



LIS Nepal Pvt. Ltd.
Setup Guide for Workshop
Prepared for:
LOCUS Workshop Participants
January 2023

Contents

| | |
|---|-----------|
| Document Revision History | ii |
| 1 Setting up the Database (Snowflake) | 1 |
| 1.1 Creating a Snowflake account..... | 1 |
| 1.2 Creating database definitions | 3 |
| 2 Setting up the environment for Python..... | 8 |
| 3 Setting up Microsoft Power BI | 11 |
| 3.1 Installing Microsoft Power BI..... | 11 |
| 3.2 Connecting Power BI to Data Source | 12 |

Document Revision History

| Version | Date | Modified By | Comments |
|---------|------------|---------------|------------------------------------|
| v0.1 | 05/09/2022 | Workshop Team | Installation Guide |
| v0.2 | 05/11/2022 | Workshop Team | Data Source connection in Power BI |
| v0.3 | 01/03/2023 | Workshop Team | Formatting and restructuring |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

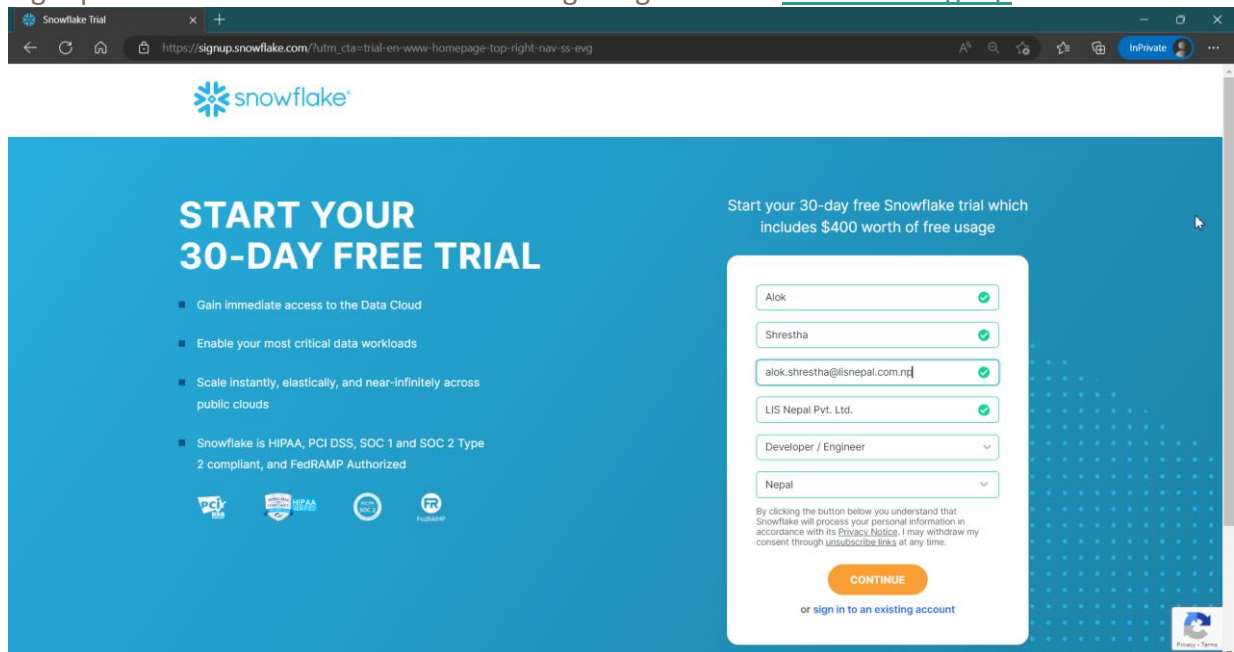
1 Setting up the Database (Snowflake)

For this workshop, we will use Snowflake, which is a cloud-based data warehousing company.

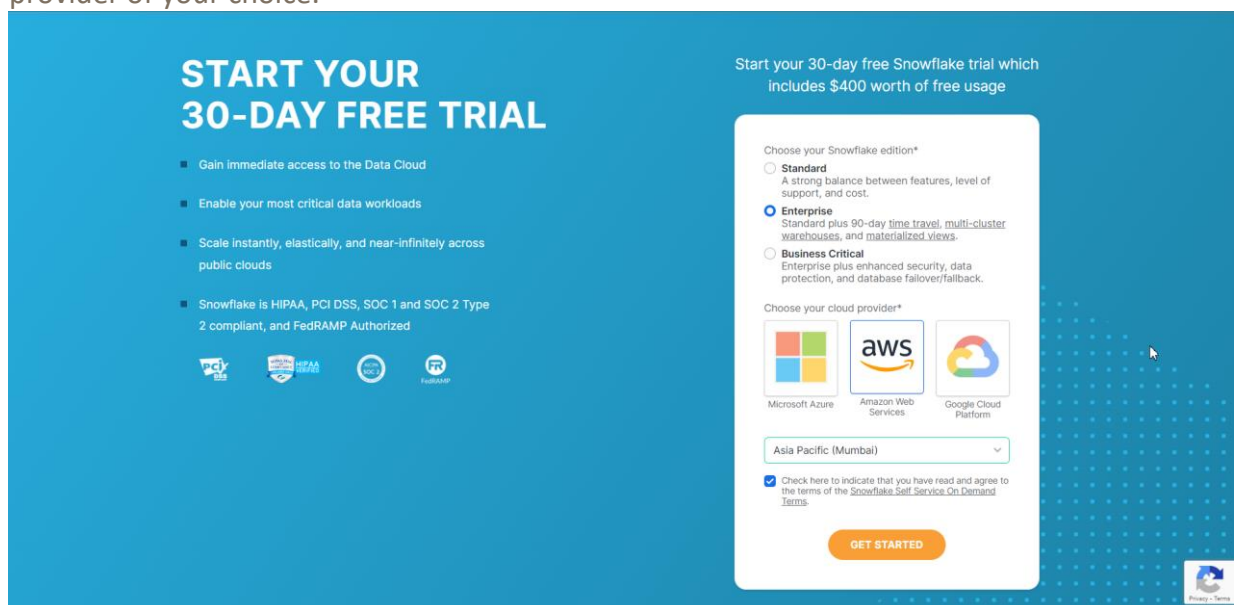
1.1 Creating a Snowflake account

Please follow the following steps to set up the Snowflake account.

1. Signup for the Snowflake trial account using the given link - [Snowflake Signup](https://signup.snowflake.com).



2. You can begin your free Snowflake trial by choosing any Snowflake edition and cloud provider of your choice.



3. Click get started to receive the email containing the activation link for the Snowflake account and the link to access the cloud database as shown below.

Start your 30-day free Snowflake trial which includes \$400 worth of free usage

YOU'RE NOW SIGNED UP! ✓

An email to activate your account has been sent to alok.shrestha@lisnepal.com.np (it may take a few minutes to arrive).


Thank you for your response!

➤ **GETTING STARTED VIDEO**
A quick tour of Snowflake's Platform.
[Watch Video](#)

➤ **VIRTUAL HANDS-ON LAB**
Instructor-led, online lab guiding you through key product features.
[Reserve Your Seat](#)

➤ **FULL SNOWFLAKE DOCUMENTATION**
A comprehensive document covering all aspects of Snowflake's Platform.
[Go To Documentation](#)

Activate Your Snowflake Account!

 snowflake

Hi Alok,

Congratulations on taking the first step to become a data-driven organization by signing up for Snowflake. Click the button below to activate your account.

[CLICK TO ACTIVATE](#)


Please note, your activation link is temporary and will expire in 72 hours. Once you activate your account, you can access it at <https://owddanm-dz37297.snowflakecomputing.com/console/login>.

Be sure to bookmark your login link to easily access your account going forward. If you experience any problems logging into your account or you forgot your username or password, please contact [Snowflake Support](#).

Best regards,

The Snowflake team

4. To activate the account, provide a username and password for the Snowflake account.



Welcome to Snowflake!

Alok Shrestha, please choose a username and password to get started

Username

Username can contain only letters and numbers.

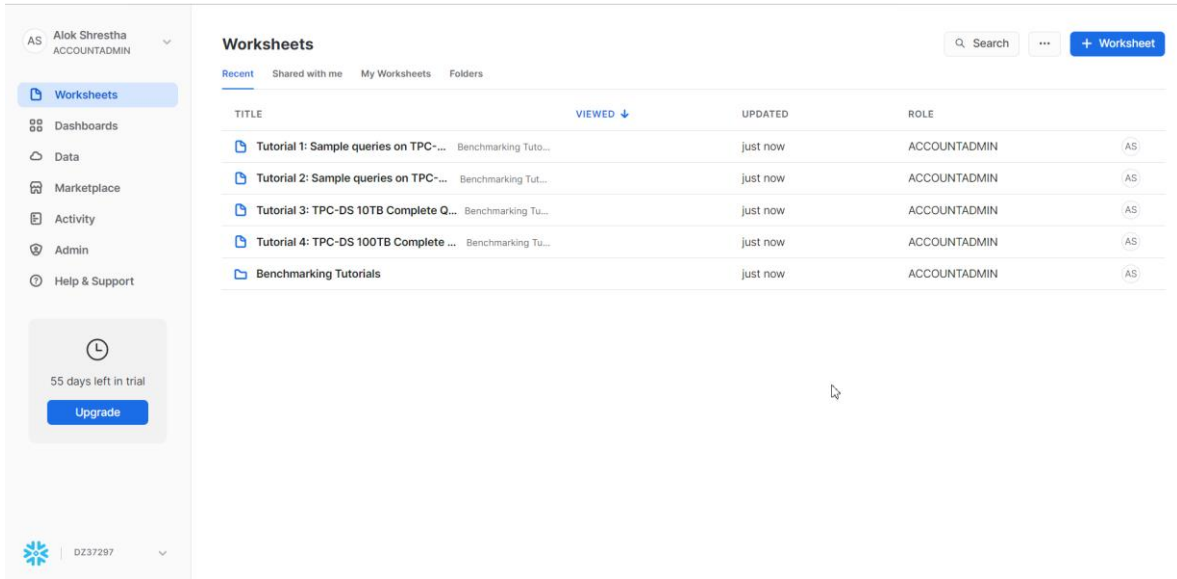
Password

Your password must be 8 - 256 characters and contain at least 1 number(s), 0 special character(s), 1 uppercase and 1 lowercase letter(s).

Confirm password

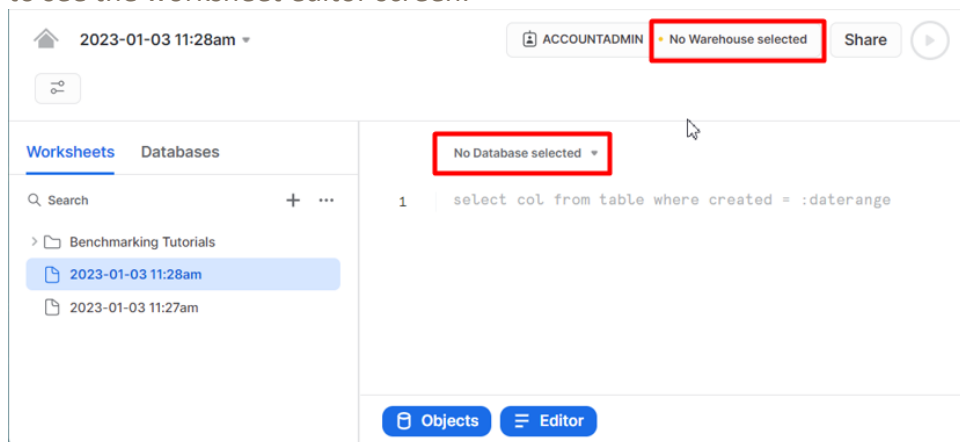
[Get started](#)

- Once all these steps are completed, your Snowflake account is ready and you are all set to create your own database. Use the link from your email and use your credentials to open the Snowflake console.

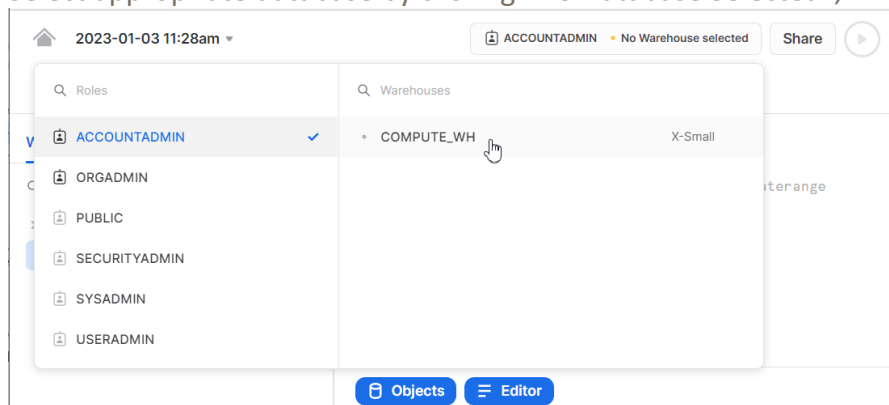


1.2 Creating database definitions

- You can run your queries using Worksheet. A worksheet is basically a workspace where you can create and submit SQL queries. Click the “+ Worksheet” button on the top right to see the worksheet editor screen.

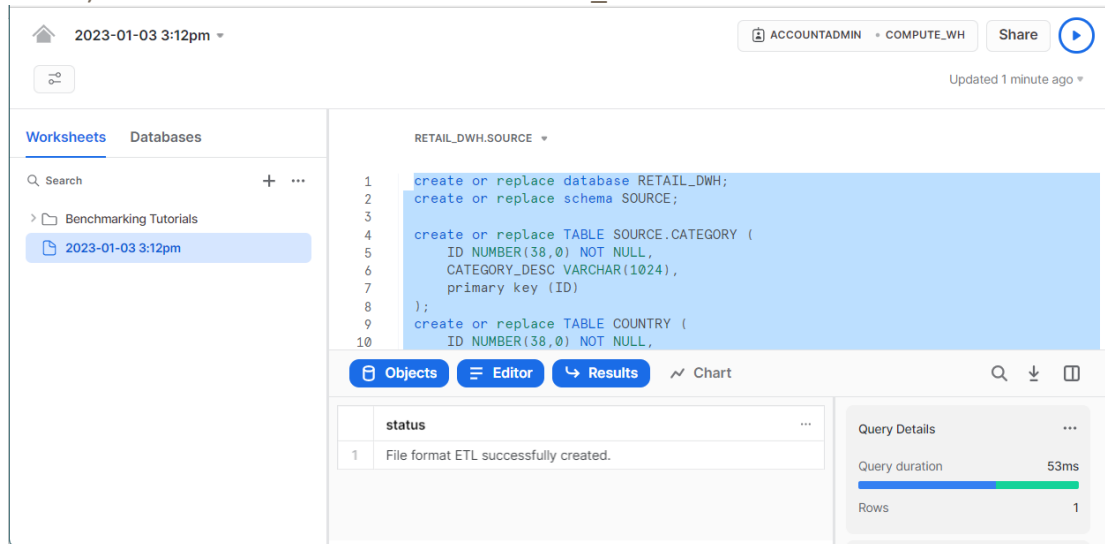


- If the screen shows “No Warehouse selected”, select default COMPUTE_WH warehouse. Select appropriate database by clicking “No Database Selected”, when needed.

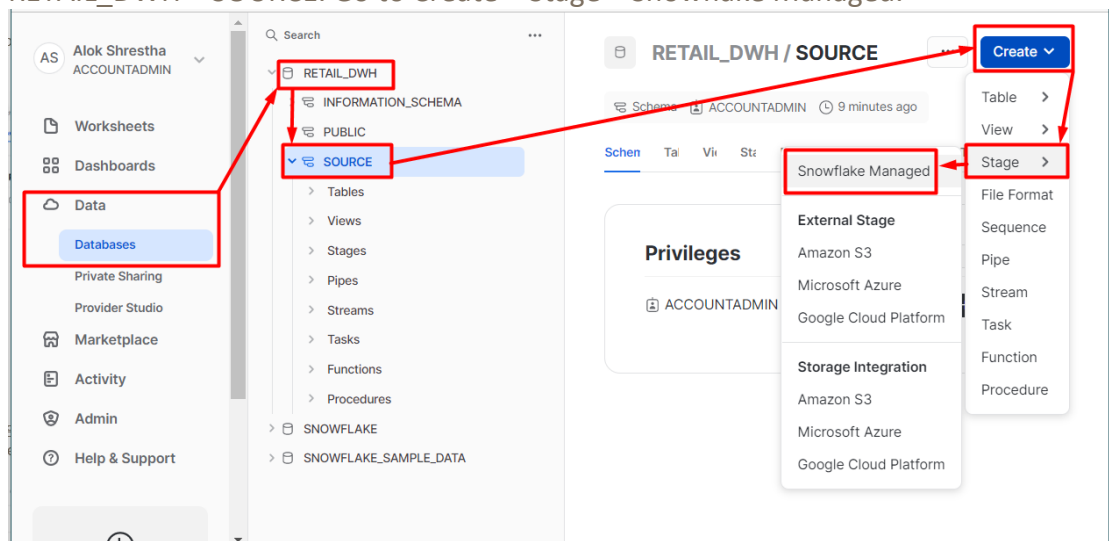


NOTES: For all the following steps please make sure to

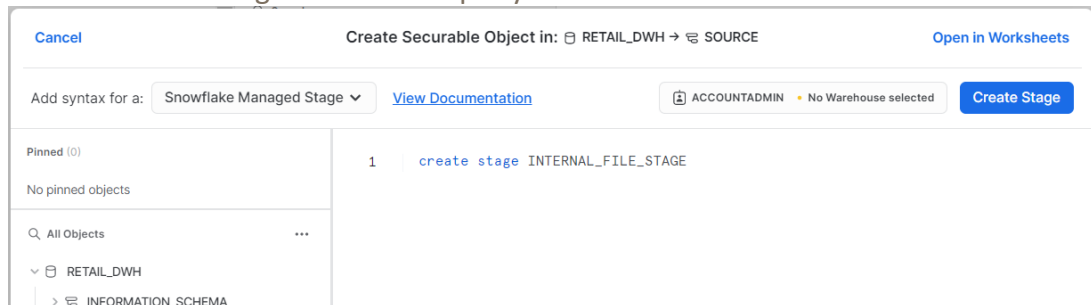
- a. Copy the queries in the provided file/s to a new worksheet (Files are provided separately)
 - b. Select the appropriate Warehouse and database from above the worksheet editor
 - c. Select the entire queries in the editor before executing them (Click Play button on the top right to execute)
3. Create database: Using the SQL DDL (Data Definition Language), we will create our own database and tables. Execute the queries in the file (sql\tables source.sql, file inside sql folder) to create the database named RETAIL_DWH.



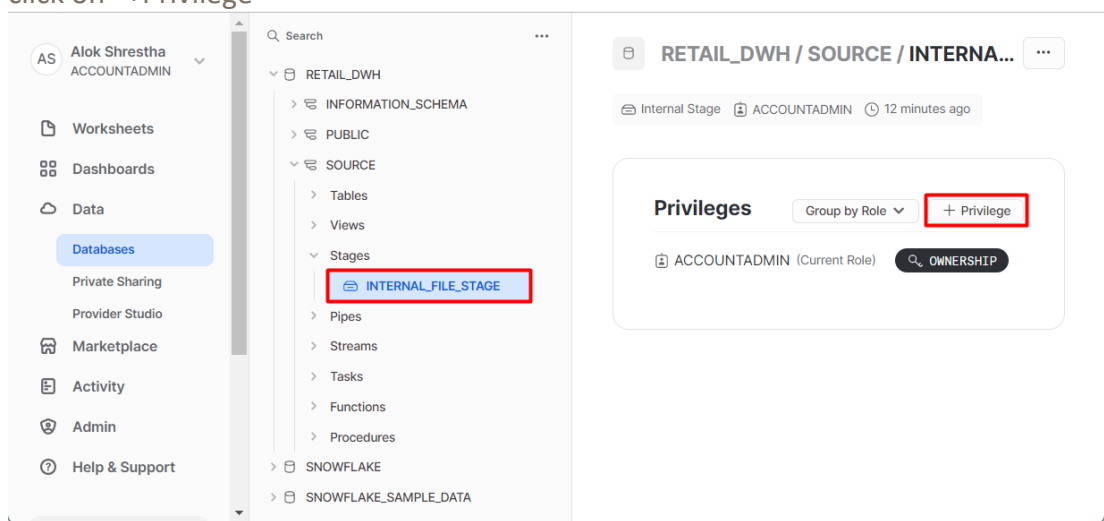
4. Insert Data to Database: Now let's insert the data into the database we just created by executing the queries in the file (sql\insert_source.sql). Select the database (RETAIL_DWH.SOURCE) before executing the query. You can verify the insertion of data by navigating to the Databases section in the leftmost panel of the Snowflake console.
6. Creating File Load Staging: We will create the stage location, to use for loading data from files into Snowflake tables and unloading data from tables into files. To do this, go to the Snowflake UI and follow the steps as below:
- a. Go to Data > Databases from the leftmost panel on Home Screen. Click on the RETAIL_DWH > SOURCE. Go to Create > Stage > Snowflake Managed.



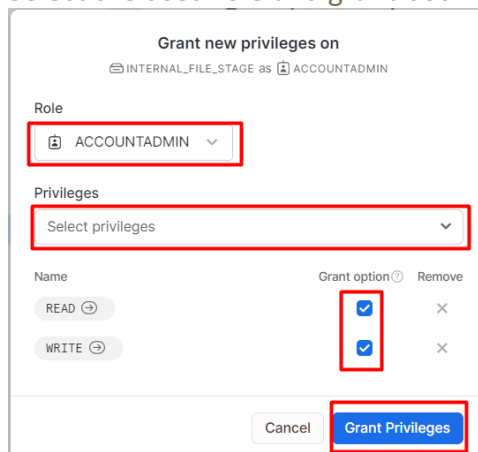
- b. This will open a new worksheet with a create stage command in it. Rename the name of the stage to “INTERNAL_FILE_STAGE” so that the query is: `create stage INTERNAL_FILE_STAGE` Make sure the name is same as mentioned above, since this is used in config. Execute this query.



- c. Next, we need to grant privileges of read and write to the current role. To do this, click on “+Privilege”



- d. Select the used role and grant both read and write privileges to the role.



7. Creating Schema: Now we will create the schema definition for our workshop. We will create 3 schemas (stage, temp, and target).

- a. First, we will create the stage schema. For that, we will execute the queries in the file (ddl\stage.ddl). Make sure to select the appropriate database before execution.

The screenshot shows the Oracle SQL Developer interface. The top bar indicates the user is ACCOUNTADMIN and the database is COMPUTE_WH. The left sidebar shows the 'Worksheets' tab with a list of worksheets, including '2023-01-03 3:44pm'. The main editor shows a SQL query in the RETAIL_DWH.STAGE schema. The query is:

```
1 create or replace schema STAGE;
2
3 create or replace TABLE STG_CATEGORY (
4   ID NUMBER(38,0) NOT NULL,
5   CATEGORY_DESC VARCHAR(1024),
6   primary key (ID)
7 );
8 create or replace TABLE STG_SUBCATEGORY (
```

The status bar shows 'Table STG_SALES successfully created.' and the query details show a duration of 249ms and 1 row.

- b. Now we will create the temp schema. For that, we will execute the queries in the file (ddl\temp.ddl). Make sure to select the appropriate database before execution.

The screenshot shows the Oracle SQL Developer interface. The top bar indicates the user is ACCOUNTADMIN and the database is COMPUTE_WH. The left sidebar shows the 'Worksheets' tab with a list of worksheets, including '2023-01-03 3:58pm'. The main editor shows a SQL query in the RETAIL_DWH.TEMP schema. The query is:

```
1 create or replace schema TEMP;
2
3 create or replace TABLE TMP_AGG_SLS_PLG_MONTH (
4   PDT_KY NUMBER(38,0),
5   STORE_KY NUMBER(38,0),
6   CTORY_KY NUMBER(38,0)
```

The status bar shows 'Table TMP_YEAR successfully created.' and the query details show a duration of 91ms and 1 row.

- c. Finally, we will create the target schema by executing the queries in the file (ddl\target.ddl). Make sure to select the appropriate database before execution.

The screenshot shows the Oracle SQL Developer interface. The top bar indicates the user is ACCOUNTADMIN and the database is COMPUTE_WH. The left sidebar shows the 'Worksheets' tab with a list of worksheets, including '2023-01-03 3:51pm'. The main editor shows a SQL query in the RETAIL_DWH.TARGET schema. The query is:

```
1 create or replace schema TARGET;
2
3 create or replace TABLE D_RETAIL_CNTRY_T (
4   CNTRY_ID NUMBER(38,0),
5   CNTRY_KY NUMBER(38,0) NOT NULL autoincrement,
6   CNTRY_DESC VARCHAR(50),
7   OPEN_CLOSE_CD VARCHAR(1),
8   ROW_INSRT_TMS TIMESTAMP_NTZ(9),
```

The status bar shows 'Table F_RETAIL_SLS_T successfully created.' and the query details show a duration of 169ms and 1 row.

8. Insert Seed Data to Calendar: Next, we perform calendar seeding in the Target schema. For that, execute the query in the file (sql\Calendar Create.sql).

The screenshot displays a database management tool interface. At the top, the status bar shows the date and time '2023-01-03 4:01pm', the user 'ACCOUNTADMIN', the database 'COMPUTE_WH', and a 'Share' button. Below the status bar, the left sidebar contains a 'Worksheets' tab and a 'Databases' section. Under 'Databases', there is a search bar and a list of files, with '2023-01-03 4:01pm' selected. The main area shows the SQL query in the 'RETAIL_DWH.TARGET' schema. The query is as follows:

```
1 TRUNCATE TABLE D_RETAIL_TIME_YEAR_T;  
2  
3 INSERT INTO RETAIL_DWH.TARGET.D_RETAIL_TIME_YEAR_T (  
4     ID  
5     , YEAR_KY  
6     , YEAR_START_DATE  
7     , YEAR_END_DATE  
8     , OPEN_CLOSE_CD
```

Below the query editor, there are tabs for 'Objects', 'Editor', 'Results', and 'Chart'. The 'Results' tab is active, showing a table with one row and one column, 'number of rows inserted', with the value '3,650'. To the right of the results table, there is a 'Query Details' panel showing 'Query duration' as '1.4s' and 'Rows' as '1'.

2 Setting up the environment for Python

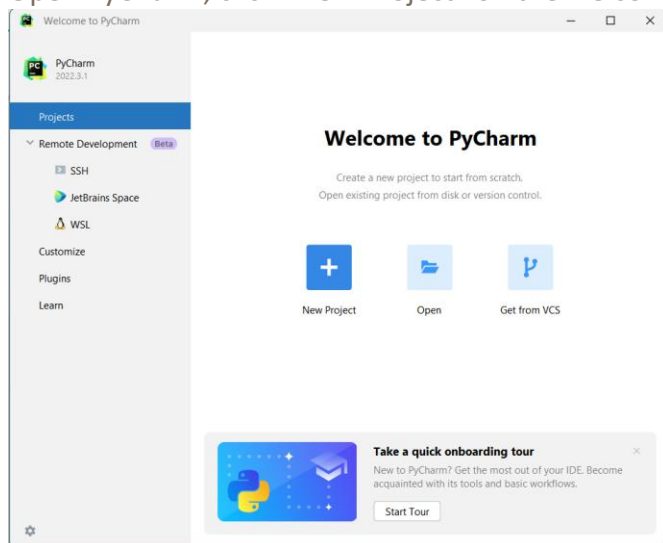
We will use the latest version of Python and PyCharm (Community Edition) as IDE/code editor.

1. Installing Python and PyCharm

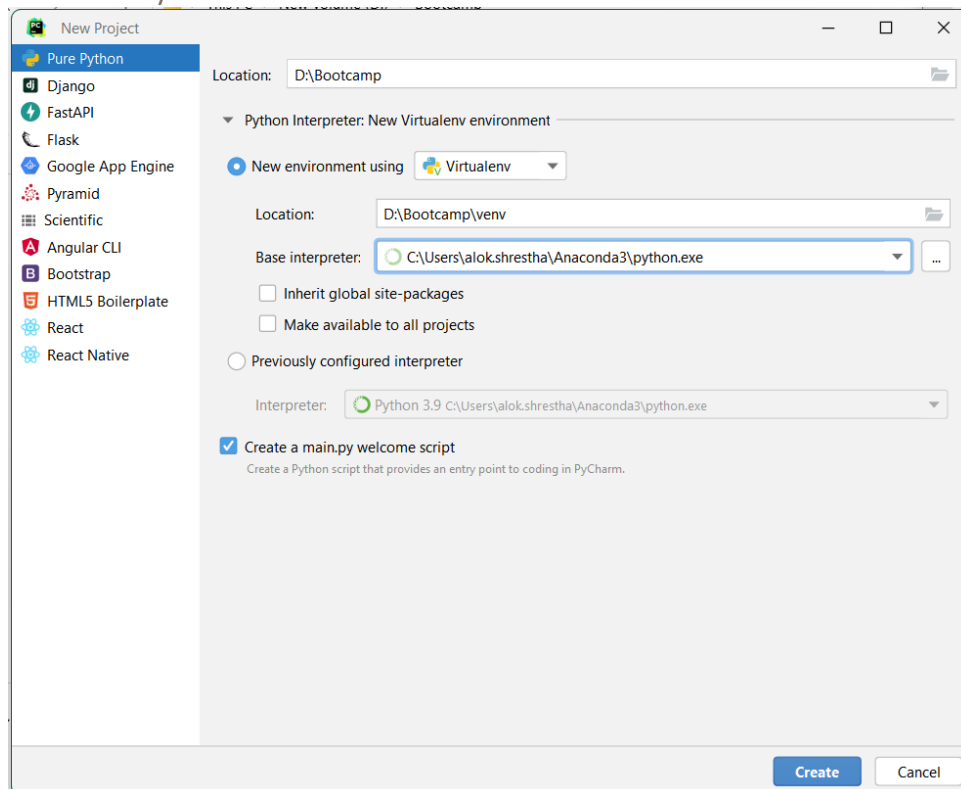
- Please download and install the latest version of the Python using the following link - [Python Download](#). Make sure to add the Python to PATH variables in Windows.
- Download and install PyCharm, using the following link - [PyCharm Download](#). While installing, make sure to set the checkbox option “Add bin folder to PATH”

2. Setting up PyCharm: Once the above are installed, we will setup project for our workshop.

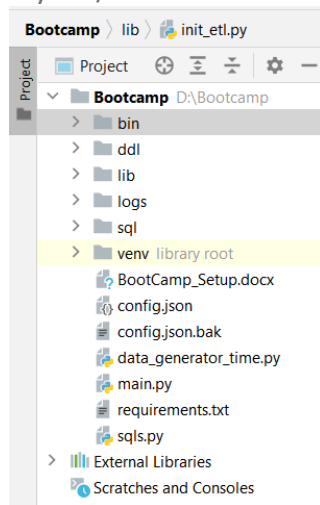
- Open PyCharm, click “New Project” on the welcome screen to create a project.



- PyCharm allows us to create a new environment in multiple ways. We will create a one for our project using Virtualenv. Make sure to update Location field to the location of your choice. Create the folder beforehand if needed.

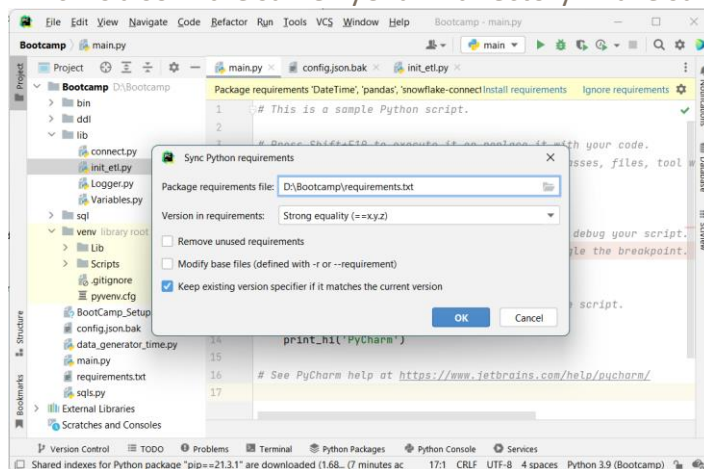


- c. Once virtualenv is created, please extract the zip file you received into the folder you entered in Location field above. Once extracted, you will see the folder structure as shown below in your PyCharm IDE. (If you do not see it, Right Click any file/folder > Reload from disk to refresh the folder structure.)



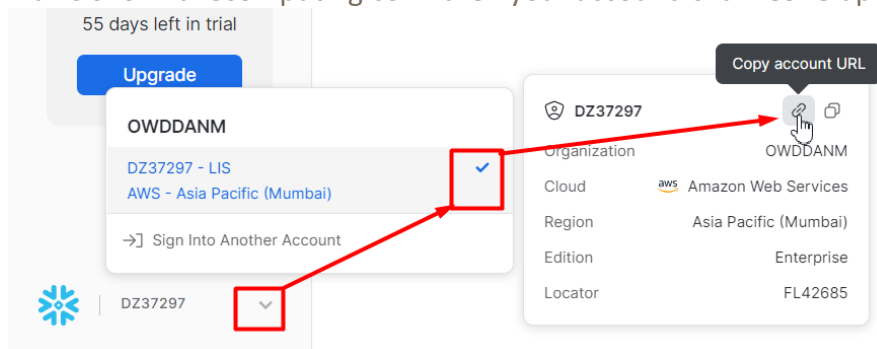
3. Installing Requirements: Now we will install the required packages for our project.

- We will import the packages using the Tool > Sync Python Requirements to sync the requirements for our project.
- Choose the requirements.txt file provided to sync the requirements for our project, which is also in the same PyCharm directory. Make sure to provide the full location.

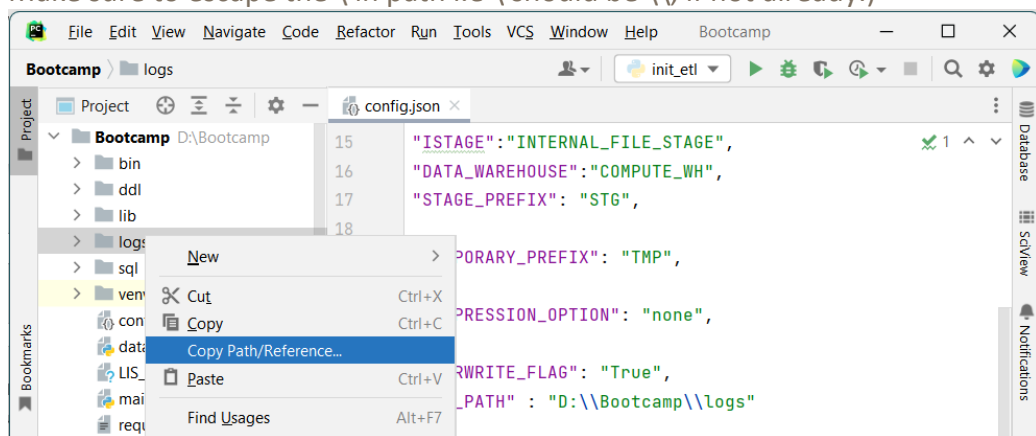


- You will see the prompt asking for the installation of the requirements. Accept the prompt to install of the packages required for our Workshop.
4. Updating Configuration: Now we will update the configuration file for this Workshop. Rename the config.json.bak file in your PyCharm directory to config.json. This file contains the configuration for the Workshop. Please open the config.json file and update the following with appropriate values.
- USER = "Your Snowflake Username",
 - PASSWORD = "Your Snowflake Password",

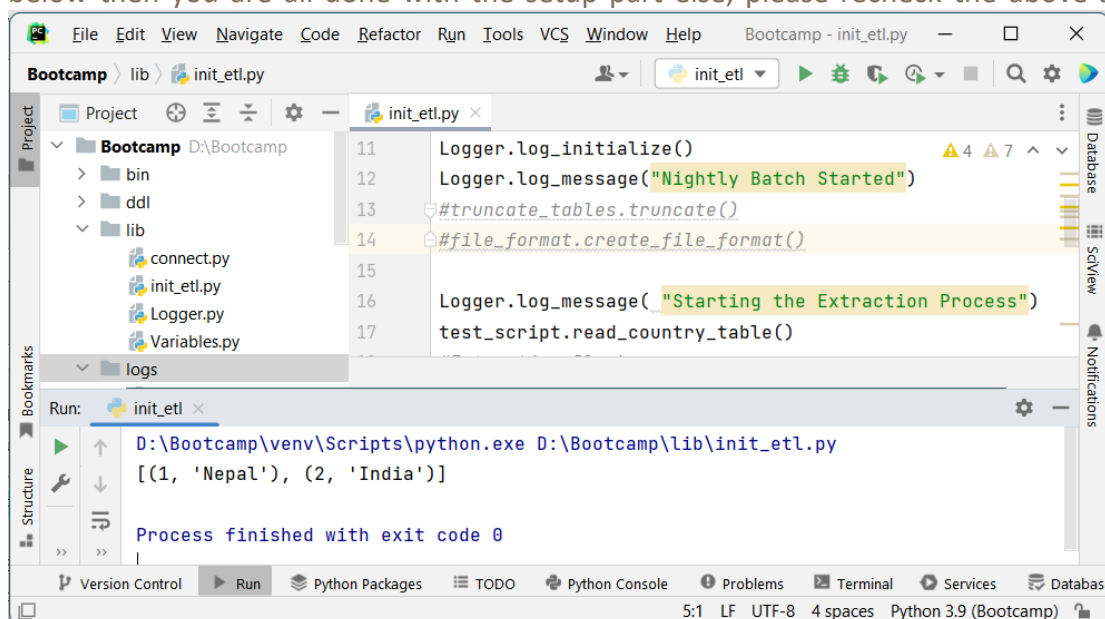
- c. ACCOUNT = (If your Snowflake account URL is 'https://uk28978.ap-south-1.aws.snowflakecomputing.com' then your account is 'uk28978.ap-south-1.aws')



- d. LOG_PATH = "Path to the log folder" (For log path, please create a folder in the project folder then we can navigate to the log folder and copy its absolute path. Make sure to escape the \ in path i.e \ should be \\, if not already.)



5. Confirming Setup so far: To make sure that all installations are correct, we have provided you with a small code snippet that will fetch the data from the database. Please execute the lib\\init_etl.py file (Right Click > Run 'init_etl') from your PyCharm directory. This should result displaying the name of the country as shown below. If you get the output as shown below then you are all done with the setup part else, please recheck the above steps.

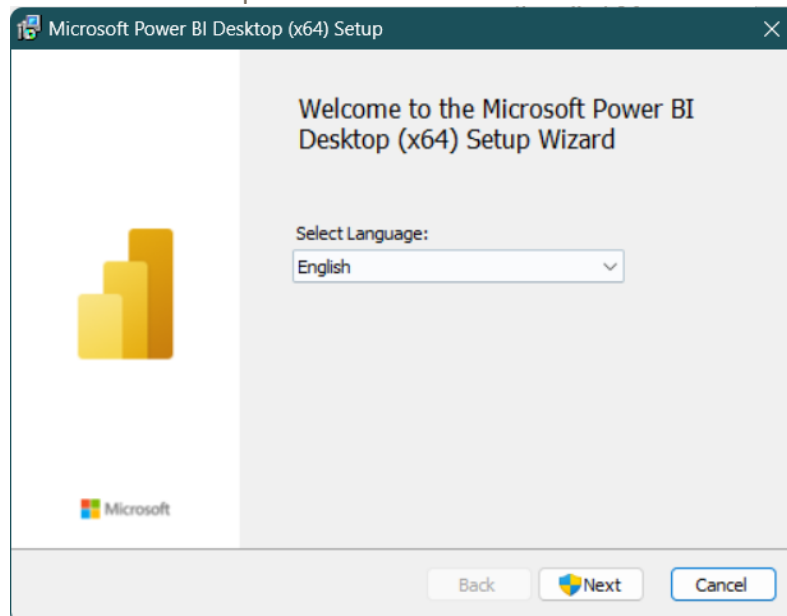


3 Setting up Microsoft Power BI

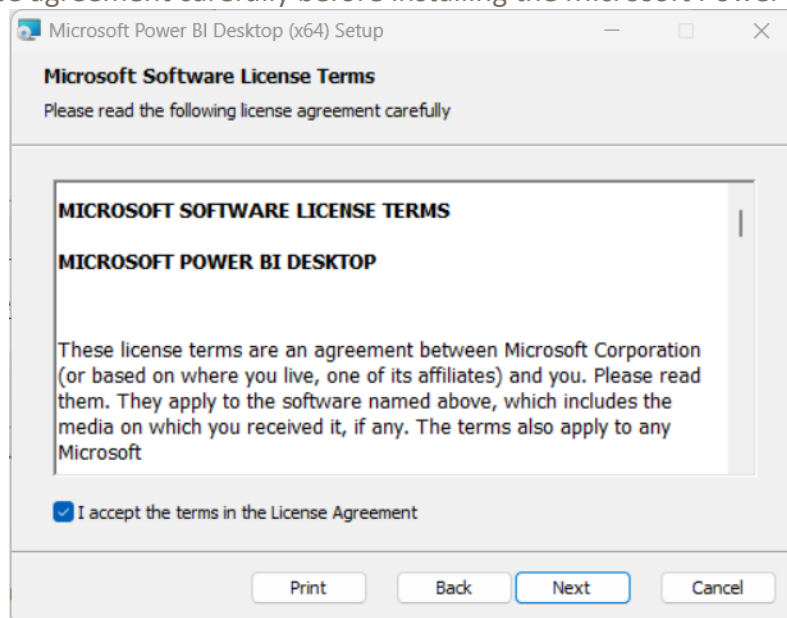
Now we will install Microsoft Power BI, which is an interactive visualization software.

3.1 Installing Microsoft Power BI

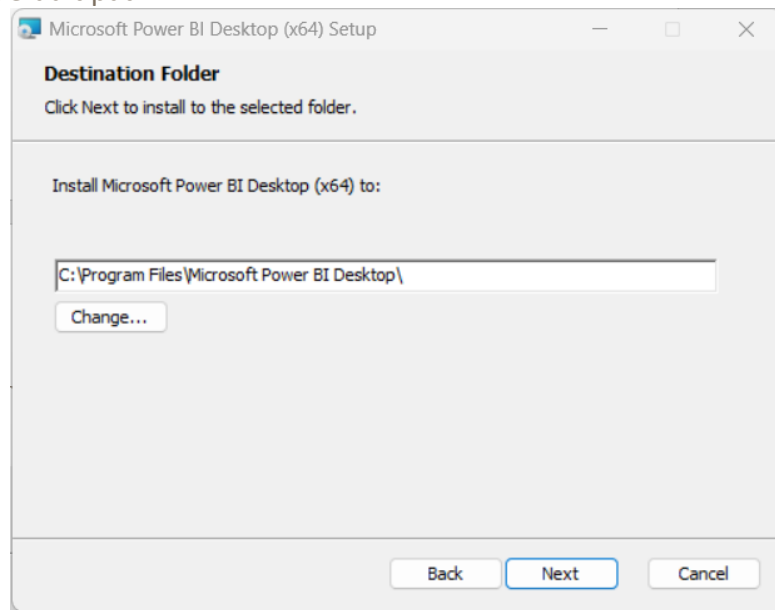
1. Using the following link - [Power BI Downloads](#), please download and install Power BI Desktop version using the executable setup.



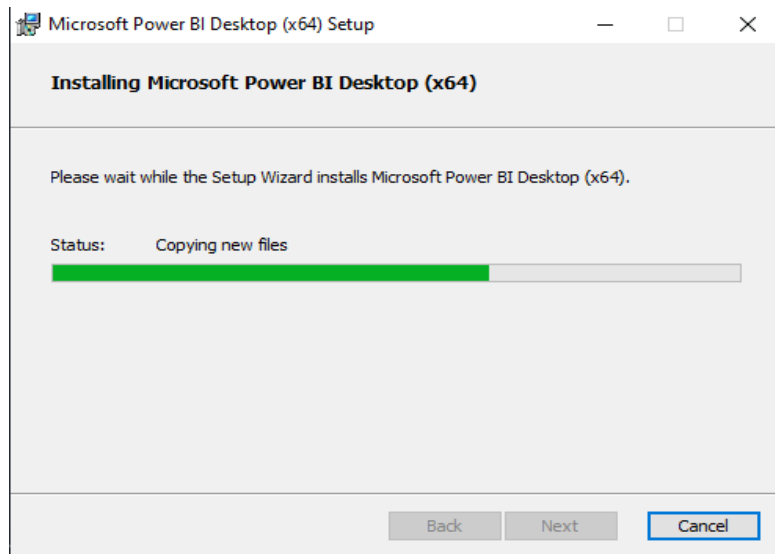
2. Read the license agreement carefully before installing the Microsoft Power BI.



3. Navigate the folder path where you want to install Microsoft Power BI, alternatively you can leave it to default path.



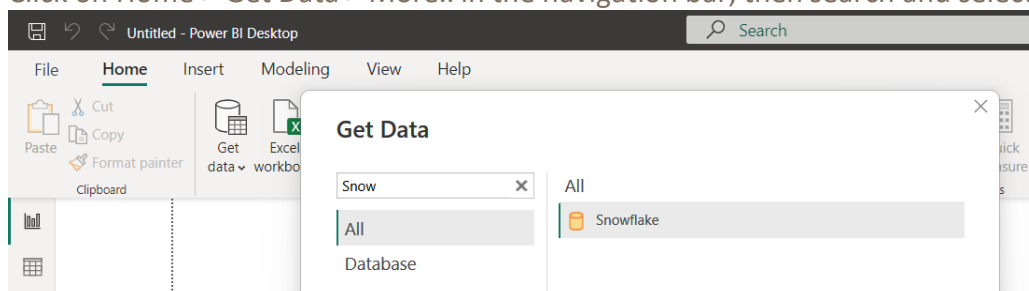
4. Clicking Next will result in installation of Microsoft Power BI which usually takes couple of minutes to complete the installation.



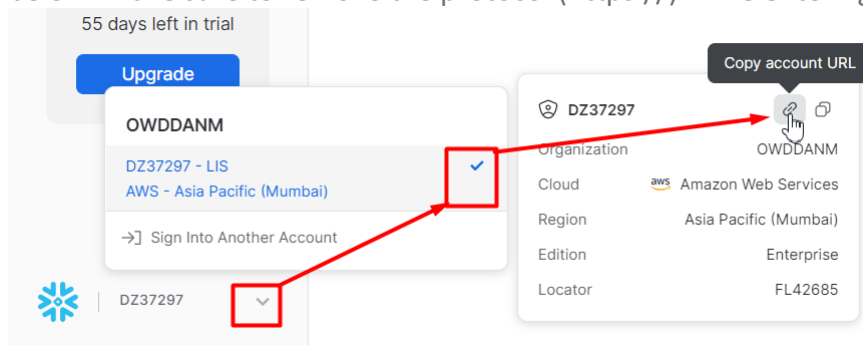
3.2 Connecting Power BI to Data Source

Once you have successfully installed the Power BI, we need to connect and import data from the source. These data will be used to build reports in Power BI.

1. Click on Home > Get Data > More.. in the navigation bar, then search and select Snowflake.



2. Enter your Snowflake server details and warehouse name in the prompt.
 - a. Server: This is URL of the Snowflake trial account you created above and sent to your email. OR you can retrieve this from the console by following the screenshot below. Make sure to remove the protocol (https://) while entering in this field.



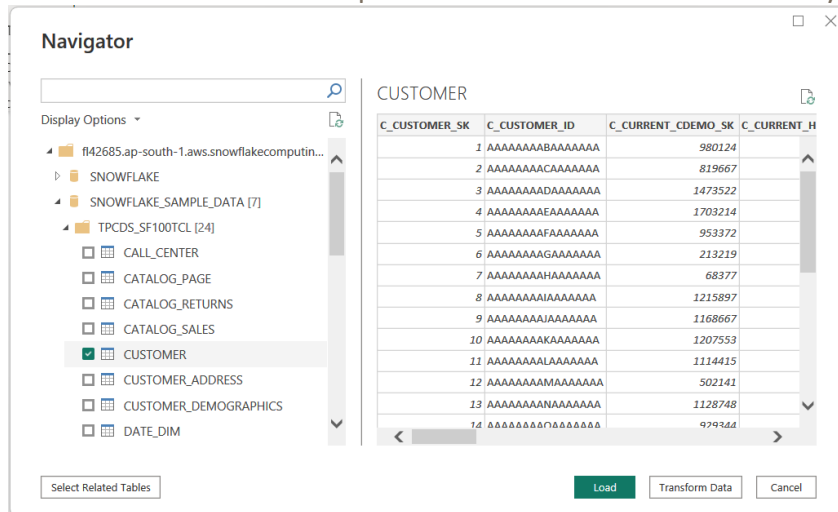
- b. Warehouse: This is the warehouse you used. Default is COMPUTE_WH

The screenshot shows a 'Snowflake' configuration dialog box. It has two input fields: 'Server' with the value 'fl42685.ap-south-1.aws.snowflakecomputing.com' and 'Warehouse' with the value 'COMPUTE_WH'. Below these fields is a link for 'Advanced options'. At the bottom right, there are 'OK' and 'Cancel' buttons.

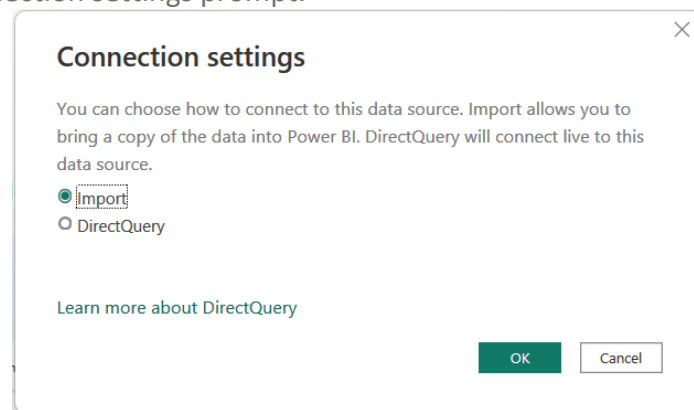
3. Click on 'OK' button of the prompt above and you'll be asked for the Snowflake credentials. Please enter your Snowflake username and password you created during setup. Click on 'Connect' button.

The screenshot shows a 'Snowflake' login dialog box. It has a sidebar on the left with 'Snowflake' and 'Microsoft Account' options. The main area shows the 'Snowflake' option selected, with a text field containing the server URL 'fl42685.ap-south-1.aws.snowflakecomputing.com;C...'. Below this are input fields for 'User name' (containing 'LIS') and 'Password'. At the bottom, there are 'Back', 'Connect', and 'Cancel' buttons.

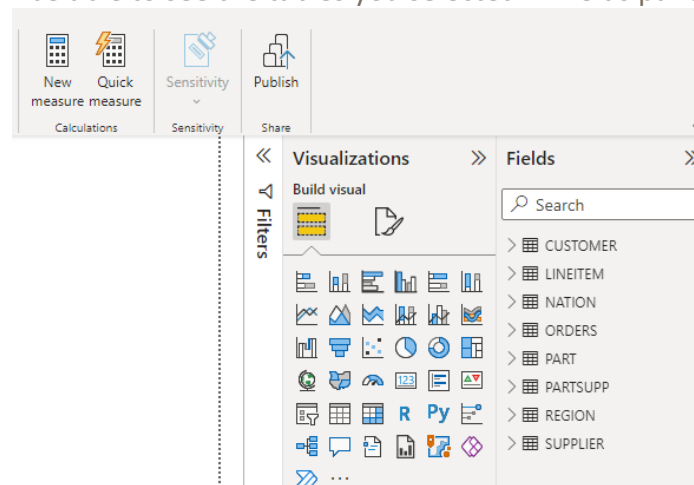
4. You will see the databases present in the warehouse. Select any dataset present.



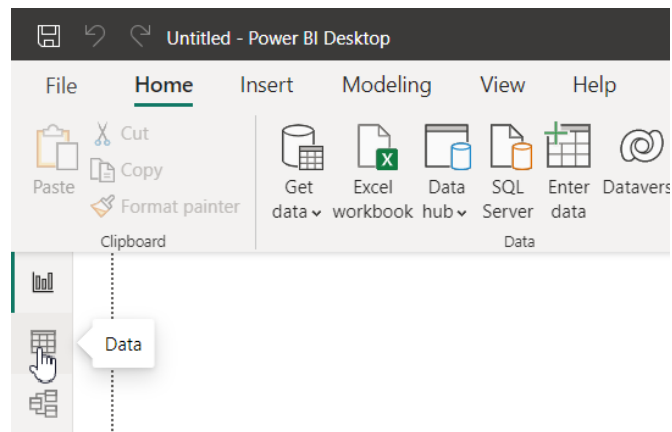
5. After selecting the tables, click Load button to load these data in Power BI. Select 'Import' option in the connection settings prompt.



6. Data import might take few minutes to complete depending upon the data size and after completion you will be able to see the tables you selected in Fields pane of Power BI.



7. Go to the Data tab by selecting the second option as shown in the screenshot below.



8. With the correct execution of above steps, you should be able to see the data in Data tab of Power BI as shown below.

The screenshot shows the Power BI Data view. A table of customer data is displayed. The table has columns: C_CUSTKEY, C_NAME, C_ADDRESS, C_NATIONKEY, C_PHONE, C_ACCTBAL, C_MKTSEGMENT, and C_COMMENT. The data is sorted by C_CUSTKEY in descending order. The right-hand pane shows the 'Fields' list with a search bar and a list of tables: CUSTOMER, LINEITEM, NATION, ORDERS, PART, PARTSUPP, REGION, and SUPPLIER. The 'CUSTOMER' table is selected.

| C_CUSTKEY | C_NAME | C_ADDRESS | C_NATIONKEY | C_PHONE | C_ACCTBAL | C_MKTSEGMENT | C_COMMENT |
|-----------|--------------------|---|-------------|-----------------|-----------|--------------|--|
| 60056 | Customer#000060056 | 2lvoe8n8b1 PJ03lq9ZKel7rH9Fahsu136 | 5 | 13-221-847-3701 | 8880.28 | BUILDING | fully furiously even deposits. final accounts nag |
| 60236 | Customer#000060236 | 5nq9eNWwE9Fk7Qltps3KCvnicgVLCviiH0k | 5 | 13-588-942-4698 | 2128.37 | BUILDING | nusual warhorses. special requests haggle slyly alongs |
| 60424 | Customer#000060424 | yolnMtgZg5QsjdVcQ | 5 | 13-115-627-1559 | 3134.46 | BUILDING | final packages. carefully regular braids acco |
| 60664 | Customer#000060664 | HTFX9WWayr9TXAFAnVikHo2j68htnMS | 5 | 13-649-190-7053 | 1201.03 | BUILDING | close, ironic instructions hagg |
| 60691 | Customer#000060691 | RFahgH8wcUSc9b2G0etXxiat6h5tTzGC | 5 | 13-198-598-3446 | 6406.36 | BUILDING | riously pending requests. accounts sleep slyly even de |
| 60742 | Customer#000060742 | QZBWkXPuNuZ4T5E65uCG968 | 5 | 13-504-343-6211 | 347.88 | BUILDING | cross the special instructions. final asymptotes wake at |
| 60781 | Customer#000060781 | B0cKZolQdA | 5 | 13-473-821-1698 | 488.81 | BUILDING | kages. quickly unusual packages doz |
| 61009 | Customer#000061009 | 3gm89hjeBkZkYoHN Mbxg7M76KANK1TlJmbi38Y | 5 | 13-831-250-2586 | -210.19 | BUILDING | ges wake ironic requests. slow, express pinto beans wa |
| 61037 | Customer#000061037 | LvFW6TYOkuKFPvQGrTldfMdWb | 5 | 13-628-354-5900 | 8373.54 | BUILDING | onic dependencies. regular accounts detect blithely ev |
| 61068 | Customer#000061068 | s7NBLcCKNoAHjb | 5 | 13-362-456-9531 | 1414.84 | BUILDING | ccording to the even, regular accounts. slyly ironic pint |
| 61099 | Customer#000061099 | FEL2j9RI JnPAWR13sdRXsv | 5 | 13-940-274-4785 | 8717.97 | BUILDING | k deposits. ironic pinto beans at the regular depths wa |
| 61186 | Customer#000061186 | v1pg7bGnxIDVZQ1gH201kAKvymG,1UXGZqQ 3 | 5 | 13-808-801-7782 | 2780.8 | BUILDING | carefully unusual requests haggle carefully at the ironi |
| 61415 | Customer#000061415 | 0Hg1Q,ulikC | 5 | 13-636-137-6607 | 7958.72 | BUILDING | y ironic deposits sleep blithely bold, final packages. sly |

9. With this, we have completed the Power BI setup.