Nadia Hassan Junior Data Engineer

San Francisco, CA · nadia.hassan@gmail.com · +1 415-555-0124 · linkedin.com/in/nadiahassan

PROFILE SUMMARY

- **Junior Data Engineer** with **2 years** building **ELT pipelines and dbt models** at a **BI software company**, after **3 years of Looker reporting** as a BI analyst at a real estate brokerage.
- Uses **Python** and **SQL** daily, builds models in **dbt** on **Snowflake**, loads sources with **Fivetran**, schedules jobs in **Apache Airflow**, and ships changes through **GitHub Actions** and **Terraform** with senior review.
- Learning **dimensional modeling**, **incremental loads**, and **Kafka** from senior engineers, and adds **dbt tests** and **row-count checks** to every model before merge, so **bad numbers stay off revenue dashboards**.
- Sits on a **4-person data team** serving **finance, sales ops, and product analysts**, joins **weekly intake** to scope new requests, and takes a monthly turn on **Airflow on-call**.
- Wrote the team's **Airflow failure runbook**, coaches **3 analysts** through their first dbt models in pairing sessions, and is booked for the **SnowPro Core** exam in January 2027.

TECHNICAL SKILLS

Languages:	Python, SQL, Bash
Processing & Transformation:	dbt Core, pandas, Polars, LookML, PySpark (basics)
Streaming & Messaging:	Apache Kafka, Kafka Connect, Snowpipe Streaming
Orchestration & Ingestion:	Apache Airflow, Fivetran, Segment
Warehouses & Lakehouses:	Snowflake, PostgreSQL, Apache Parquet
Cloud & Infrastructure:	AWS (S3, IAM, Lambda), Terraform, Docker, GitHub Actions
Data Quality & Observability:	dbt tests, Great Expectations, Elementary, SQLFluff, Datadog

EDUCATION

DataTalks.Club — Data Engineering Zoomcamp (certificate) — Online · Jan 2024 to Apr 2024

University of California, Davis — B.S. in Statistics — Davis, CA · Sep 2017 to Jun 2021

WORK EXPERIENCE

Sigma Computing — Junior Data Engineer — San Francisco, CA · Oct 2024 to Present

- Built Python loaders for the internal billing and licensing APIs next to **Fivetran** syncs for Salesforce and Zendesk, with incremental pulls and duplicate-safe reruns, bringing **14 sources** into **Snowflake**.
- Split the nightly **Airflow** refresh into per-source DAGs with sensors, retries, and Slack alerts, so stale sales dashboards at 8am dropped from 6 mornings a month to **1**.
- Paired with a senior engineer to move product usage events from hourly S3 loads to **Kafka Connect** and **Snowpipe Streaming**, cutting table lag from 70 minutes to **4**.
- Rewrote the slowest **dbt** models in the product usage project as incremental merges clustered on event date, which cut their nightly **Snowflake** credit use by **43%**.
- Set up **GitHub Actions** jobs that lint SQL with SQLFluff and build only changed dbt models in a dev schema, catching **31 broken models** before merge in the first quarter.

Compass — BI Analyst — San Francisco, CA · Jul 2021 to Sep 2024

- Modeled closed deals, agents, and offices as a **dbt** star schema in **Snowflake** for commission reporting, cutting the monthly agent payout close from 3 days to **6 hours**.
- Added **dbt tests** for duplicates and nulls on the listings models, plus row-count checks against the MLS feeds, which caught **38 bad loads** before agents saw them in Looker.
- Met biweekly with brokerage ops managers to turn report requests into tickets with agreed metric definitions, and documented each field in **LookML**, cutting duplicate dashboard requests by **41%**.