JSON to RDBMS

In this experiment, we will learn to ingest JSON files into Snowflake.

1. Note, you must use SnowSQL for this, thus open SnowSQL in the terminal

Open Terminal
Enter: snowsql -a <account> -u <user>
Where account is your account id and user is your user id
Enter your password when prompted

2. Download sales.json to local environment (remember location)

```
{"location": {"state_city": "MA-Lexington","zip": "40503"},"sale_date":
"2017-3-5","price": "275836"}
{"location": {"state_city": "MA-Belmont","zip": "02478"},"sale_date":
"2017-3-17","price": "392567"}
{"location": {"state_city": "MA-Winchester","zip": "01890"},"sale_date":
"2017-3-21","price": "389921"}
```

3. Create database and home_sales table

```
USE ROLE ACCOUNTADMIN;

CREATE OR REPLACE DATABASE SNOWTEST;
CREATE OR REPLACE SCHEMA SNOWTEST.PUBLIC;

CREATE OR REPLACE TEMPORARY TABLE SNOWTEST.PUBLIC.home_sales (
  city STRING,
  zip STRING,
  state STRING,
  type STRING DEFAULT 'Residential',
  sale_date timestamp_ntz,
  price STRING
  );
```

4. Create file format to hold the JSON file

```
CREATE OR REPLACE FILE FORMAT sf_tut_json_format
  TYPE = JSON;
```

5. Create stage for ingesting external file

```
CREATE OR REPLACE TEMPORARY STAGE sf_tut_stage
 FILE_FORMAT = sf_tut_json_format;
```

6. Use PUT command to place the local file into Snowflake

```
PUT 'file://G:/Shared drives/IDSTS Shared Drive/Innovation In Software/Citi
Training/experiments/data-load-internal/sales.json' @sf_tut_stage
AUTO_COMPRESS=TRUE;
```

- Windows if there are no spaces in the file path
    - PUT file://C:\<file_path>\sales.json @sf_tut_stage AUTO_COMPRESS=TRUE;

- Windows if there are spaces in the file path
    - PUT 'file://C:/<file_path>/sales.json' @sf_tut_stage AUTO_COMPRESS=TRUE;

- Linux/MacOS
    - PUT file://C:/<file_path>/sales.json @sf_tut_stage AUTO_COMPRESS=TRUE

7. Copy file from stage into the database table

```
COPY INTO home_sales(city, state, zip, sale_date, price)
  FROM (SELECT SUBSTR($1:location.state_city,4),
               SUBSTR($1:location.state_city,1,2),
               $1:location.zip,
               to_timestamp_ntz($1:sale_date),
               $1:price
        FROM @sf_tut_stage/sales.json.gz t)
  ON_ERROR = 'continue';
```

8. View the file

```
SELECT * FROM SNOWTEST.PUBLIC.home_sales;
```

You can also ingest the JSON file without defining a table

9. Create new table but use the variant column

```
DROP TABLE home_sales;
CREATE OR REPLACE TABLE home_sales (
```

```
  json_column variant
  );
```

10. Ingest the data as before, but into the variant column

```
CREATE OR REPLACE FILE FORMAT sf_tut_json_format
  TYPE = JSON;

CREATE OR REPLACE TEMPORARY STAGE sf_tut_stage
 FILE_FORMAT = sf_tut_json_format;

PUT 'file://G:/Shared drives/IDSTS Shared Drive/Innovation In Software/Citi
Training/experiments/data-load-internal/sales.json' @sf_tut_stage
AUTO_COMPRESS=TRUE;
PUT file://C:/Users/kwame/Downloads/sales.json @sf_tut_stage
AUTO_COMPRESS=TRUE;

COPY INTO home_sales(json_column)
   FROM (SELECT *
         FROM @sf_tut_stage/sales.json.gz t)
   ON_ERROR = 'continue';
```

11. When you select the data, you'll see everything stored as a JSON

```
SELECT * FROM home_sales;
```

12. You can extract data from the JSON and create columns as well

```
SELECT SUBSTR($1:location.state_city,4),
              SUBSTR($1:location.state_city,1,2),
              $1:location.zip,
              to_timestamp_ntz($1:sale_date),
              $1:price
         FROM SNOWTEST.PUBLIC.home_sales t
```

13. Clear resources

```
DROP DATABASE SNOWTEST;
```

Test Your Skills

Using the students.json, ingest the file into Snowflake. Use both approaches above.