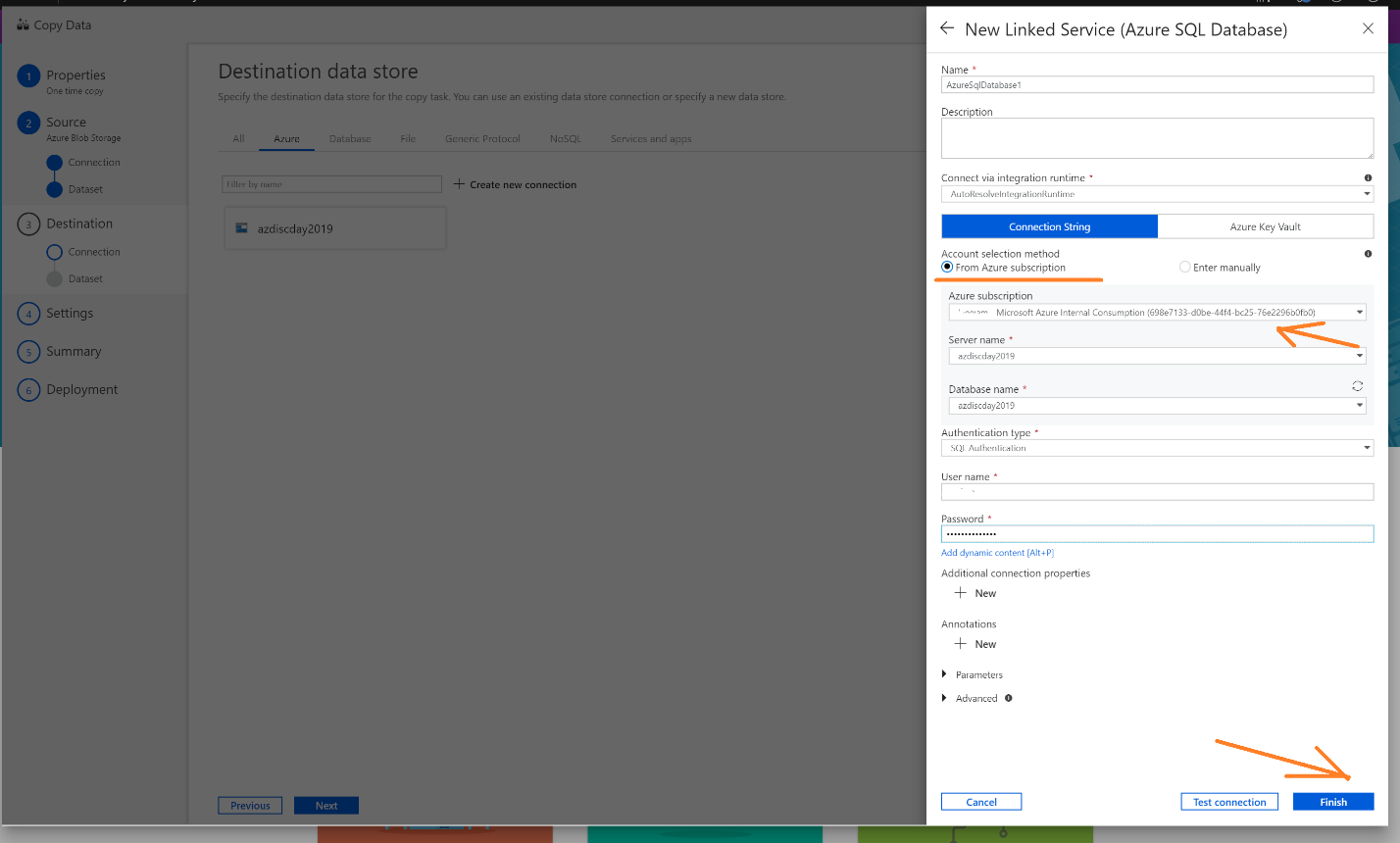
Please also have a look at schema by clicking on schema tab.

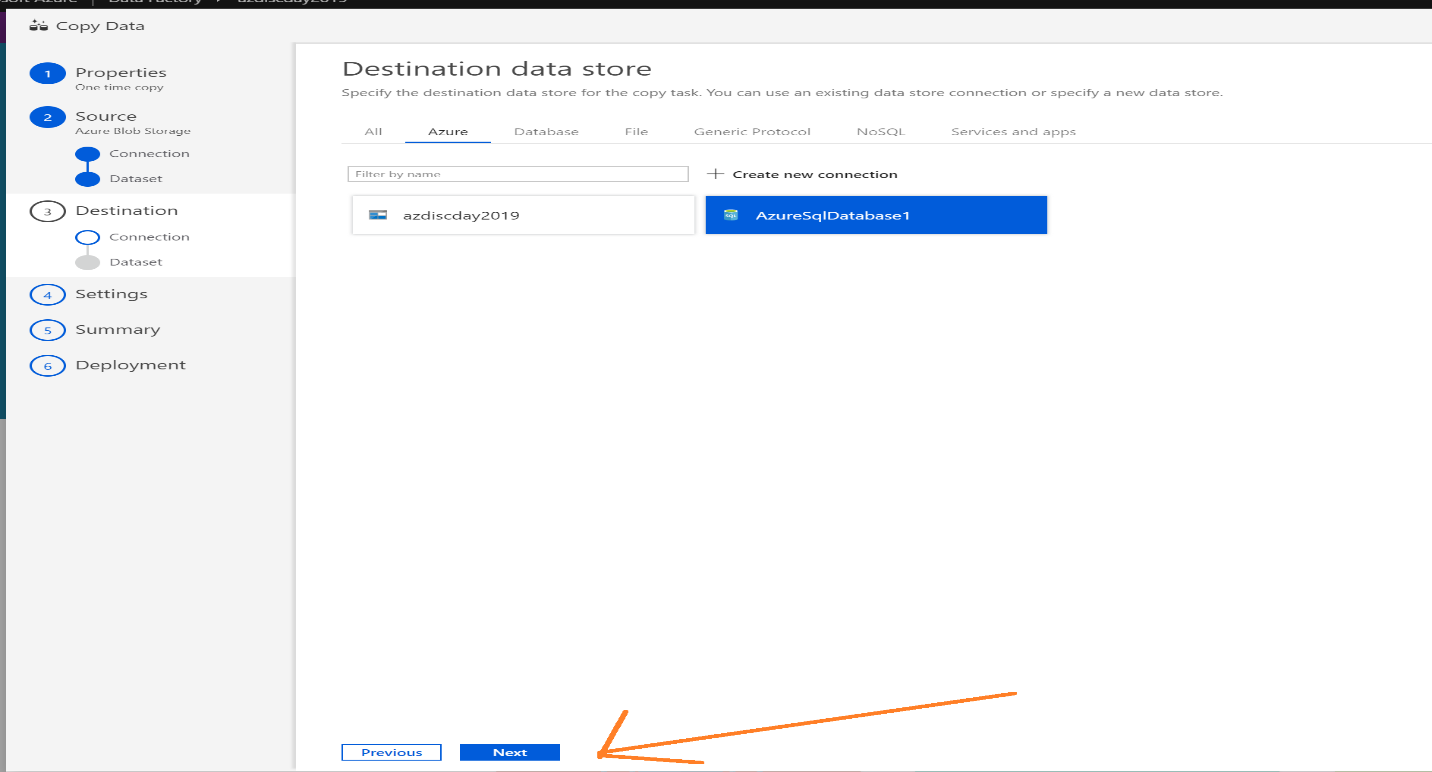


Next step is to configure destination – Azure SQL Database in this case.

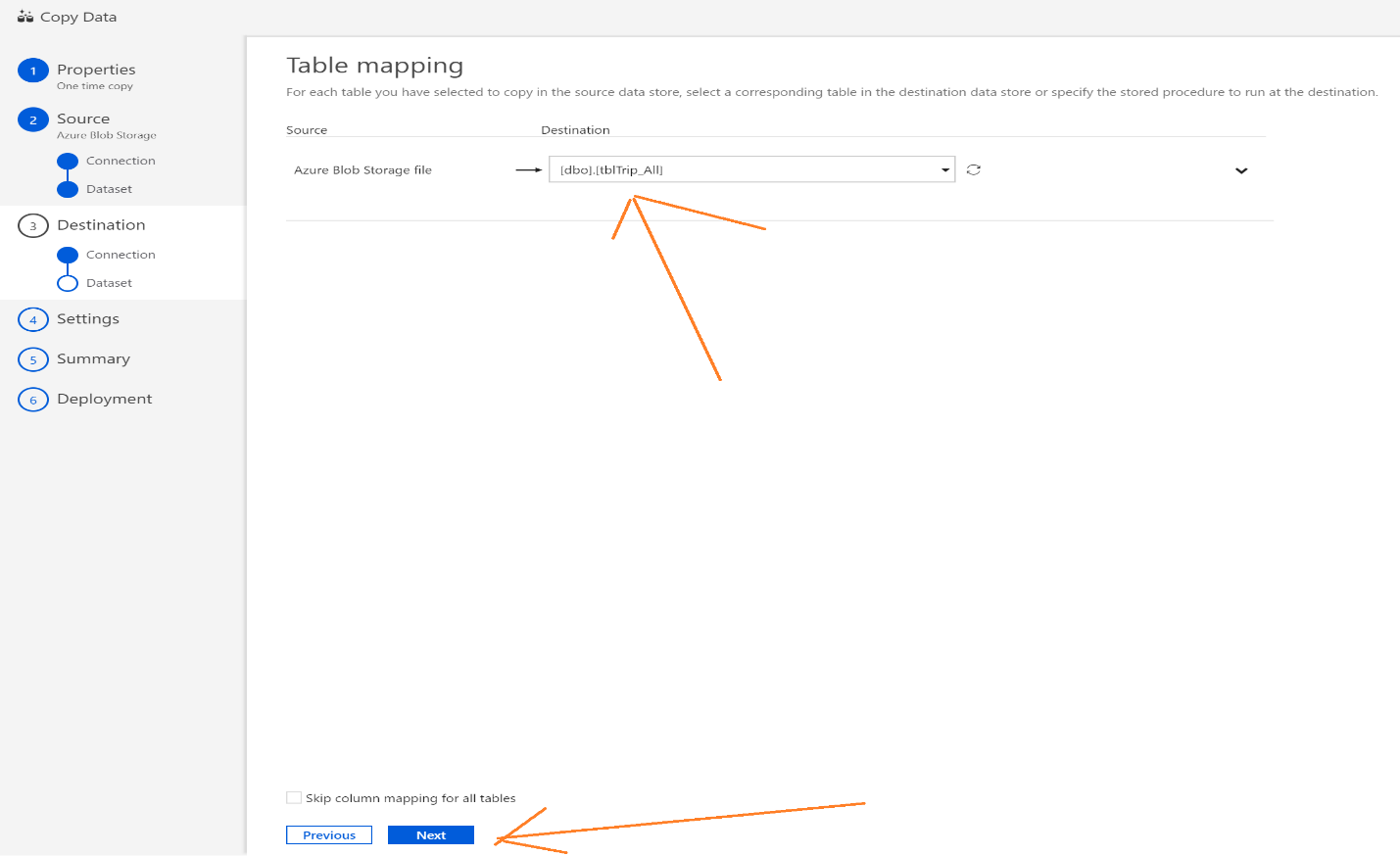


Select the option from Azure subscription and select your database.

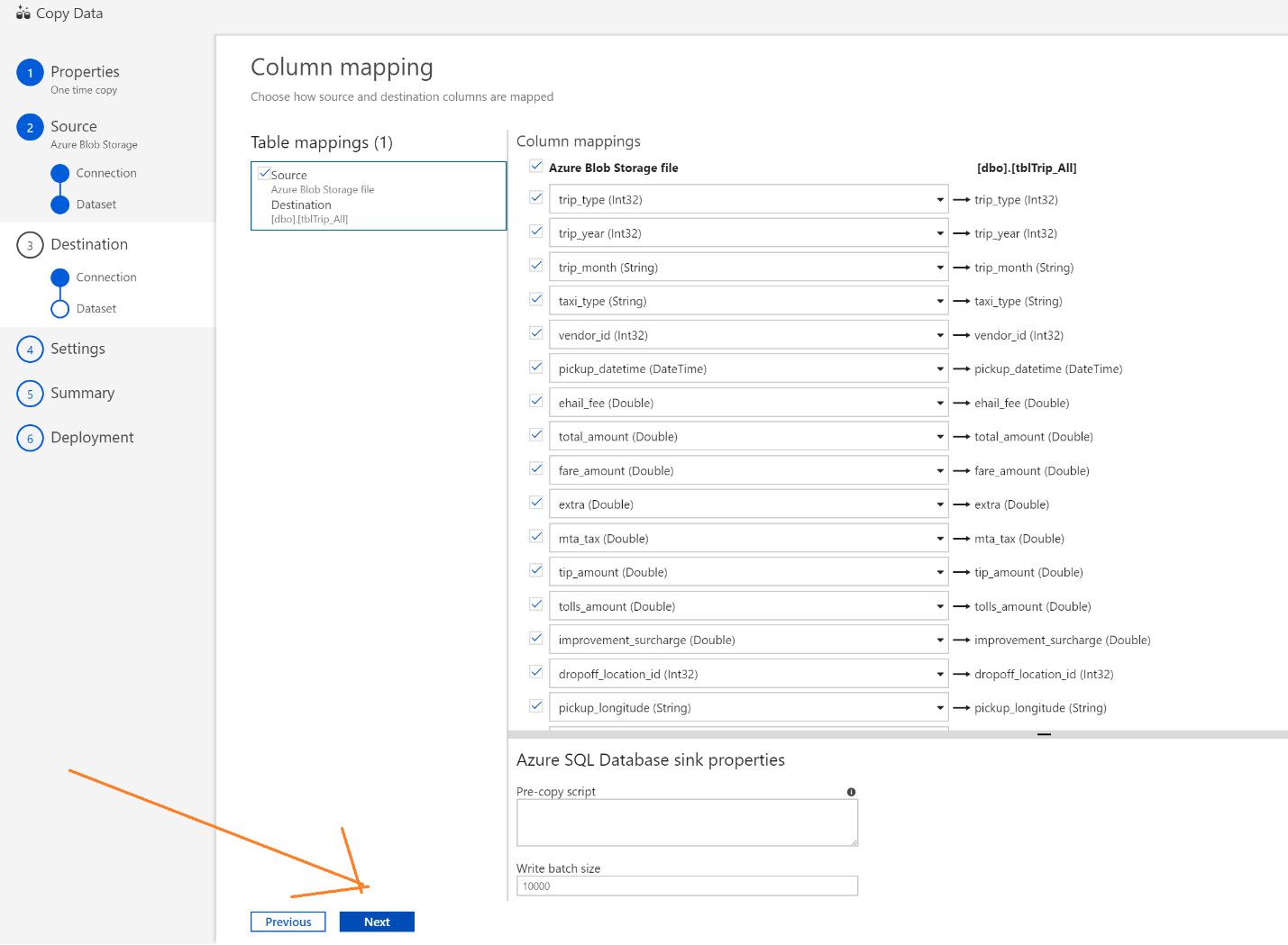




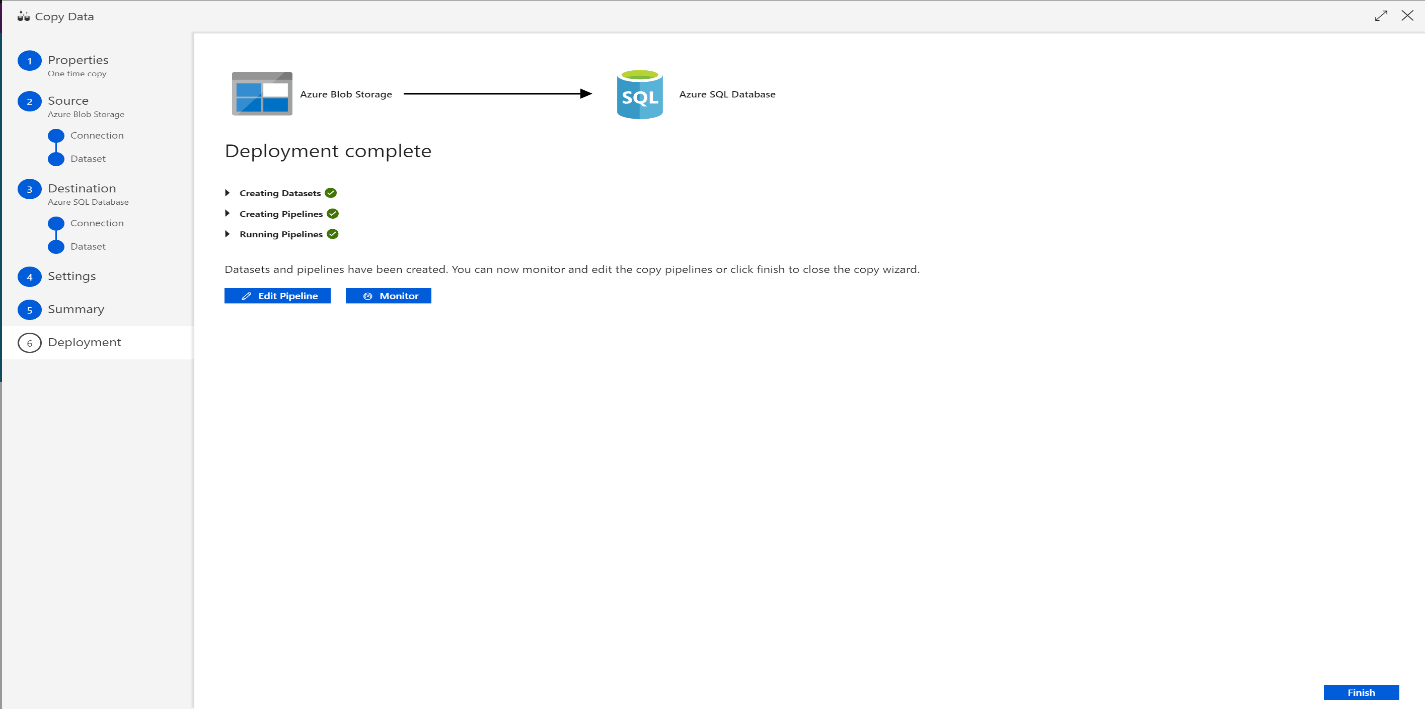
Next screen is table mapping – select the table in Azure SQL Database where you want to copy this data.

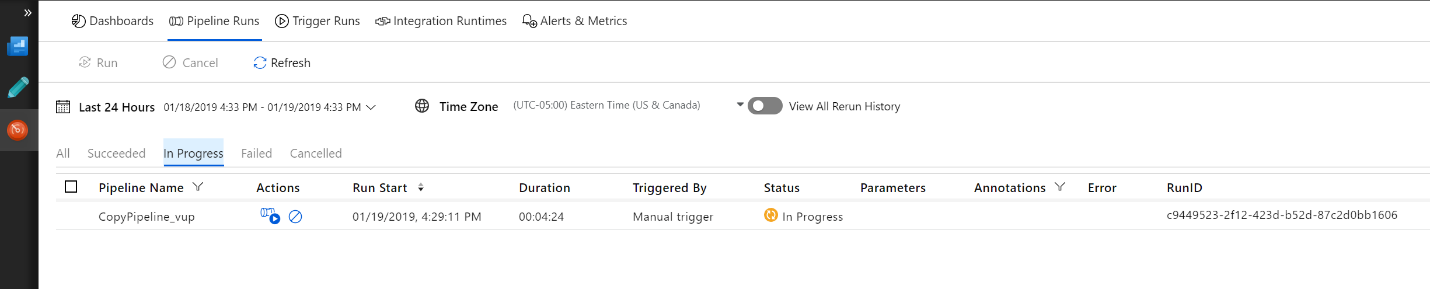


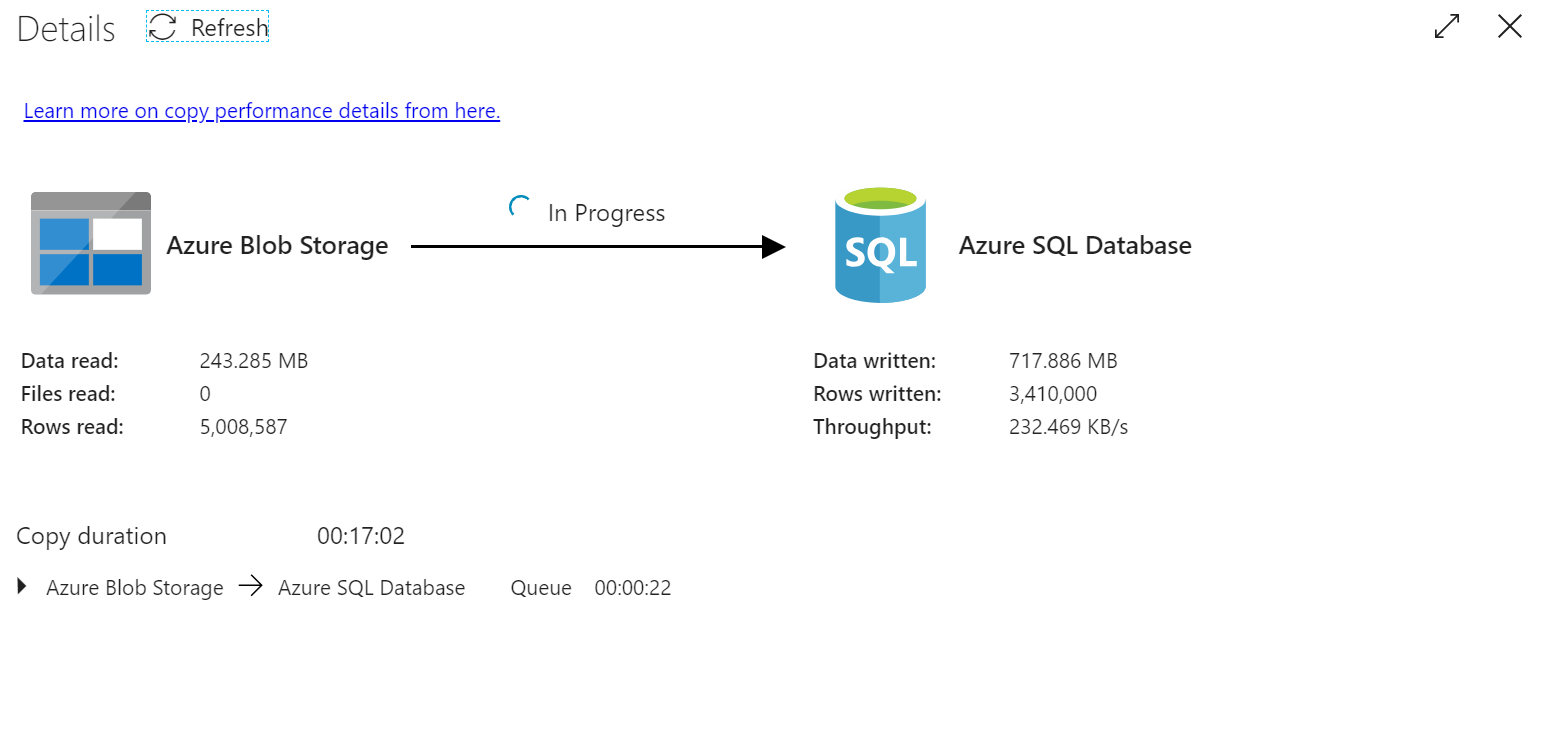
Please review columns mappings and click next.



Now Azure Data factory copy activity is running, and you can monitor the progress by clicking on monitor button.

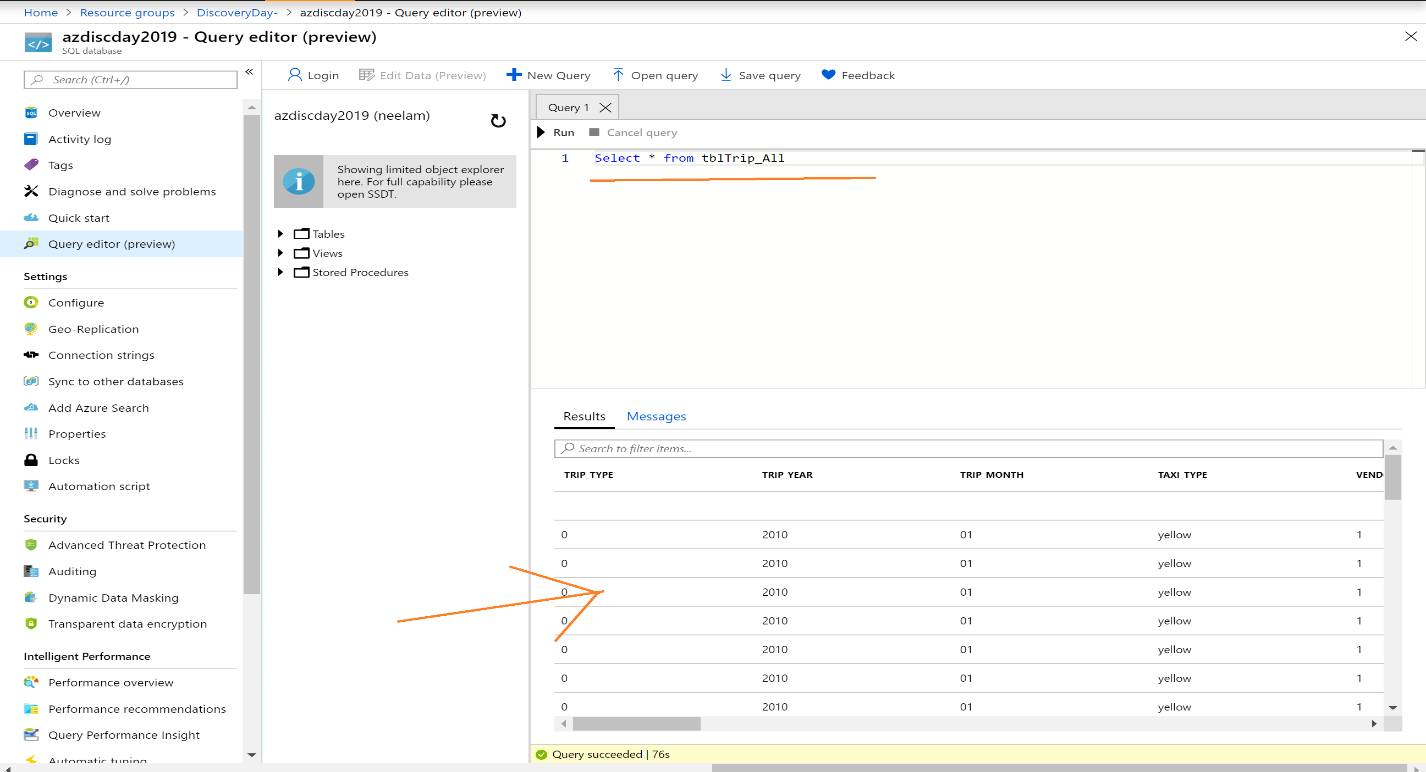






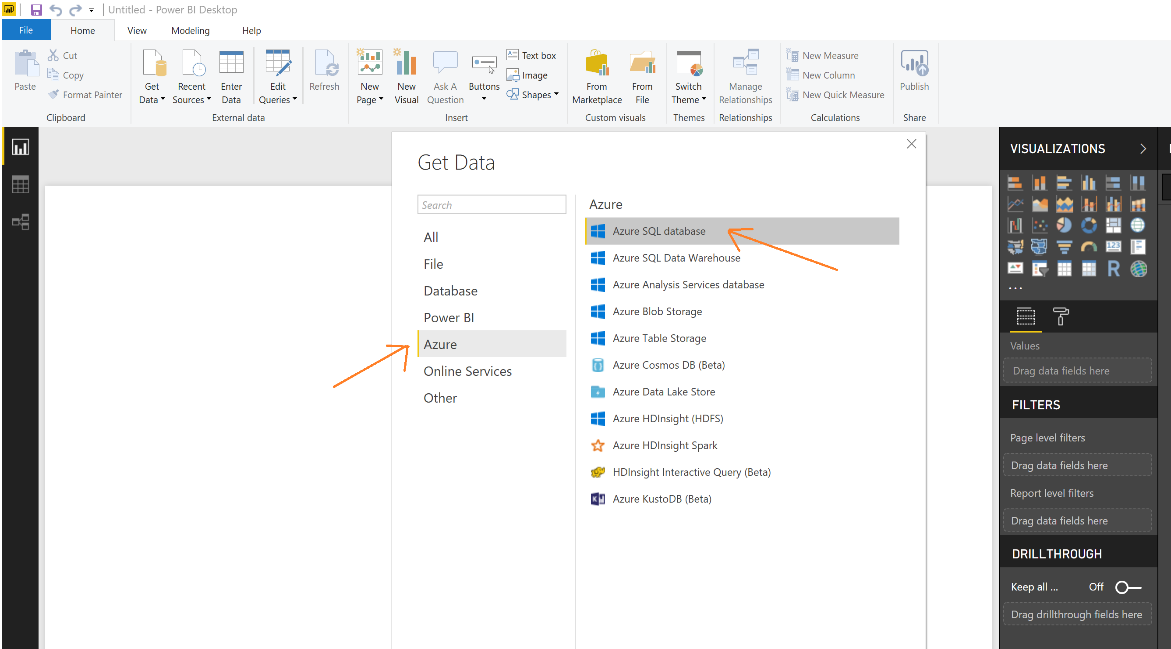
To see data in SQL Server tables. To go query editor under Azure SQL database on the Azure portal. And enter select query and then click run.

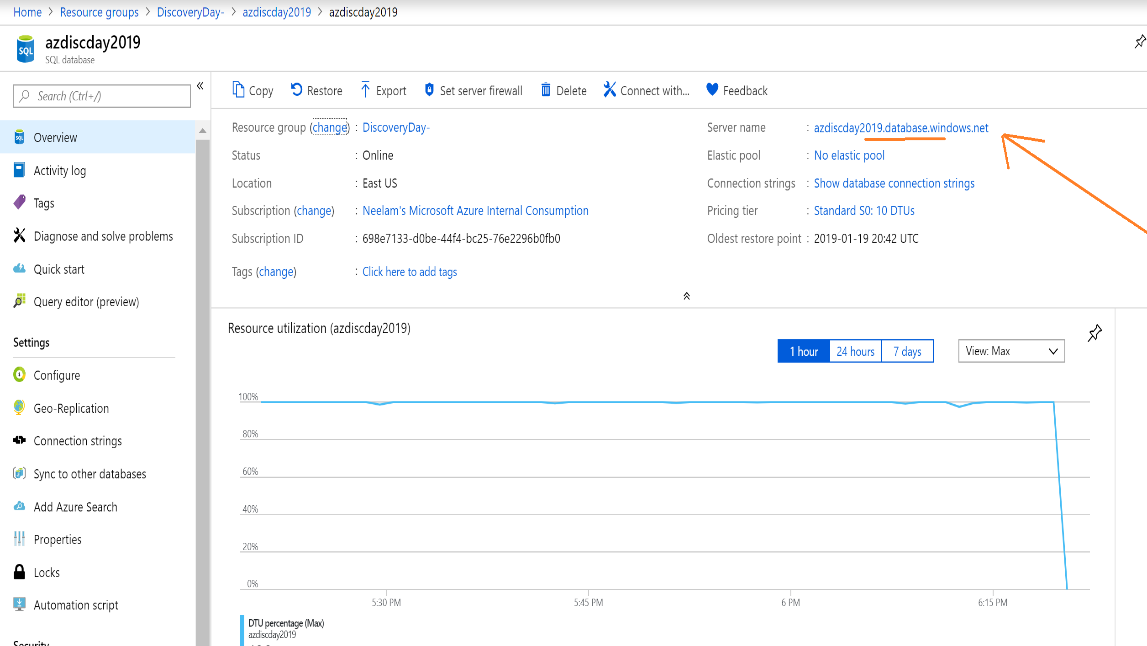
Select \* from tblTrip\_All



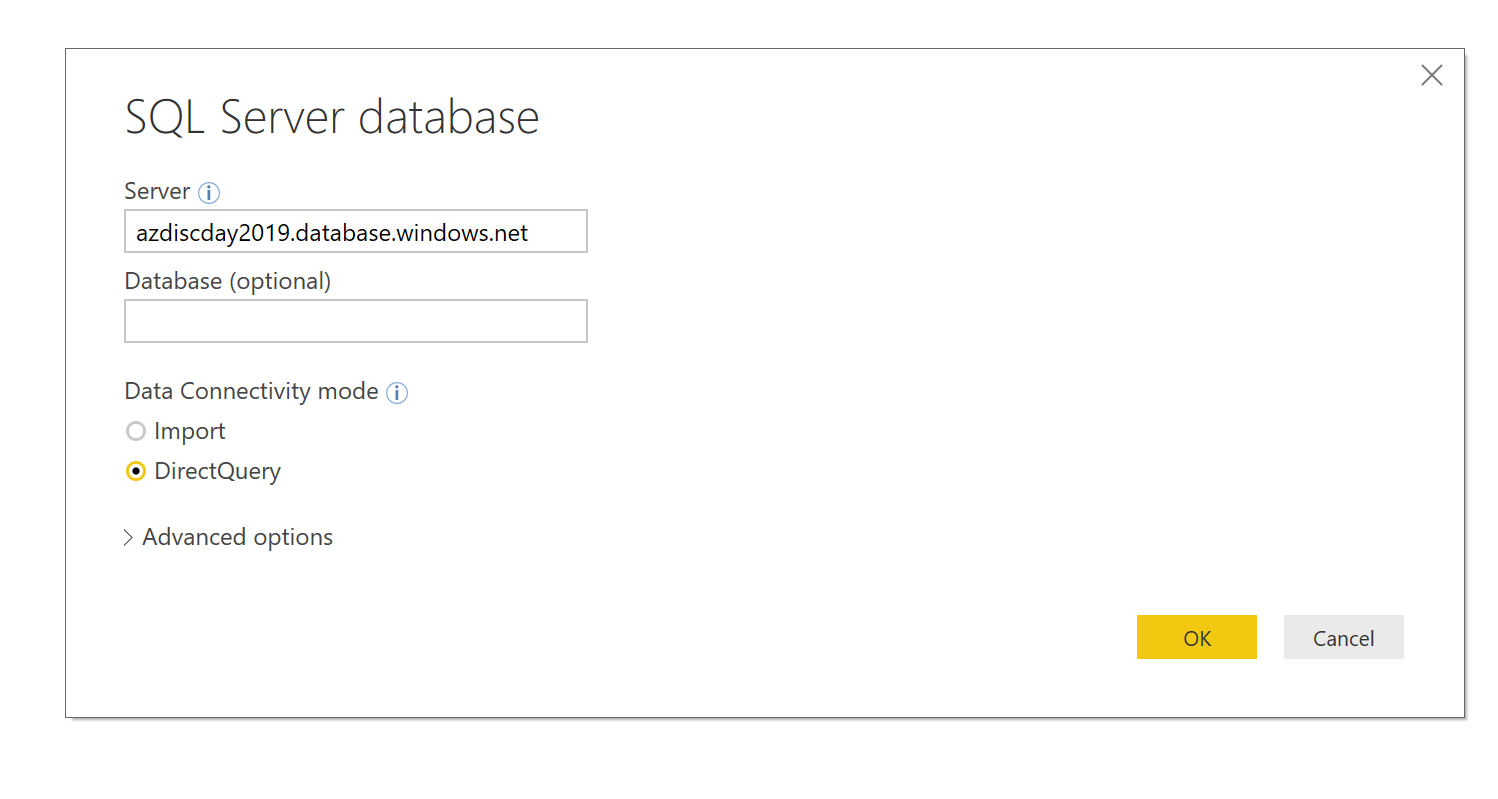
### Task 3 – Create Power BI Reports

Open Power BI desktop application and connect to Azure sql database.

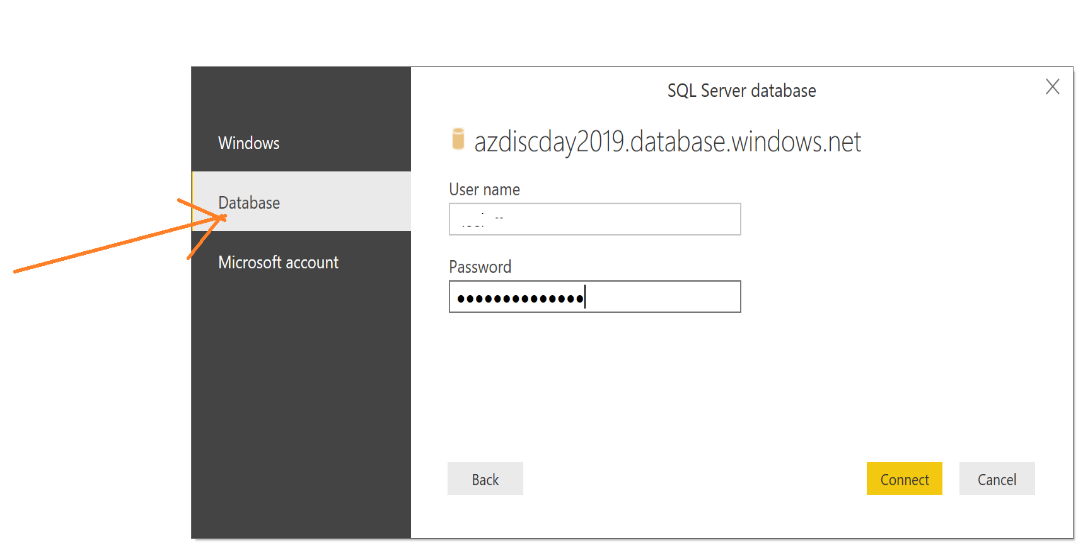




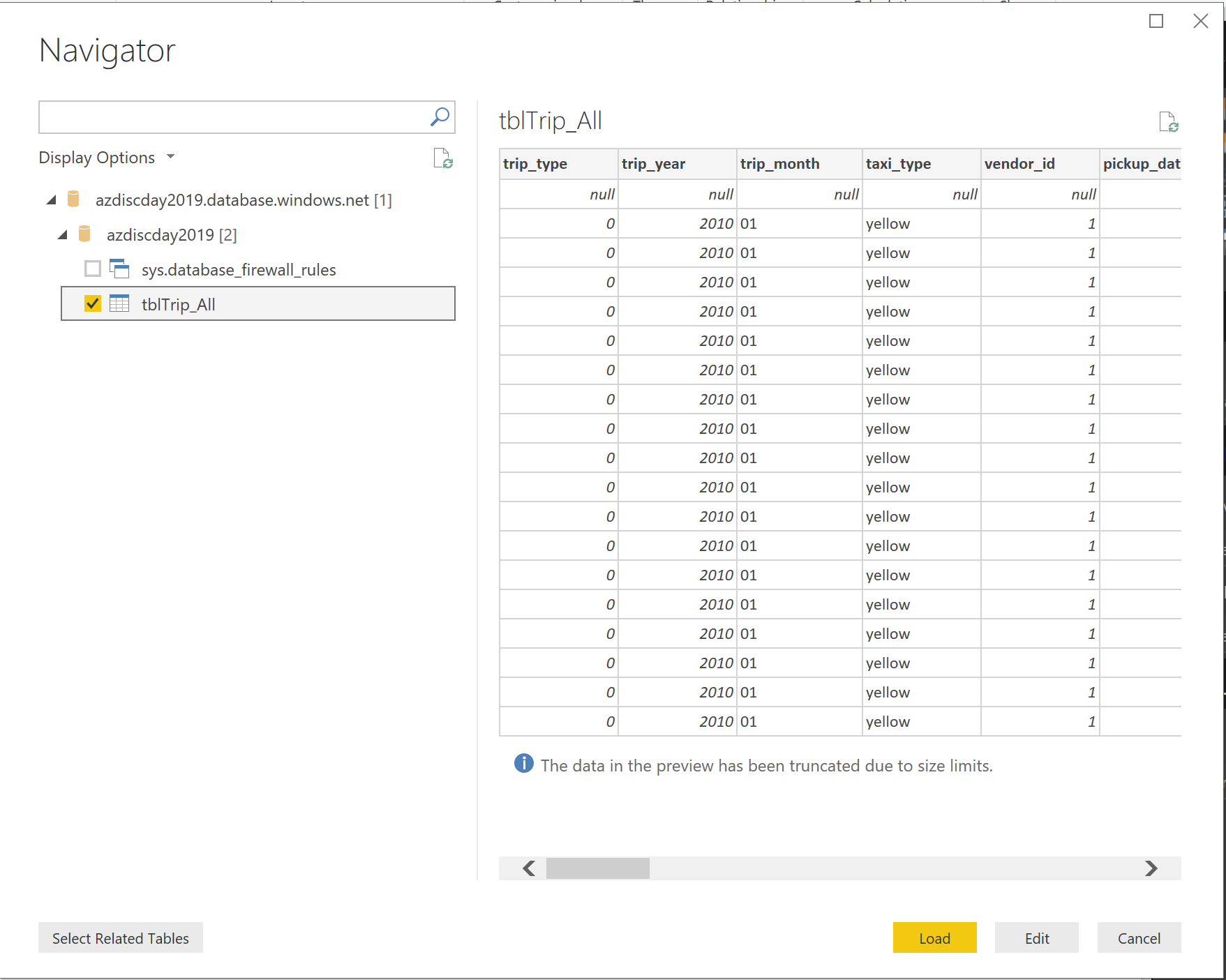
Enter Server name and select direct query and click ok

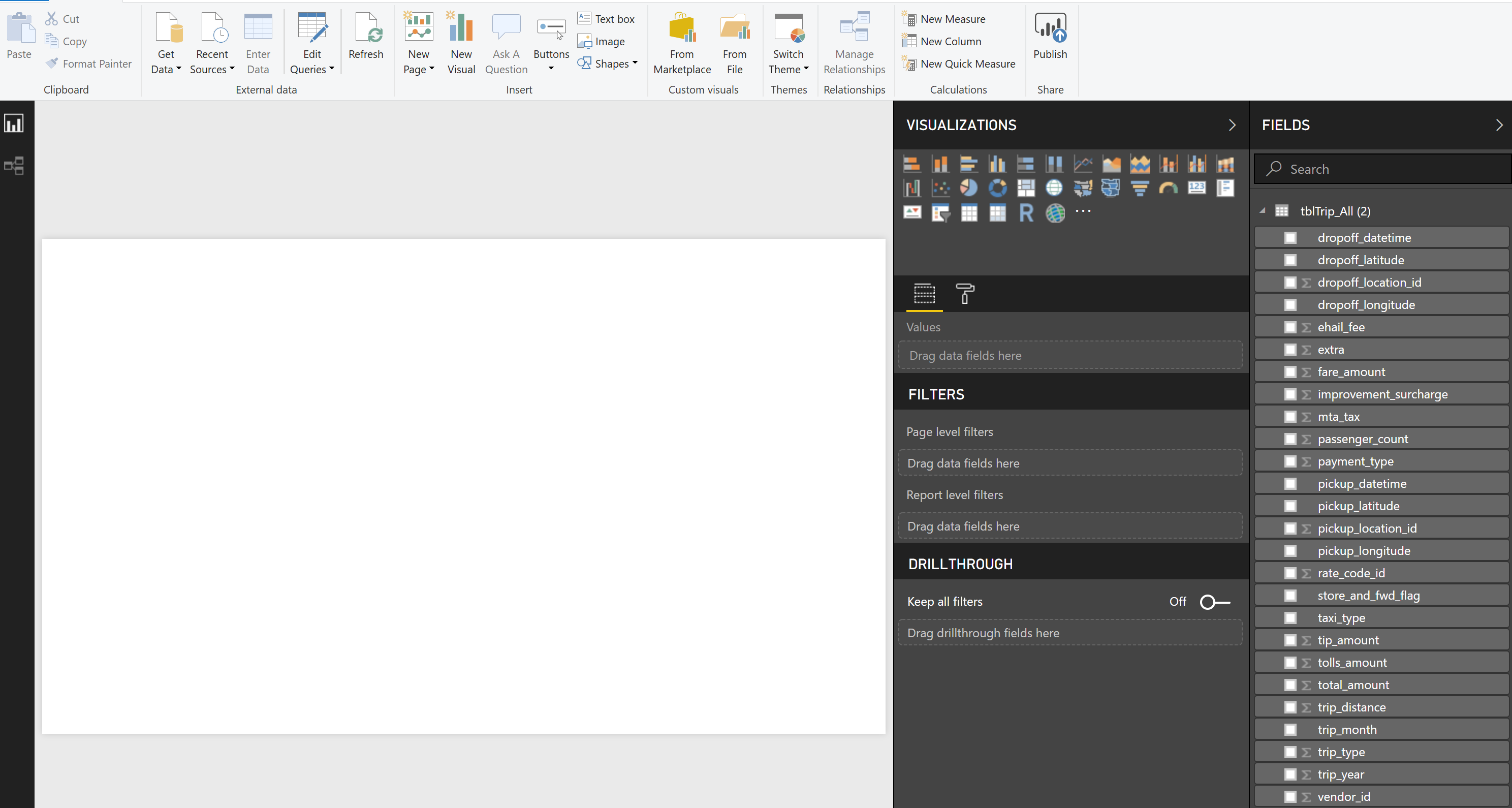


Select Database as the option and enter credentials



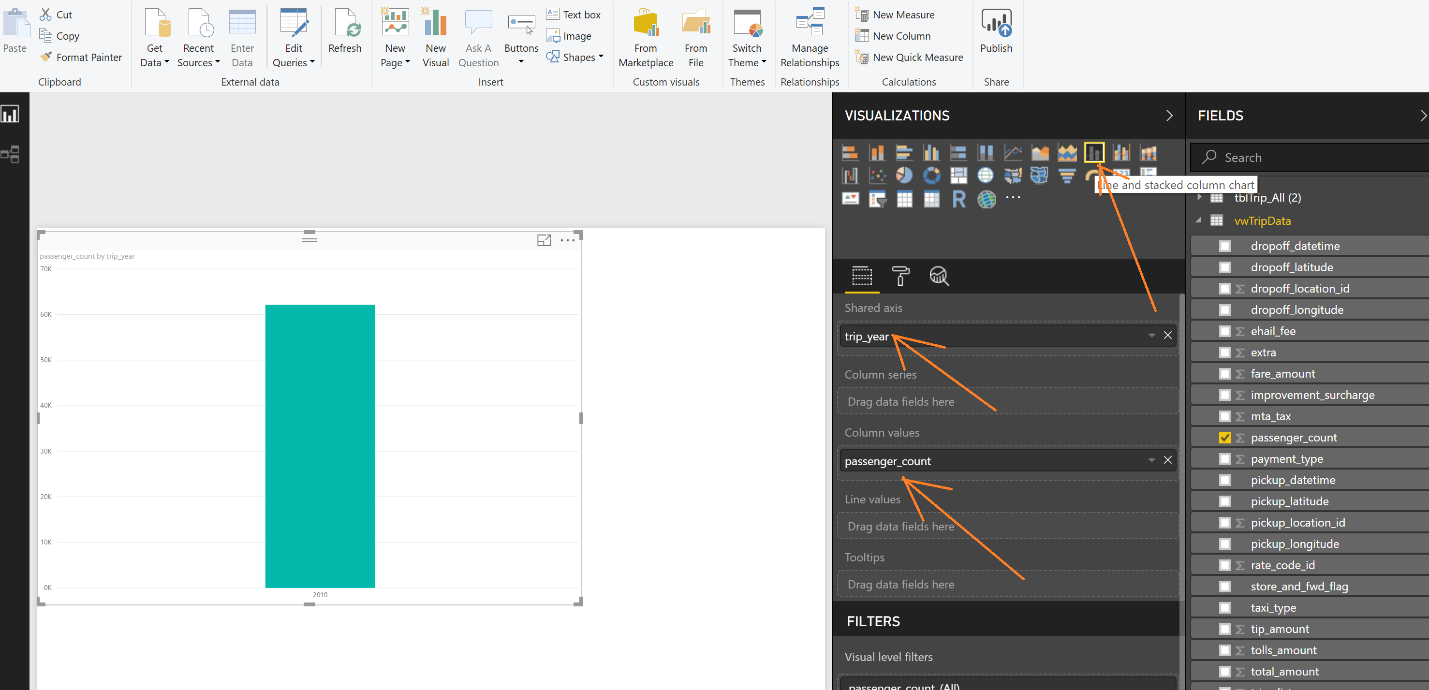
Select table and click load



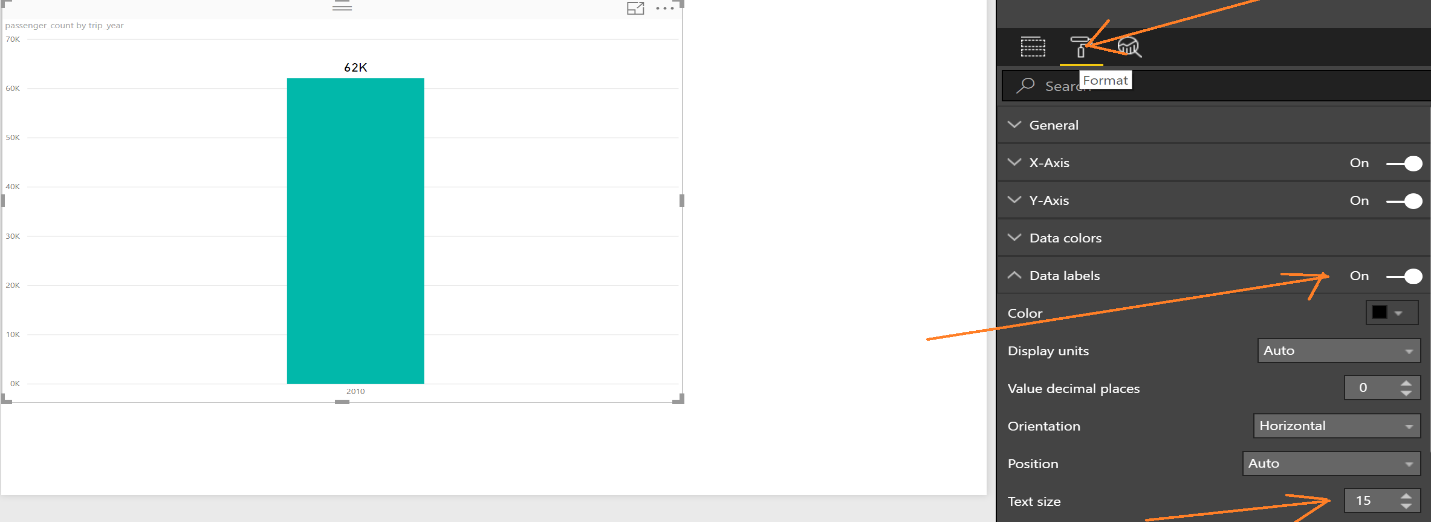


Create a Power BI Chart 🡪 Total passengers by year.

Drag year into Axis and passengers count in values.



Please format labels using below options –



Rename chart title –

