

Homework 8

1. Create a dbt project with snowflake connector (1 pt)

```
PROBLEMS OUTPUT DEBUG CONSOLE PORTS TERMINAL AZURE
• https://docs.getdbt.com/docs/configure-your-profile

One more thing:

Need help? Don't hesitate to reach out to us via GitHub issues or on Slack:

https://community.getdbt.com/

Happy modeling!

01:48:32 Setting up your profile.
Which database would you like to use?
[1] snowflake

(Don't see the one you want? https://docs.getdbt.com/docs/available-adapters)

Enter a number: 1
account (https://<this_value>.snowflakecomputing.com): kzcjwgd-rdb31003
user (dev username): niranjanrao07
[1] password
[2] keypair
[3] sso
Desired authentication type option (enter a number): 1
password (dev password):
role (dev role): ACCOUNTADMIN
warehouse (warehouse name): compute_wh
database (default database that dbt will build objects in): dev
schema (default schema that dbt will build objects in): analytics
threads (1 or more) [1]: 1
01:50:43 Profile build_test2 written to C:\Users\Niranja Rao\.dbt\profiles.yml using target's p
rofile_template.yml and your supplied values. Run 'dbt debug' to validate the connection.
(dbt_env) PS C:\Users\Niranja Rao> 
```

2. Set up Input models (2pt)(Input tables should be built as CTE,Show contents of models/input/user_session_channel.sql and input/session_timestamp.sql)

```
(dbt_env) PS C:\Users\Niranja Rao\build_test2> dbt run
• 02:37:39 Running with dbt=1.8.7
02:37:39 Registered adapter: snowflake=1.8.4
02:37:39 Unable to do partial parsing because a project config has changed
02:37:41 [WARNING]: Configuration paths exist in your dbt_project.yml file which do not apply t
o any resources.
There are 3 unused configuration paths:
- models.build_mau.output
- models.build_mau.input
- models.build_mau
02:37:41 Found 5 models, 1 snapshot, 6 data tests, 3 sources, 459 macros
02:37:41
02:37:44 Concurrency: 1 threads (target='dev')
02:37:44
02:37:44 1 of 5 START sql table model analytics.my_first_dbt_model ..... [RUN]
02:37:46 1 of 5 OK created sql table model analytics.my_first_dbt_model ..... [SUCC
ESS 1 in 2.13s]
02:37:46 2 of 5 START sql view model analytics.session_timestamp ..... [RUN]
02:37:47 2 of 5 OK created sql view model analytics.session_timestamp ..... [SUCC
ESS 1 in 1.16s]
02:37:47 3 of 5 START sql view model analytics.user_session_channel ..... [RUN]
02:37:49 3 of 5 OK created sql view model analytics.user_session_channel ..... [SUCC
ESS 1 in 1.09s]
02:37:49 4 of 5 START sql view model analytics.my_second_dbt_model ..... [RUN]
02:37:50 4 of 5 OK created sql view model analytics.my_second_dbt_model ..... [SUCC
ESS 1 in 1.09s]
02:37:50 5 of 5 START sql view model analytics.session_summary ..... [RUN]
02:37:51 5 of 5 OK created sql view model analytics.session_summary ..... [SUCC
ESS 1 in 1.37s]
02:37:51
02:37:51 Finished running 1 table model, 4 view models in 0 hours 0 minutes and 9.96 seconds (9
.96s).
02:37:51
02:37:51 Completed successfully
02:37:51
02:37:51 Done. PASS=5 WARN=0 ERROR=0 SKIP=0 TOTAL=5
• (dbt_env) PS C:\Users\Niranja Rao\build_test2> 
```

```
user_session_channel.sql X
C: > Users > Niranja Rao > build_test2 > models > input > user_session_channel.sql
1 SELECT
2     userId,
3     sessionId,
4     channel
5 FROM {{ source('raw_data', 'user_session_channel') }}
6
```

```
user_session_channel.sql session_timestamp.sql X
C: > Users > Niranja Rao > build_test2 > models > input > session_timestamp.sql
1 SELECT
2     sessionId,
3     ts
4 FROM {{ source('raw_data', 'session_timestamp') }}
5
```

3. Set up Output models (1pt)(Show contents of models/output/session_summary.sql)

```
(dbt_env) PS C:\Users\Niranja Rao\build_test2> dbt snapshot
02:45:55 Running with dbt=1.8.7
02:45:56 Registered adapter: snowflake=1.8.4
02:45:56 [WARNING]: Configuration paths exist in your dbt_project.yml file which do not apply to
any resources.
There are 3 unused configuration paths:
- models.build_mau.output
- models.build_mau
- models.build_mau.input
02:45:56 Found 5 models, 1 snapshot, 6 data tests, 3 sources, 459 macros
02:45:56
02:46:00 Concurrency: 1 threads (target='dev')
02:46:00
02:46:00 1 of 1 START snapshot snapshot.session_summary ..... [RUN]
02:46:04 1 of 1 OK snapshot snapshot.session_summary ..... [SUCCESS 1 in 3.51s]
02:46:04
02:46:04 Finished running 1 snapshot in 0 hours 0 minutes and 7.50 seconds (7.50s).
02:46:04 Completed successfully
02:46:04 Done. PASS=1 WARN=0 ERROR=0 SKIP=0 TOTAL=1
(dbt_env) PS C:\Users\Niranja Rao\build_test2>
```

```
user_session_channel.sql session_timestamp.sql session_summary.sql X
C: > Users > Niranja Rao > build_test2 > models > output > session_summary.sql
1 WITH u AS (
2     SELECT * FROM {{ ref("user_session_channel") }}
3 ), st AS (
4     SELECT * FROM {{ ref("session_timestamp") }}
5 )
6 SELECT u.userId, u.sessionId, u.channel, st.ts
7 FROM u
8 JOIN st ON u.sessionId = st.sessionId
9
```

4. Add Snapshot to the output table (session_summary) (2pt)(Show contents of snapshots/snapshot_session_summary.sql)

```
user_session_channel.sql session_timestamp.sql session_summary.sql snapshot_session_summary.sql X
C: > Users > Niranja Rao > build_test2 > snapshots > snapshot_session_summary.sql
1 {% snapshot snapshot_session_summary %}
2
3 {{
4     config(
5         target_schema='snapshot',
6         unique_key='sessionId',
7         strategy='timestamp',
8         updated_at='ts',
9         invalidate_hard_deletes=True
10    )
11 }}
12
13 SELECT * FROM {{ source('analytics', 'session_summary') }}
14
15 {% endsnapshot %}
16
```

5. Add at least 2 tests to sessionId field of session_summary (2pt)(Show contents of models/schema.yml, For each step, capture screenshots accordingly (examples to follow), Create a github repo (the entire dbt project needs to be uploaded) for this and add the link to the report)

```
(dbt_env) PS C:\Users\Niranja Rao\build_test2> dbt test
03:00:53 Running with dbt=1.8.7
03:00:54 Registered adapter: snowflake=1.8.4
03:00:54 Found 3 models, 1 snapshot, 2 data tests, 3 sources, 459 macros
03:00:54 Concurrency: 1 threads (target='dev')
03:00:56 1 of 2 START test not_null_session_summary_sessionId ..... [RUN]03:00:57 1 of 2 PASS not_null_session_summary_sessionId ..... [PASS in 0.99s]
03:00:57 2 of 2 START test unique_session_summary_sessionId ..... [RUN]03:00:58 2 of 2 PASS unique_session_summary_sessionId ..... [PASS in 0.93s]
03:00:58
03:00:58 Finished running 2 data tests in 0 hours 0 minutes and 4.04 seconds (4.04s).
03:00:58 Completed successfully
03:00:58 Done. PASS=2 WARN=0 ERROR=0 SKIP=0 TOTAL=2
(dbt_env) PS C:\Users\Niranja Rao\build_test2>
```

```
! dbt_project.yml  ! profiles.yml  ! schema.yml X
models > ! schema.yml
1  version: 2
2
3  models:
4    - name: session_summary
5      description: "Analytics model for session data"
6      columns:
7        - name: sessionId
8          description: "Unique identifier for each stock price record"
9          data_tests:
10            - unique
11            - not_null
12
```

<https://github.com/NiranjanRao07/dbt-data-project>