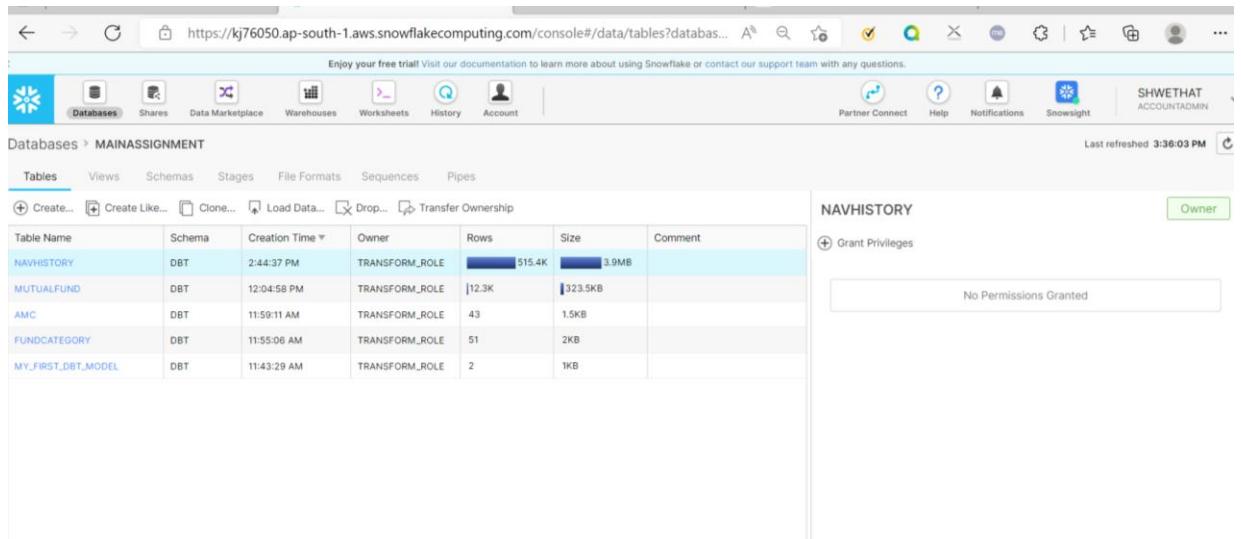


# Shwetha. T – [shwetht@deloitte.com](mailto:shwetht@deloitte.com)

## Main Assignment

1. Create Tables and load csv data to those tables  
All tables created successfully



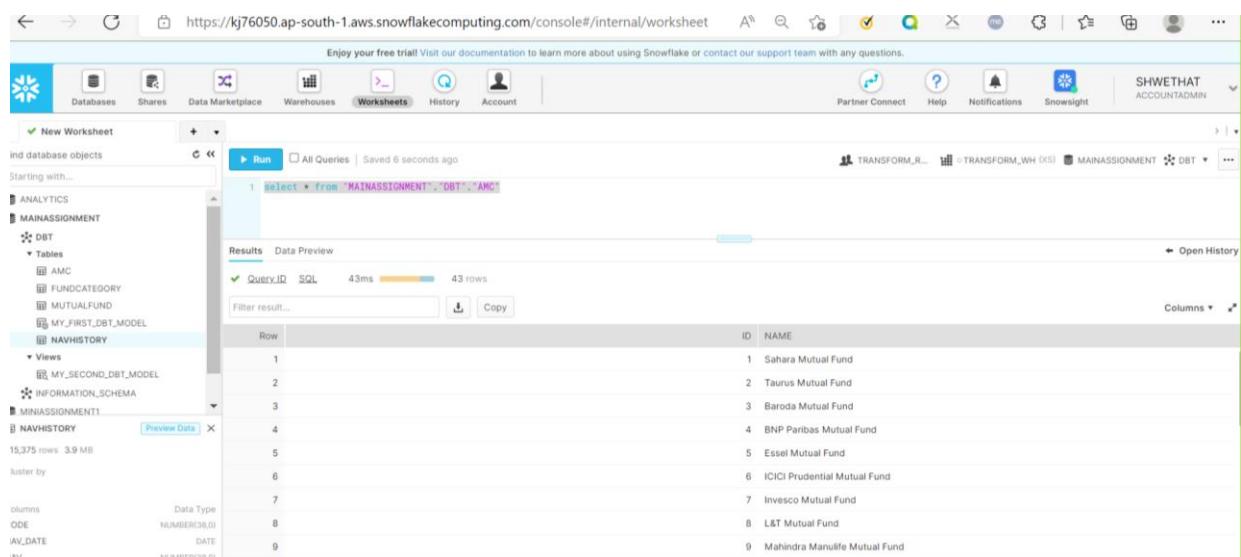
Databases > MAINASSIGNMENT

Table Name	Schema	Creation Time	Owner	Rows	Size	Comment
NAVHISTORY	DBT	2:44:37 PM	TRANSFORM_ROLE	515.4K	3.9MB	
MUTUALFUND	DBT	12:04:58 PM	TRANSFORM_ROLE	12.3K	323.5KB	
AMC	DBT	11:59:11 AM	TRANSFORM_ROLE	43	1.5KB	
FUNDCATEGORY	DBT	11:55:06 AM	TRANSFORM_ROLE	51	2KB	
MY_FIRST_DBT_MODEL	DBT	11:43:29 AM	TRANSFORM_ROLE	2	1KB	

2. Load csv data to the tables:

Here I have loaded AMC.csv , Fundcategory.csv, mutualfund.csv files using snowflake UI directly since size of file is less than 16MB.

### AMC table



New Worksheet

```
select * from 'MAINASSIGNMENT'.'DBT'.AMC
```

Row	ID	NAME
1	1	Sahara Mutual Fund
2	2	Taurus Mutual Fund
3	3	Baroda Mutual Fund
4	4	BNP Paribas Mutual Fund
5	5	Essel Mutual Fund
6	6	ICICI Prudential Mutual Fund
7	7	Invesco Mutual Fund
8	8	L&T Mutual Fund
9	9	Mahindra Manulife Mutual Fund

## FundCategory Table

The screenshot shows the Snowflake Worksheet interface with the following details:

- URL:** https://kj76050.ap-south-1.aws.snowflakecomputing.com/console#/internal/worksheet
- Worksheet Title:** New Worksheet
- Query:** `select * from 'MAINASSIGNMENT'.DBT.'FUNDCATEGORY'`
- Results:**
  - 51 rows
  - 47ms execution time

ID	CATEGORY
1	Open Ended Schemes ( Growth )
2	Open Ended Schemes ( Liquid )
3	Open Ended Schemes ( Money Market )
4	Open Ended Schemes ( ELSS )
5	Open Ended Schemes ( Equity Scheme - Multi Cap Fund )
6	Open Ended Schemes ( Equity Scheme - Large Cap Fund )
7	Open Ended Schemes ( Equity Scheme - Large & Mid Cap Fund )
8	Open Ended Schemes ( Equity Scheme - Mid Cap Fund )
9	Open Ended Schemes ( Equity Scheme - Small Cap Fund )
10	Open Ended Schemes ( Equity Scheme - Dividend Yield Fund )

## MutualFunds Table

The screenshot shows the Snowflake Worksheet interface with the following details:

- URL:** https://kj76050.ap-south-1.aws.snowflakecomputing.com/console#/internal/worksheet
- Worksheet Title:** New Worksheet
- Query:** `select * from 'MAINASSIGNMENT'.DBT.'MUTUALFUND'`
- Results:**
  - 12,261 rows
  - 35ms execution time

Row	CODE	AMC_ID	CATEGORY_ID	NAME	ISIN_GROWTH	ISIN_DIVIDEND_PAYOUT	ISIN_DIVIDEND_REINVESTN
1	120373	1	1	SAHARA BANKING & FIN...	INF515L01AJ6	INF515L01AJ6	NULL
2	109493	1	1	SAHARA BANKING & FIN...	INF515L01494	INF515L01494	NULL
3	120374	1	1	SAHARA BANKING & FIN...	INF515L01AH0	INF515L01AH0	INF515L01AIB
4	109494	1	1	SAHARA BANKING & FIN...	INF515L01478	INF515L01478	INF515L01486
5	101529	1	1	Sahara Growth Fund-Div...	INF515L01155	INF515L01155	INF515L01163
6	120288	1	1	Sahara Growth Fund-Div...	INF515L01759	INF515L01759	INF515L01767
7	101528	1	1	Sahara Growth Fund-Gro...	INF515L01171	INF515L01171	NULL
8	120289	1	1	Sahara Growth Fund-Gro...	INF515L01775	INF515L01775	NULL
9	120352	1	1	Sahara Infrastructure Fun...	INF515L01924	INF515L01924	INF515L01932

In order to load navHistory data I am following these methods:(since size of file is 20MB +)

### Method1:-

Using snowsql

```

cycle Bin Deloitte - LogMeln ... ARC FINAL ASSIGNME... NavHistory...
Deloitte VPN Multi-Factor Authentica... Extra files FundCateg...
Google Chrome Hashedin Certificates Weekly MutualFund...
Postman Deloitte Remote ... AnypointS... - Shortcut AMC.csv
Microsoft Office Exc... Microsoft Office Po... Talend Open Studio for ...
Microsoft Word Microsoft Powerpoint NavHistory...
Atom

Command Prompt - snowsql -a kj76050.ap-south-1 -u shwethat
+-----+
| Statement executed successfully. |
+-----+
1 Row(s) produced. Time Elapsed: 0.078s
shwethat#TRANSFORM_WH@MAINASSIGNMENT.DBT>create or replace file format csvfileformat
    type = 'CSV'
    field_delimiter = ','
    skip_header = 1;
+-----+
| status |
+-----+
| File format CSVFILEFORMAT successfully created. |
+-----+
1 Row(s) produced. Time Elapsed: 0.393s
shwethat#TRANSFORM_WH@MAINASSIGNMENT.DBT>create stage stag_navhistory file_format = csvfileformat;
+-----+
| status |
+-----+
| Stage area STAG_NAVHISTORY successfully created. |
+-----+
1 Row(s) produced. Time Elapsed: 0.113s
shwethat#TRANSFORM_WH@MAINASSIGNMENT.DBT>put file:///C:/Users/shwethat/Desktop/navHistory.csv @stag_navhistory auto_compr
    ess=true;
-00001 (n/a): GET with url b'https://sfc-in-ds1-39-customer-stage.s3.amazonaws.com/https://sfc-in-ds1-39-customer-stage.s3.amazonaws.com/?accelerate' failed for exceeding maximum retries.
shwethat#TRANSFORM_WH@MAINASSIGNMENT.DBT>

```

Fails because of some s3 error in the command prompt. Put method is not working

## Method2:

### Using Python Script to load data into snowflake tables

```

File Edit Format Run Options Window Help
import pandas as pd
import snowflake.connector
data = pd.read_csv(r'C:\Users\shwethat\Desktop\navHistory.csv')
df = pd.DataFrame(data, columns= ['code','nav_date','nav','repurchase_price','sale_price'])
print(df)

#gets version
conn = snowflake.connector.connect(
    user = 'Shwethat',
    password = 'Sunrise.123',
    account = 'kj76050.ap-south-1'
)

#create cursor
cur = conn.cursor()

#Execute query
cur.execute("use warehouse transform_wh")
cur.execute("use mainassignment")

#insert data to table
for row in df.itertuples():
    cur.execute(SQLCommand = "insert into navhistory values(code,nav_date,nav,repurchase_price,sale_price) values(?, ?, ?, ?, ?)",row.code,row.nav_date,row.nav,row.repurchase_price,row.sale_price)
    conn.commit()

#fetch result
SQLShow = "select * from navhistory"
cur.execute(SQLShow)
conn.commit()
for record in cur:
    print(record)

```

## Query:

```

import pandas as pd
import snowflake.connector

```

```

data = pd.read_csv(r'C:\Users\shweth\Desktop\navHistory.csv')
df = pd.DataFrame(data, columns= ['code','nav_date','nav','repurchase_price','sale_price'])
print(df)

#gets version
conn = snowflake.connector.connect(
    user = 'Shwethath',
    password = 'Sunrise.123',
    account = 'kj76050.ap-south-1'
)

#create cursor
cur = conn.cursor()

#Execute query
cur.execute("use warehouse transform_wh;")
cur.execute("use mainassignment")

#insert data to table
for row in df.itertuples():
    cur.execute(SQLCommand = "insert into navhistory
values(code,nav_date,nav,repurchase_price,sale_price)
values(?, ?, ?, ?, ?)",row.code,row.nav_date,row.nav,row.repurchase_price,row.sale_price)
conn.commit()

#fetch result
SQLShow = "select * from navhistory"
cur.execute(SQLShow)
conn.commit()
for record in cur:
    print(record)

```

Couldn't load data because of error in file format

The screenshot shows a Windows desktop environment. In the background, there is a File Explorer window titled 'Analytics\_Main\_Assignment1' showing a directory structure with various files and folders like 'analyses', 'dbt\_packages', 'logs', etc. In the foreground, there is a Python IDLE Shell window titled 'IDLE Shell 3.9.13'. The shell window contains a Python script named 'pythonnav.py' which is reading a CSV file named 'navhistory.csv'. The script uses pandas to read the CSV and then prints the first few rows of the DataFrame.

```

File Edit Shell Debug Options Window Help
File Edit Format Run Options Window Help
Python 3.9.13 (tags/v3.9.13:6de2ca5, May 17 2022, 16:36:42) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: C:\Users\shweth\Desktop\pythonnav.py =====
   code  nav_date  nav repurchase_price sale_price
0  120373  2018-12-03  688138.0      0        0
1  120373  2018-12-04  686524.0      0        0
2  120373  2018-12-05  681853.0      0        0
3  120373  2018-12-06  674551.0      0        0
4  120373  2018-12-07  680586.0      0        0
...
5448038  120798  2021-01-22  182736.0      0        0
5448039  120798  2021-01-25  182783.0      0        0
5448040  120798  2021-01-27  182818.0      0        0
5448041  120798  2021-01-28  182834.0      0        0
5448042  120798  2021-01-29  182850.0      0        0
[5448043 rows x 5 columns]
Traceback (most recent call last):
  File "C:\Users\shweth\Desktop\pythonnav.py", line 23, in <module>
    cur.execute(''SQLCommand = "insert into navhistory values(code,nav_date,nav,repurchase_price,sale_price) values(?, ?, ?, ?, ?)"'',row.code,row.nav_date,row.nav,row.repurchase_price,row.sale_price)
  File "C:\Users\shweth\AppData\Local\Programs\Python\Python39\lib\site-packages\snowflake\connector\cursor.py", line 656, in execute
    if params is not None and len(params) == 0:
TypeError: object of type 'int' has no len()
>>>

```

## Method3:

By chunking the csv file – I am using Python script to chunk the csv file navhistory.csv

The screenshot shows a Windows desktop environment. In the background, there is a File Explorer window titled 'Analytics\_Main\_Assignment1' showing a directory structure with various files and folders like 'analyses', 'dbt\_packages', 'logs', etc. In the foreground, there is a Python IDLE Shell window titled 'chukFile.py - C:\Users\shweth\Analytics\_Main\_Assignment1\mainassignment\chuk...'. The script reads a large CSV file 'NAVHistorynew.csv' in chunks of size 400000 and writes each chunk to a separate CSV file named 'NAVHistorynew.csv' followed by a batch number (e.g., 'NAVHistorynew.csv1', 'NAVHistorynew.csv2', etc.).

```

File Edit Format Run Options Window Help
File Edit Format Run Options Window Help
chunk_size = 400000
batch_no = 1
for chunk in pd.read_csv('NAVHistorynew.csv', chunksize=chunk_size):
    chunk.to_csv('NAVHistorynew.csv' + str(batch_no) + '.csv', index=False)
    batch_no +=1

```

Query:

```

import pandas as pd
chunk_size = 400000
batch_no = 1
for chunk in pd.read_csv('NAVHistorynew.csv', chunksize=chunk_size):
    chunk.to_csv('NAVHistorynew.csv' + str(batch_no) + '.csv', index=False)
    batch_no +=1

```

The screenshot shows the Snowflake Worksheet interface. A query is being run:

```
select * from "MAINASSIGNMENT", "DBT", "NAVHISTORY"
```

The results show 515,375 rows. The data preview table has columns: Row, CODE, NAV\_DATE, NAV, REPURCHASE\_PRICE, and SALE\_PRICE. The first 10 rows of data are as follows:

Row	CODE	NAV_DATE	NAV	REPURCHASE_PRICE	SALE_PRICE
1	139565	2019-01-11	119853	1	18
2	139565	2019-01-12	119877	15	20
3	139565	2019-01-13	119900	7	2
4	139565	2019-01-14	119940	44	16
5	139565	2019-01-15	119968	33	4
6	139565	2019-01-16	119993	47	23
7	139565	2019-01-17	120017	18	1
8	139565	2019-01-18	120043	6	25
9	139565	2019-01-19	120067	24	36
10	139565	2019-01-20	120090	32	18

Data Loaded successfully !!!

## Milestone-1:

Gather basic data insights from the given data:

1. Calculate Monthly average NAV, Repurchase & Sale Price for each scheme.

The screenshot shows the Atom code editor with two files open:

- schema.yml**: A DBT schema configuration file with the following content:

```
1 {{config(materialized='table')}}
2 select f.category as Scheme,date_part(Month, nav_date) as Month,date_part(Year, nav_date) as Year, avg(nav) as Avg_Nav ,av
3 from fundcategory f,mutualfund m, navhistory n
4 where f.id = m.category_id and m.code = n.code
5 group by f.category, date_part(Year, nav_date), date_part(Month, nav_date)
6
```
- milestone1monthlyaverage.sql**: A SQL query file with the following content:

```
mainassignment\models\example\milestone1monthlyaverage.sql 6:1
```

Query:

```

{{config(materialized='table')}}
select f.category as Scheme,date_part(Month, nav_date) as Month,date_part(Year, nav_date)
as Year, avg(nav) as Avg_Nav ,avg(repurchase_price) as Avg_Repurchase_Price,
avg(sale_price) as Avg_Sales_Price
from fundcategory f,mutualfund m, navhistory n
where f.id = m.category_id and m.code = n.code
group by f.category, date_part(Year, nav_date), date_part(Month, nav_date)

```

## Result:

The screenshot shows the Snowflake Worksheet interface with the following details:

- Worksheet Title:** New Worksheet
- Query:**

```

1 select * from "MAINASSIGNMENT"."DBT"."MILESTONE1MONTHLYAVERAGE"
2
3
4

```
- Results:** The results table displays 129 rows of data with the following columns:
 

Row	SCHEME	MONTH	YEAR	AVG_NAV	AVG_REPURCHASE_PRICE	AVG_SALES_PRICE
1	Close Ended Schemes ( Incom... )	1	2019	314482.366535	25.541982	25.487840
2	Close Ended Schemes ( ELSS )	1	2019	133344.064545	25.442403	25.236665
3	Interval Fund Schemes ( Grow... )	1	2019	100671.282609	21.380435	24.739130
4	Open Ended Schemes ( Liquid )	2	2019	16518968.566964	25.691964	24.058036
5	Open Ended Schemes ( Mone... )	2	2019	113966.712500	23.925000	25.512500
6	Open Ended Schemes ( ELSS )	2	2019	441769.225000	27.112500	23.162500
7	Open Ended Schemes ( Equity... )	2	2019	585268.562931	25.750000	26.275000
8	Open Ended Schemes ( Equity... )	2	2019	673483.207477	25.412617	25.377850
9	Open Ended Schemes ( Equity... )	2	2019	363527.777083	25.058333	25.381250
10	Open Ended Schemes ( Equity... )	2	2019	514550.040625	25.118750	25.308594
11	Open Ended Schemes ( Equity... )	2	2019	323738.621622	25.868243	25.610811

shweth@USBLRSHWETHT1 MINGW64 ~/Desktop/Analytics\_Main\_Assignment1/mainassignment

```
$ dbt run
11:21:22 Running with dbt=1.1.0
11:21:22 Found 3 models, 1 test, 0 snapshots, 0 analyses, 181 macros, 0 operations, 0 seed files, 0 sources, 0 exposures, 0 metrics
11:21:24 Concurrency: 1 threads (target='dev')
Starting with...
11:21:24 1 of 3 START table model dbt.milestone1monthlyaverage ..... [RUN]
11:21:24 1 of 3 OK created table model dbt.milestone1monthlyaverage ..... [SUCCESS 1 in 1.65s]
11:21:24 2 of 3 START table model dbt.my_first_dbt_model ..... [RUN]
11:21:24 2 of 3 OK created table model dbt.my_first_dbt_model ..... [SUCCESS 1 in 1.44s]
11:21:27 3 of 3 START view model dbt.my_second_dbt_model ..... [RUN]
11:21:29 3 of 3 OK created view model dbt.my_second_dbt_model ..... [SUCCESS 1 in 1.06s]
11:21:29 Finished running 2 table models, 1 view model in 6.14s.
11:21:29 Completed successfully
11:21:29 Done. PASS=3 WARN=0 ERROR=0 SKIP=0 TOTAL=3
```

shweth@USBLRSHWETHT1 MINGW64 ~/Desktop/Analytics\_Main\_Assignment1/mainassignment

```
$ dbt test
11:21:41 Running with dbt=1.1.0
11:21:41 Found 3 models, 1 test, 0 snapshots, 0 analyses, 181 macros, 0 operations, 0 seed files, 0 sources, 0 exposures, 0 metrics
11:21:42 Concurrency: 1 threads (target='dev')
11:21:42 1 of 1 START test not_null_milestone1monthlyaverage_Scheme ..... [RUN]
11:21:43 1 of 1 PASS not_null_milestone1monthlyaverage_Scheme ..... [PASS in 0.82s]
11:21:43 Finished running 1 test in 1.81s.
11:21:43 Completed successfully
11:21:43 Done. PASS=1 WARN=0 ERROR=0 SKIP=0 TOTAL=1
```

shweth@USBLRSHWETHT1 MINGW64 ~/Desktop/Analytics\_Main\_Assignment1/mainassignment

```
$ |
```

Type here to search

## 2. Find out each scheme's Max and Min NAV value and Date it occurred

File Edit View Selection Find Packages Help

Project

schema.yml

```
1 {{config(materialized='table')}}
2 select category as Scheme, min(nav) as Min_Nav, max(nav) as Max_Nav, nav_date as Date
3 from navhistory, fundcategory, mutualfund
4 where mutualfund.code = navhistory.code and
5 fundcategory.id = mutualfund.category_id
6 group by category, nav_date
7
```

milestone1\_2\_schememaxminnav.sql

milestone1monthlyaverage.sql

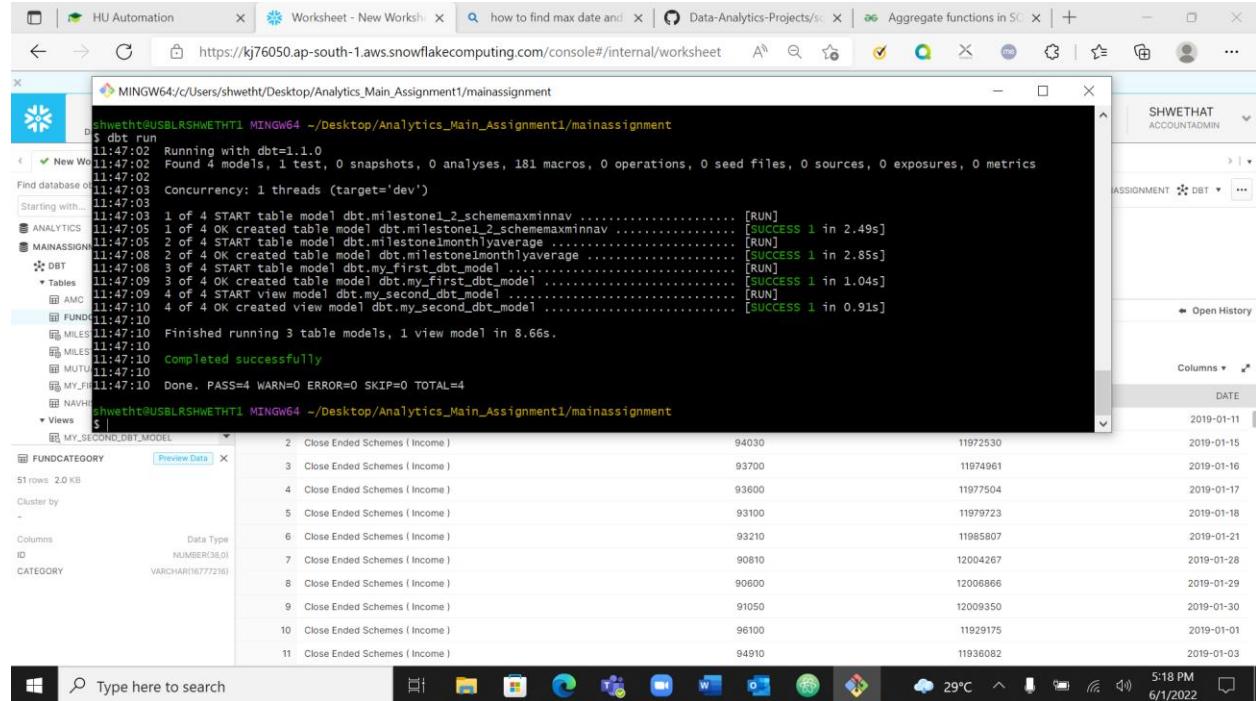
mainassignment\models\example\milestone1\_2\_schememaxminnav.sql

Type here to search

**Query:**

```
{{config(materialized='table')}}
select category as Scheme, min(nav) as Min_Nav, max(nav) as Max_Nav, nav_date as Date
from navhistory, fundcategory, mutualfund
```

where mutualfund.code = navhistory.code and  
 fundcategory.id = mutualfund.category\_id  
 group by category, nav\_date



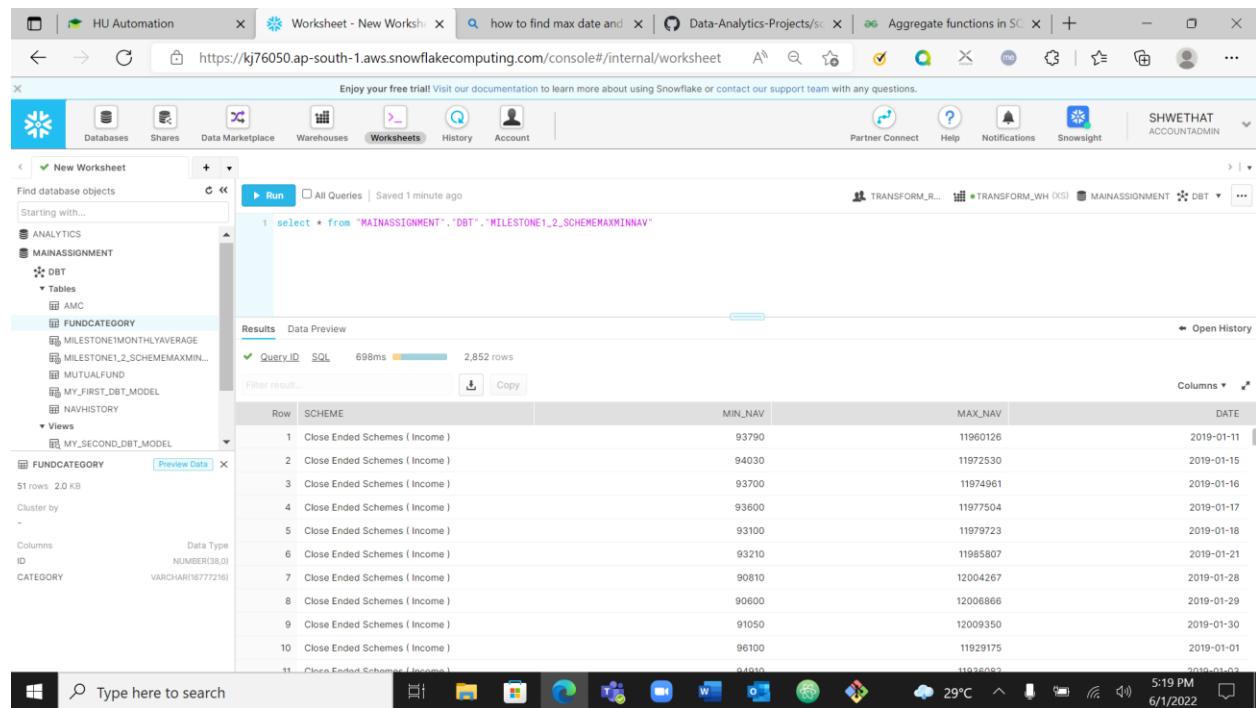
```

MINGW64:/c/Users/shweth/Desktop/Analytics_Main_Assignment1/mainassignment
$ dbt run
11:47:02 Running with dbt=1.1.0
11:47:02 Found 4 models, 1 test, 0 snapshots, 0 analyses, 181 macros, 0 operations, 0 seed files, 0 sources, 0 exposures, 0 metrics
11:47:02 Concurrency: 1 threads (target='dev')
Starting with...
11:47:03 1 of 4 START table model dbt.milestone1_2_schememaxminav ..... [RUN]
11:47:05 1 of 4 OK created table model dbt.milestone1_2_schememaxminav ..... [SUCCESS 1 in 2.49s]
11:47:05 2 of 4 START table model dbt.milestone1monthlyaverage ..... [RUN]
11:47:08 2 of 4 OK created table model dbt.milestone1monthlyaverage ..... [SUCCESS 1 in 2.85s]
11:47:08 3 of 4 START table model dbt.my_first_dbt_model ..... [RUN]
11:47:09 3 of 4 OK created table model dbt.my_first_dbt_model ..... [SUCCESS 1 in 1.04s]
11:47:10 4 of 4 START view model dbt.my_second_dbt_model ..... [RUN]
11:47:10 4 of 4 OK created view model dbt.my_second_dbt_model ..... [SUCCESS 1 in 0.91s]
11:47:10 Finished running 3 table models, 1 view model in 8.66s.
11:47:10 Completed successfully
11:47:10 Done. PASS=4 WARN=0 ERROR=0 SKIP=0 TOTAL=4
  
```

Views

FUNDCATEGORY

	SCHEME	MIN_NAV	MAX_NAV	DATE
1	Close Ended Schemes ( Income )	93790	11960126	2019-01-11
2	Close Ended Schemes ( Income )	94030	11972530	2019-01-15
3	Close Ended Schemes ( Income )	93700	11974961	2019-01-16
4	Close Ended Schemes ( Income )	93600	11977504	2019-01-17
5	Close Ended Schemes ( Income )	93100	11979723	2019-01-18
6	Close Ended Schemes ( Income )	93210	11985807	2019-01-21
7	Close Ended Schemes ( Income )	90810	12004267	2019-01-28
8	Close Ended Schemes ( Income )	90600	12006866	2019-01-29
9	Close Ended Schemes ( Income )	91050	12009350	2019-01-30
10	Close Ended Schemes ( Income )	96100	11929175	2019-01-01
11	Close Ended Schemes ( Income )	94910	11936082	2019-01-03



Enjoy your free trial! Visit our documentation to learn more about using Snowflake or contact our support team with any questions.

Databases Shares Data Marketplace Warehouses Worksheets History Account

TRANSFORM\_R... TRANSFORM\_WH (X) MAINASSIGNMENT DBT ...

New Worksheet

Find database objects Starting with...

ANALYTICS MAINASSIGNMENT

DBT

Tables

FUNDCATEGORY

MILESTONE1MONTHLYAVERAGE MILESTONE1\_2\_SCHEMAMINNAV MUTUALFUND MY\_FIRST\_DBT\_MODEL

NAV HISTORY

Views

FUNDCATEGORY

Results Data Preview

Query\_ID SQL 698ms 2,852 rows

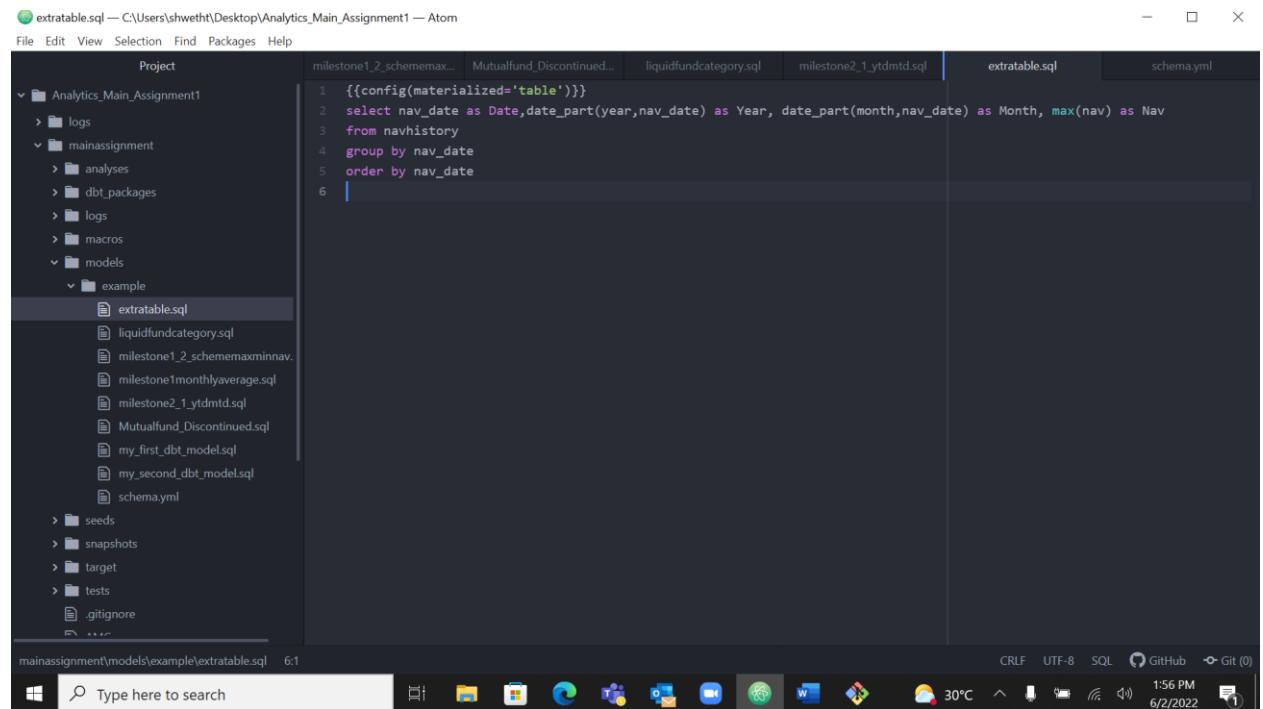
Row	SCHEME	MIN_NAV	MAX_NAV	DATE
1	Close Ended Schemes ( Income )	93790	11960126	2019-01-11
2	Close Ended Schemes ( Income )	94030	11972530	2019-01-15
3	Close Ended Schemes ( Income )	93700	11974961	2019-01-16
4	Close Ended Schemes ( Income )	93600	11977504	2019-01-17
5	Close Ended Schemes ( Income )	93100	11979723	2019-01-18
6	Close Ended Schemes ( Income )	93210	11985807	2019-01-21
7	Close Ended Schemes ( Income )	90810	12004267	2019-01-28
8	Close Ended Schemes ( Income )	90600	12006866	2019-01-29
9	Close Ended Schemes ( Income )	91050	12009350	2019-01-30
10	Close Ended Schemes ( Income )	96100	11929175	2019-01-01
11	Close Ended Schemes ( Income )	94910	11936082	2019-01-03

Successful!!

## Milestone-2 :

1. Simple: Mutual fund performance - including YTD, MTD, 1 Year and Since Inception returns. See the PPT for formula to calculate this

Here In order to get YTD , MTD I am going to create one extra table as subtable through which I will get YTD,MTD and Inception



The screenshot shows the Atom code editor interface. The left sidebar displays a project structure for 'Analytics\_Main\_Assignment1' containing 'logs', 'mainassignment' (with 'analyses', 'dbt\_packages', 'logs', 'macros', 'models', and 'example' subfolders), and 'seeds', 'snapshots', 'target', 'tests', and '.gitignore'. The right pane shows an SQL file named 'extratable.sql' with the following content:

```
1 {{config(materialized='table')}}
2 select nav_date as Date,date_part(year,nav_date) as Year, date_part(month,nav_date) as Month, max(nav) as Nav
3 from navhistory
4 group by nav_date
5 order by nav_date
6
```

The status bar at the bottom indicates the file is at line 6, character 1. The system tray shows the date as 6/2/2022 and the time as 1:56 PM.

Query:

```
{{config(materialized='table')}}

select nav_date as Date,date_part(year,nav_date) as Year,
date_part(month,nav_date) as Month, max(nav) as Nav
from navhistory
group by nav_date
order by nav_date
```

The screenshot shows the Atom code editor interface. The title bar reads "milestone2\_1\_ytdmtd.sql — C:\Users\shweth\Desktop\Analytics\_Main\_Assignment1 — Atom". The menu bar includes File, Edit, View, Selection, Find, Packages, Help. The left sidebar is titled "Project" and shows a tree structure for "Analytics\_Main\_Assignment1" with various sub-directories like logs, mainassignment, analyses, dbt\_packages, macros, models, example, seeds, snapshots, target, tests, and .gitignore. The right pane displays the contents of the "milestone2\_1\_ytdmtd.sql" file, which contains a DBT SQL script. The status bar at the bottom shows file statistics (CRLF, UTF-8, SQL), GitHub integration, and system information (30°C, 156 PM, 6/2/2022).

```
1 {{config(materialized='table')}}  
2 select nav,  
3 sum(nav) over(partition by year order by date) as YTD,  
4 sum(nav) over(partition by year,month order by date) as MTD,  
5 pow((nav/nav),1) -1 as Inception  
6 from "MAINASSIGNMENT"."DBT"."EXTRATABLE"
```

Query:

```
{{config(materialized='table')}}
```

```
select nav,  
sum(nav) over(partition by year order by date) as YTD,  
sum(nav) over(partition by year,month order by date) as MTD,  
pow((nav/nav),1) -1 as Inception  
from "MAINASSIGNMENT"."DBT"."EXTRATABLE"
```

Row	NAV	YTD	MTD	INCEPTION
1	44521411	44521411	44521411	0
2	44530602	89052013	89052013	0
3	1260854244	1349906257	1349906257	0
4	115521813	1465428070	1465428070	0
5	114790653	1580218723	1580218723	0
6	112856162	1693074885	1693074885	0
7	113841363	1806916248	1806916248	0
8	44587653	1851503901	1851503901	0
9	44596710	1896100611	1896100611	0
10	111658918	2007759529	2007759529	0
11	112304625	2120064154	2120064154	0

Successful!!!

### 3. Medium: List of mutual funds which were discontinued in a given year

```

Mutalfund_Discontinued.sql — C:\Users\shweth\Desktop\Analytics_Main_Assignment1 — Atom
File Edit View Selection Find Packages Help
Project
Analytics_Main.Assignment1
  logs
  mainassignment
    analyses
    dbt_packages
    logs
    macros
    models
      example
      seeds
      snapshots
      target
      tests
      .gitignore
      AMC.csv
      chukfile.py
      dbt_project.yml
      FundCategory.csv
      MutualFund.csv
      NAVHistorynew.csv
      NAVHistorynew.csv1.csv
      NAVHistorynew.csv2.csv
      pythonnav.py
      README.md
mainassignment\models\example\Mutalfund_Discontinued.sql 7:1
schema.yaml milestone1_2_schememaxminnav.sql Mutualfund_Discontinued.sql milestone1monthlyaverage.sql
1 {{config(materialized='table')}}
2 select name as Mutual_Funds
3   from navhistory, mutualfund
4  where mutualfund.code = navhistory.code
5 group by name
6 having max(date_part(year,nav_date))='2018'

```

Query:

```

{{config(materialized='table')}}
select name as Mutual_Funds
from navhistory, mutualfund
where mutualfund.code = navhistory.code
group by name
having max(date_part(year,nav_date))='2018'

```

```

MINGW64:/c/Users/shweth/Desktop/Analytics_Main_Assignment1/mainassignment
_Discontinued.sql' which was not found
15:59:41 Found 5 models, 1 test, 0 snapshots, 0 analyses, 181 macros, 0 operations, 0 seed files, 0 sources, 0 exposures, 0 metrics
15:59:41   Concurrency: 1 threads (target='dev')
15:59:43 1 of 5 START table model dbt.MutualFund_Discontinued ...
15:59:45 1 of 5 OK created table model dbt.MutualFund_Discontinued ...
15:59:45 2 of 5 START table model dbt.milestone1_2_schememaxminnav ...
15:59:46 2 of 5 OK created table model dbt.milestone1_2_schememaxminnav ...
15:59:46 3 of 5 START table model dbt.milestone1monthlyaverage ...
15:59:48 3 of 5 OK created table model dbt.milestone1monthlyaverage ...
15:59:48 4 of 5 START table model dbt.my_first_dbt_model ...
15:59:49 4 of 5 OK created table model dbt.my_first_dbt_model ...
15:59:49 5 of 5 START view model dbt.my_second_dbt_model ...
15:59:50 5 of 5 OK created view model dbt.my_second_dbt_model ...
15:59:50 Finished running 4 table models, 1 view model in 9.04s.
15:59:50
15:59:50 Completed successfully
15:59:50
15:59:50 Done. PASS=5 WARN=0 ERROR=0 SKIP=0 TOTAL=5

shweth@USBLRSHWETH1 MINGW64 ~/Desktop/Analytics_Main_Assignment1/mainassignment
$ dbt run
16:00:38 Running with dbt=1.1.0
16:00:38 [WARNING]: Did not find matching node for patch with name 'MutualFund_Discontinued.sql' in the 'models' section of file 'models\example\schema.yml'
16:00:38 Found 5 models, 1 test, 0 snapshots, 0 analyses, 181 macros, 0 operations, 0 seed files, 0 sources, 0 exposures, 0 metrics
16:00:38   Concurrency: 1 threads (target='dev')
16:00:41
16:00:43 1 of 5 START table model dbt.MutualFund_Discontinued ...
16:00:43 1 of 5 OK created table model dbt.MutualFund_Discontinued ...
16:00:43 2 of 5 START table model dbt.milestone1_2_schememaxminnav ...
16:00:45 2 of 5 OK created table model dbt.milestone1_2_schememaxminnav ...
16:00:45 3 of 5 START table model dbt.milestone1monthlyaverage ...
16:00:47 3 of 5 OK created table model dbt.milestone1monthlyaverage ...
16:00:47 4 of 5 START table model dbt.my_first_dbt_model ...
16:00:47 4 of 5 OK created table model dbt.my_first_dbt_model ...
16:00:56 5 of 5 START view model dbt.my_second_dbt_model ...
16:00:56 5 of 5 OK created view model dbt.my_second_dbt_model ...
16:00:56 Finished running 4 table models, 1 view model in 18.54s.
16:00:57
16:00:57 Completed successfully
16:00:57
16:00:57 Done. PASS=5 WARN=0 ERROR=0 SKIP=0 TOTAL=5

```

shweth@USBLRSHWETH1 MINGW64 ~/Desktop/Analytics\_Main\_Assignment1/mainassignment Worksheet - New Worksheet (1/1) in 1.19s - Done and 11 more pages - Work - Microsoft Edge

Row	MUTUAL_FUNDS
1	Sundaram Multi Asset Fund Dividend Payout
2	HSBC Short Duration Fund - Quarterly Dividend Direct
3	Indiabulls Gilt Fund - Indirect Plan - Growth Option
4	Indiabulls Gilt Fund - Direct Plan - Monthly Dividend Option
5	Indiabulls Gilt Fund - Direct Plan - Growth Option
6	Indiabulls Gilt Fund - Direct Plan - Fortnightly Dividend Option
7	Indiabulls Gilt Fund - Direct Plan - Weekly Dividend Option
8	HSBC Dynamic Asset Allocation Fund-Dividend Direct
9	HSBC Dynamic Asset Allocation Fund-Growth Direct
10	Aditya Birla Sun Life Fixed Term Plan - Series NB (1099 days) - Direct Plan-Growth
11	Axis Fixed Term Plan - Series 94 (177 Days) - Direct Plan - Growth Option

Successful!!!

#### 4. Medium: Weekly, monthly and ytd results for all funds in liquid fund category

```

File Edit View Selection Find Packages Help
Project
Analytics_Main_Assignment1
  logs
  mainassignment
    analyses
    dbt_packages
    logs
    macros
    models
      example
        seeds
        snapshots
        target
        tests
        .gitignore
        AMC.csv
        chukFile.py
        dbt_project.yml
        FundCategory.csv
        Mutual_Fund_Report.pdf
        MutualFund.csv
        NAVHistorynew.csv
        NAVHistorynew.csv1.csv
        NAVHistorynew.csv2.csv
        pythonnnav.py
milestone1_2_schemamaxminna... Mutualfund_Discontinued.sql extratable.sql liquidfundcategory.sql schemayml
1 {{config(materialized='table')}}
2 select code,
3   sum(nav) over(partition by date_part(year,nav_date) order by nav_date) as YTD,
4   sum(nav) over(partition by date_part(month,nav_date) order by nav_date) as MTD
5   from navhistory
6   where code in (select m.code
7     from fundcategory f, mutualfund m
8       where f.category like '%Liquid%' and m.category_id = f.id)
9
mainassignment\models\example\liquidfundcategory.sql 9:1
Windows Taskbar: Type here to search, File Explorer, Task View, Start, Taskbar icons, 30°C, 1:39 PM, 6/2/2022, Battery

```

Enjoy your free trial! Visit our documentation to learn more about using Snowflake or contact our support team with any questions.

New Worksheet

Find database objects Starting with...

- LIQUIDFUNDCATEGORY
- MILESTONE1MONTHLYAVERAGE
- MILESTONE1\_2\_SCHEMAMAXMIN...
- MUTUALFUND
- MUTUALFUND\_DISCONTINUED
- MY\_FIRST\_DBT\_MODEL
- NAVHISTORY
- Views
- MY\_SECOND\_DBT\_MODEL
- INFORMATION\_SCHEMA
- MINASSIGNMENT1
- INFORMATION\_SCHEMA
- MUTUALFUND

12,261 rows 323.5 kB Cluster by -

Columns Data Type

CODE	NUMBER(38,0)	AMC_ID	NUMBER(38,0)	CATEGORY_ID	NUMBER(38,0)	NAME	VARCHAR(16777216)	ISIN_GROWTH	VARCHAR(16777216)	ISIN_DIVIDEND_PAYOUT	VARCHAR(16777216)
1		101397		101397				146786953441			4859299008
2		101397		101397				150040497013			8112643480
3		101397		101397				154900825042			129731715099
4		101397		101397				159701683505			17834029972
5		101397		101397				164623054971			22695401438
6		101397		101397				169485014420			27557360887
7		101397		101397				174347745928			32420092395
8		101397		101397				179211086006			37283432473
9		101397		101397				182467593317			40539939784
10		101397		101397				187332123296			45404469763

Query:

```
{{config(materialized='table')}}
```

```
select code,
       sum(nav) over(partition by date_part(year,nav_date) order by nav_date) as YTD,
       sum(nav) over(partition by date_part(month,nav_date) order by nav_date) as MTD
  from navhistory
 where code in (select m.code
                  from fundcategory f, mutualfund m
                 where f.category like '%Liquid%' and m.category_id = f.id)
```

Worksheet - New Worksheet (1/1) | New tab

Enjoy your free trial! Visit our documentation to learn more about using Snowflake or contact our support team with any questions.

```
08:07:42 Running with dbt=1.1.0
08:07:42 Found 7 models, 1 test, 0 snapshots, 0 analyses, 181 macros, 0 operations, 0 seed files, 0 sources, 0 exposures, 0 metrics
08:07:42 Concurrency: 1 threads (target='dev')
08:07:43 1 of 7 START table model dbt.MutualFund_Discontinued ..... [RUN]
08:07:46 1 of 7 OK created table model dbt.MutualFund_Discontinued ..... [SUCCESS 1 in 3.06s]
08:07:46 2 of 7 START table model dbt.Extratable ..... [RUN]
08:07:49 2 of 7 OK created table model dbt.Extratable ..... [SUCCESS 1 in 2.66s]
08:07:49 3 of 7 START table model dbt.LiquidFundCategory ..... [RUN]
08:07:51 3 of 7 OK created table model dbt.LiquidFundCategory ..... [SUCCESS 1 in 1.83s]
08:07:51 4 of 7 START table model dbt.Milestone1_2_SchemeMaxMinNav ..... [RUN]
08:07:52 4 of 7 OK created table model dbt.Milestone1_2_SchemeMaxMinNav ..... [SUCCESS 1 in 1.11s]
08:07:52 5 of 7 START table model dbt.Milestone1MonthAvg ..... [RUN]
08:07:53 5 of 7 OK created table model dbt.Milestone1MonthAvg ..... [SUCCESS 1 in 1.53s]
08:07:53 6 of 7 START table model dbt.MyFirst_DbT_Model ..... [RUN]
08:07:55 6 of 7 OK created table model dbt.MyFirst_DbT_Model ..... [SUCCESS 1 in 1.23s]
08:07:55 7 of 7 START view model dbt.MySecond_DbT_Model ..... [RUN]
08:07:55 7 of 7 OK created view model dbt.MySecond_DbT_Model ..... [SUCCESS 1 in 0.55s]
08:07:55 Finished running 6 table models, 1 view model in 13.44s.
08:07:55 Completed successfully
08:07:55 Done. PASS=7 WARN=0 ERROR=0 SKIP=0 TOTAL=7
```

shweth@USBLRSHWETH1: MINGW64 ~/Desktop/Analytics\_Main\_Assignment1/mainassignment

Column	Data Type	Value	Value	Value
CODE	NUMBER(38,0)	4	101397	159761683505
AMC_ID	NUMBER(38,0)	5	101397	164623054971
CATEGORY_ID	NUMBER(38,0)	6	101397	169485014420
NAME	VARCHAR(16777216)	7	101397	174347745928
ISIN_GROWTH	VARCHAR(16777216)	8	101397	179211086006
ISIN_DIVIDEND_PAYOUT	VARCHAR(16777216)	9	101397	182467593317
		10	101397	187332123296

Successfull!!

- Medium: Best and worst performing fund in a given category on the basis of 1 year returns

Here for Best performance I am taking separate model to show best performance fund.

bestperformance.sql — C:\Users\shweth\Desktop\Analytics\_Main\_Assignment1 — Atom

File Edit View Selection Find Packages Help

Project

Analytics\_Main\_Assignment1

- logs
- mainassignment
- analyses
- dbt\_packages
- logs
- macros
- models
- example
- seeds
- snapshots
- target
- tests
- .gitignore
- AMC.csv
- chukfile.py
- dbt\_project.yml
- FundCategory.csv
- Mutual\_Fund\_Report.pdf
- MutualFund.csv
- NAVHistorynew.csv
- NAVHistorynew.csv1.csv
- NAVHistorynew.csv2.csv
- pythonnav.py

mainassignment\models\example\bestperformance.sql 5:1

CRLF UTF-8 SQL GitHub Git (0)

Windows Taskbar: Type here to search, File Explorer, Edge, File History, Mail, Photos, OneDrive, Task View, Taskbar icons, Weather, 30°C, 2:11 PM, 6/2/2022, Notifications

Worksheet - New Worksheet (1/1) New tab

https://kj76050.ap-south-1.aws.snowflakecomputing.com/console#/internal/worksheet

Enjoy your free trial! Visit our documentation to learn more about using Snowflake or contact our support team with any questions.

Databases Shares Data Marketplace Warehouses Worksheets History Account SHWETHAT ACCOUNTADMIN

New Worksheet

Find database objects Starting with... ANALYTICS MAINASSIGNMENT DBT Tables BESTPERFORMANCE EXTRATABLE FUNDCATEGORY LIQUIDFUNDCATEGORY MILESTONE1MONTHLYAVERAGE MILESTONE1\_2\_SCHEMEMAXMIN MILESTONE2\_1\_YTDMTD MUTUALFUND

EXTRATABLE Preview Data

90 rows 4.0 KB Cluster by

Columns	Data Type
DATE	DATE
YEAR	NUMBER(4,0)
MONTH	NUMBER(2,0)
NAV	NUMBER(18,0)

Run All Queries | Saved 0 seconds ago

1. select \* from 'MAINASSIGNMENT'.'DBT'.'BESTPERFORMANCE'

Results Data Preview

Query\_ID SQL 166ms 1 rows

Row	NAME
1	SAHARA BANKING & FINANCIAL SERVICES FUND- GROWTH - Direct

Open History Columns

Windows Taskbar: Type here to search, File Explorer, Edge, File History, Mail, Photos, OneDrive, Task View, Taskbar icons, Weather, 30°C, 2:11 PM, 6/2/2022, Notifications

For worst performance

The screenshot shows the Snowflake Worksheet interface. The left sidebar displays the schema structure under 'New Worksheet'. A query is run against the 'WORSTPERFORMANCE' table:

```
1 select * from "MAINASSIGNMENT"."DBT"."WORSTPERFORMANCE"
```

The results show 1,103,490 rows. The first few rows of the result set are:

Row	NAME
1	Aditya Birla Sun Life Fixed Term Plan - Series PD (1177 days) - Direct Plan-Quarterly Dividend
2	Aditya Birla Sun Life Fixed Term Plan - Series PF (1148 days) - Regular Plan - Normal Dividend
3	Aditya Birla Sun Life Fixed Term Plan - Series PF (1148 days) - Regular Plan - Growth
4	ICICI Prudential US Bluechip Equity Fund - Direct Plan - Dividend
5	Aditya Birla Sun Life Fixed Term Plan - Series PF (1148 days) - Direct Plan - Quarterly Dividend
6	ADITYA BIRLA SUN LIFE OVERNIGHT FUND-REGULAR PLAN-DAILY DIVIDEND REINVESTMENT
7	UTI - Fixed Term Income Fund Series XXXI-VIII (1153 Days) - Regular Plan - Quarterly Dividend Option
8	Aditya Birla Sun Life Fixed Term Plan - Series PE (1159 days) - Regular Plan-Quarterly Dividend
9	Aditya Birla Sun Life Fixed Term Plan - Series PE (1159 days) - Regular Plan-Normal Dividend
10	Aditya Birla Sun Life Fixed Term Plan - Series PE (1159 days) - Regular Plan-Growth
11	Aditya Birla Sun Life Fixed Term Plan - Series PE (1159 days) - Direct Plan-Quarterly Dividend

The screenshot shows the Atom code editor with the file 'worstperformance.sql' open. The project structure on the left includes 'Analytics\_Main\_Assignment1' with subfolders like 'logs', 'mainassignment', 'models', and 'example'. The code editor shows the following SQL query:

```
1 {{config(materialized='table')}}
2 select name
3 from "MAINASSIGNMENT"."DBT"."MILESTONE2_1_YTDMDT", "MAINASSIGNMENT"."DBT"."MUTUALFUND"
4 order by nav asc
```

The status bar at the bottom indicates the file is at line 5, character 1.

Successful!!!!

## Milestone-3 : Visualize analytics and build dashboards via Data Studio

2. Count of mutual funds discontinued each year.

Count of mutual funds discontinued each year.

MUTUAL_FUNDS	YEAR
UTI Fixed Interval Income Fund (Annual Interval Plan) Series IV - Insta Growth Option	2018
UTI FTIF Series XXIII-XI (1100 Days)- Regular Plan - Growth Option	2018
UTI FTIF Series XXIII-XI (1100 Days)- Regular Plan - Annual Div Option	2018
UTI FTIF Series XXIII-XI (1100 Days)- Direct Plan - Growth Option	2018
UTI FTIF Series XXIII-XI (1100 Days)- Direct Plan - Annual Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Regular Plan - Quarterly Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Regular Plan - Maturity Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Regular Plan - Growth Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Regular Plan - Annual Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Direct Plan - Quarterly Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Direct Plan - Maturity Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Direct Plan - Growth Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Direct Plan - Flex Div Option	2018
UTI FTIF Series XXIII-X (1100 Days)- Direct Plan - Annual Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Regular Plan - Quarterly Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Regular Plan - Maturity Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Regular Plan - Growth Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Regular Plan - Flex Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Regular Plan - Annual Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Direct Plan - Quarterly Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Direct Plan - Maturity Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Direct Plan - Growth Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Direct Plan - Flex Div Option	2018
UTI FTIF Series XXIII-IX (1100 Days)- Direct Plan - Annual Div Option	2018

3. Distribution of latest mutual funds based on category

