Name: Abhinita Sanabada
SJSU ID: 018320874

Q1. Create a dbt project with snowflake connector (1 pt)

Ans:

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

Enter a number: 1
account (https://<this_value>.snowflakecomputing.com): SFEDU02-LVB17920
user (dev username): CATFISH
[1] password
[2] keypair
[3] sso
Desired authentication type option (enter a number): 1
password (dev password):
role (dev role): TRAINING_ROLE
warehouse (warehouse name): CAT_QUERY_WH
database (default database that dbt will build objects in): USER_DB_CATFISH
schema (default schema that dbt will build objects in): ANALYTICS
threads (1 or more) [1]: 2
21:10:37  Profile build_mau written to /Users/tachyon_hunter/.dbt/profiles.yml using target's profile_template.yml and your s
upplied values. Run 'dbt debug' to validate the connection.
```

Terminal

```
Enter a number: 1
account (https://<this_value>.snowflakecomputing.com): SFEDU02-LVB17920
user (dev username): CATFISH
[1] password
[2] keypair
[3] sso
Desired authentication type option (enter a number): 1
password (dev password):
role (dev role): TRAINING_ROLE
warehouse (warehouse name): CAT_QUERY_WH
database (default database that dbt will build objects in): USER_DB_CATFISH
schema (default schema that dbt will build objects in): ANALYTICS
threads (1 or more) [1]: 2
21:10:37  Profile build_mau written to /Users/tachyon_hunter/.dbt/profiles.yml using target's profile_template.yml and your s
upplied values. Run 'dbt debug' to validate the connection.
Abhinita Sanabada ~ % ls -tl build_mau
total 16
-rw-r--r--@ 1 tachyon_hunter  staff  1241 Oct 30 13:58 dbt_project.yml
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 tests
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 snapshots
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 seeds
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 models
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 macros
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 analyses
-rw-r--r--@ 1 tachyon_hunter  staff   571 Oct 30 13:57 README.md
Abhinita Sanabada ~ %
```

Q2.Set up Input models (2pt)

1. Input tables should be built as CTE

```
Abhinita Sanabada build_mau % ls -tl
total 16
-rw-r--r--@ 1 tachyon_hunter  staff  1241 Oct 30 13:58 dbt_project.yml
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 tests
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 snapshots
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 seeds
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 models
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 macros
drwxr-xr-x@ 3 tachyon_hunter  staff    96 Oct 30 13:57 analyses
-rw-r--r--@ 1 tachyon_hunter  staff   571 Oct 30 13:57 README.md
Abhinita Sanabada build_mau % cd models
Abhinita Sanabada models % ls -tl
total 0
drwxr-xr-x@ 5 tachyon_hunter  staff  160 Oct 30 13:57 example
Abhinita Sanabada models % rm -rf example
```

```
dbt_project.yml  ×

Users > tachyon_hunter > build_mau >  ! dbt_project.yml
1    name: 'build_mau'
2    version: '1.0.0'
3    config-version: 2
4    profile: 'build_mau'
5
6    model-paths: ["models"]
7    snapshot-paths: ["snapshots"]    # ensure this is present
8
9    clean-targets:
10      - "target"
11      - "dbt_packages"
12
13   models:
14     build_mau:
15       +materialized: view          # default (unused by our folders below)
16       input:
17         +materialized: ephemeral  # CTEs (no objects created)
18       output:
19         +materialized: table       # physical table in ANALYTICS
20
```

2. Show contents of models/input/user_session_channel.sql and input/session_timestamp.sql

```
Abhinita Sanabada models % mkdir input output
Abhinita Sanabada models % ls -tl
total 0
drwxr-xr-x@ 2 tachyon_hunter  staff  64 Oct 30 14:44 output
drwxr-xr-x@ 2 tachyon_hunter  staff  64 Oct 30 14:44 input
Abhinita Sanabada models % cd input
Abhinita Sanabada input % ls -tl
total 16
-rw-r--r--@ 1 tachyon_hunter  staff  116 Oct 30 14:54 user_session_channel.sql
-rw-r--r--@ 1 tachyon_hunter  staff   98 Oct 30 14:54 session_timestamp.sql
Abhinita Sanabada input %
```

Inside the input files:

```
Abhinita Sanabada input % cat user_session_channel.sql
SELECT
  userId,
  sessionId,
  channel
FROM {{ source('raw', 'user_session_channel') }}
WHERE sessionId IS NOT NULL%
Abhinita Sanabada input % cat session_timestamp.sql
SELECT
  sessionId,
  ts
FROM {{ source('raw', 'session_timestamp') }}
WHERE sessionId IS NOT NULL%
Abhinita Sanabada input %
```

## Q3. Set up Output models (1pt):

1. Show contents of models/output/session_summary.sql

```
Abhinita Sanabada models % cd output
Abhinita Sanabada output % ls
session_summary.sql
Abhinita Sanabada output % cat session_summary.sql
SELECT
  u.userId,
  u.sessionId,
  u.channel,
  st.ts
FROM {{ ref('user_session_channel') }} AS u
JOIN {{ ref('session_timestamp') }} AS st
  ON u.sessionId = st.sessionId%
Abhinita Sanabada output %
```

Dbt run:

```
Abhinita Sanabada build_mau % dbt run
21:59:59  Running with dbt=1.11.0-b3
22:00:17  Registered adapter: snowflake=1.10.2
22:00:18  Unable to do partial parsing because a project config has changed
22:00:19  Found 3 models, 3 sources, 492 macros
22:00:19
22:00:19  Concurrency: 1 threads (target='dev')
22:00:19
22:00:23  1 of 1 START sql table model ANALYTICS.session_summary ........................ [RUN]
22:00:28  1 of 1 OK created sql table model ANALYTICS.session_summary ................... [SUCCESS 1 in 4.82s]
22:00:28
22:00:28  Finished running 1 table model in 0 hours 0 minutes and 9.54 seconds (9.54s).
22:00:28
22:00:28  Completed successfully
22:00:28
22:00:28  Done. PASS=1 WARN=0 ERROR=0 SKIP=0 NO-OP=0 TOTAL=1
Abhinita Sanabada build_mau %
```

## Q4. Add Snapshot to the output table (session_summary) (2pt):

1. Show contents of snapshots/snapshot_session_summary.sql

```
1   │ CREATE SCHEMA IF NOT EXISTS SNAPSHOT;
2
```

**Results (28 minutes ago)**

[ Table ]  [ Chart ]

| ⬚⬚⬚ | A status |
|-----|----------|
| 1   | Schema SNAPSHOT successfully created. |

```
Abhinita Sanabada build_mau % cd snapshots
Abhinita Sanabada snapshots % ls -tl
total 8
-rw-r--r--@ 1 tachyon_hunter  staff  259 Oct 30 15:16 snapshot_session_summary.sql
Abhinita Sanabada snapshots % cat snapshot_session_summary.sql
{% snapshot snapshot_session_summary %}
{{
  config(
    target_schema='SNAPSHOT',
    unique_key='sessionId',
    strategy='timestamp',
    updated_at='ts',
    invalidate_hard_deletes=true
  )
}}
select * from {{ ref('session_summary') }}
{% endsnapshot %}
Abhinita Sanabada snapshots %
```

Dbt snapshot run:

```
Abhinita Sanabada build_mau % dbt snapshot -s snapshot_session_summary

22:16:15  Running with dbt=1.11.0-b3
22:16:16  Registered adapter: snowflake=1.10.2
22:16:17  Found 3 models, 1 snapshot, 3 sources, 492 macros
22:16:17
22:16:17  Concurrency: 1 threads (target='dev')
22:16:17
22:16:20  1 of 1 START snapshot SNAPSHOT.snapshot_session_summary ...................... [RUN]
22:16:26  1 of 1 OK snapshotted SNAPSHOT.snapshot_session_summary ...................... [SUCCESS 1 in 5.70s]
22:16:26
22:16:26  Finished running 1 snapshot in 0 hours 0 minutes and 9.44 seconds (9.44s).
22:16:26
22:16:26  Completed successfully
22:16:26
22:16:26  Done. PASS=1 WARN=0 ERROR=0 SKIP=0 NO-OP=0 TOTAL=1
Abhinita Sanabada build_mau %
```

Checking in snowflake:

```
1      CREATE SCHEMA IF NOT EXISTS SNAPSHOT;
2
3      SELECT *
4      FROM snapshot.snapshot_session_summary
5      LIMIT 10;
```

Results (just now)

| | USERID | SESSIONID | CHANNEL | TS | DBT_SCD_ID | DBT_UPDATED |
|---|---|---|---|---|---|---|
| 1 | 184 | c41dd99a69df04044aa4e33ece9c924 | Naver | 2019-05-01 00:35:59.897 | cbf977c5099c3b5feb71ec791dbfb85 | 2019-05-01 00 |
| 2 | 184 | c41dd99a69df04044aa4e33ece9c924 | Naver | 2019-05-01 00:35:59.897 | cbf977c5099c3b5feb71ec791dbfb85 | 2019-05-01 00 |
| 3 | 184 | c41dd99a69df04044aa4e33ece9c924 | Naver | 2019-05-01 00:35:59.897 | cbf977c5099c3b5feb71ec791dbfb85 | 2019-05-01 00 |
| 4 | 80 | fdc0eb412a84fa549afe68373d9087e9 | Organic | 2019-05-01 02:53:48.503 | f1106b98f87644982e9b5e82f5454ea | 2019-05-01 02 |
| 5 | 80 | fdc0eb412a84fa549afe68373d9087e9 | Organic | 2019-05-01 02:53:48.503 | f1106b98f87644982e9b5e82f5454ea | 2019-05-01 02 |
| 6 | 80 | fdc0eb412a84fa549afe68373d9087e9 | Organic | 2019-05-01 02:53:48.503 | f1106b98f87644982e9b5e82f5454ea | 2019-05-01 02 |
| 7 | 251 | 0a54b19a13b6712dc04d1b49215423 | Facebook | 2019-05-01 12:18:26.637 | b4706917a7e8ee94cc25608c9bbfcb | 2019-05-01 12 |
| 8 | 251 | 0a54b19a13b6712dc04d1b49215423 | Facebook | 2019-05-01 12:18:26.637 | b4706917a7e8ee94cc25608c9bbfcb | 2019-05-01 12 |
| 9 | 251 | 0a54b19a13b6712dc04d1b49215423 | Facebook | 2019-05-01 12:18:26.637 | b4706917a7e8ee94cc25608c9bbfcb | 2019-05-01 12 |
| 10 | 264 | a914ecef9c12ffdb9bede64bb703d877 | Google | 2019-05-01 13:41:29.267 | 61ffaf490c97e7e01d3213f0e8a0daet | 2019-05-01 13 |

Q5. Add at least 2 tests to sessionId field of session_summary (2pt):

1. Show contents of models/schema.yml

```
Abhinita Sanabada build_mau % cd models
Abhinita Sanabada models % ls
input          output          schema.yml          sources.yaml
Abhinita Sanabada models % cat schema.yml
version: 2

models:
  - name: session_summary
    description: "Analytics model for session data"
    columns:
      - name: sessionId
        description: "Unique identifier for each session"
        tests:            # use 'tests' (works across dbt versions)
          - not_null
          - unique
Abhinita Sanabada models %
```

Running test for snapshot:

```
Abhinita Sanabada models % dbt snapshot

22:40:52  Running with dbt=1.11.0-b3
22:40:53  Registered adapter: snowflake=1.10.2
22:40:54  Found 3 models, 1 snapshot, 2 data tests, 3 sources, 492 macros
22:40:54
22:40:54  Concurrency: 1 threads (target='dev')
22:40:54
22:40:57  1 of 1 START snapshot SNAPSHOT.snapshot_session_summary ........................ [RUN]
22:41:03  1 of 1 OK snapshotted SNAPSHOT.snapshot_session_summary ........................ [SUCCESS 0 in 5.59s]
22:41:03
22:41:03  Finished running 1 snapshot in 0 hours 0 minutes and 9.39 seconds (9.39s).
22:41:03
22:41:03  Completed successfully
22:41:03
22:41:03  Done. PASS=1 WARN=0 ERROR=0 SKIP=0 NO-OP=0 TOTAL=1
Abhinita Sanabada models %
```

**git hub link:**

https://github.com/abhinitasanabada-web/Homework/tree/8d0d9015cad7124ff2537
3e3c75da7d5a3aa11c3/Homework%237