

Sneha Ekka

Boston, MA | (857) 333-5810 | sekka@bu.edu | [linkedin.com/in/snehaekka/](https://www.linkedin.com/in/snehaekka/) | github.com/SnehaEkka | www.snehaekka.com/

PROFESSIONAL SUMMARY

Analytics Engineer with 4 years of experience across startup and enterprise environments, owning ELT pipelines, analytics-ready datasets, and metrics that power business decision-making. Partnered with stakeholders to translate business requirements through data modeling and self-serve analytics, with expertise in SQL, Snowflake, dbt, Python, automation, and end-to-end data pipelines.

SKILLS SUMMARY

Languages & Frameworks: SQL, Python, PySpark, Spark SQL, Excel

Data Platforms & Cloud: Snowflake, BigQuery, Databricks, Teradata, AWS, GCP, MotherDuck

Modeling & Orchestration: dbt, Airflow, Prefect, CI/CD, Git/GitHub

BI & Analytics Enablement: Tableau, Looker, Power BI, Sigma, Streamlit

Core Strengths: ELT/ETL pipelines, modular data modeling, analytics-ready datasets, metric definitions, query optimization, data validation, data quality automation, self-service analytics, stakeholder management, cross-functional collaboration

ACADEMIC EXPERIENCE

M.S. in Business Analytics (STEM)

Boston University, Questrom School of Business | Boston, MA

Sep 2023 – Jan 2025

B.S., M.S. in Statistics

University of Mumbai | Mumbai, India

Jun 2015 – May 2020

WORK EXPERIENCE

Senior Business Analyst

IDORI (Startup) | Boston, MA

Mar 2025 – Present

- Engineered and deployed ELT pipelines on GCP for marketing and web data (Meta Ads, Google Merchant), enabling ROI tracking and performance visibility to optimize CAC, targeting reduction in ad spend per converted customer by 37%.
- Led modeling of customer profile datasets using behavioral and demographic signals to support GTM strategy, improving conversion rate from 1% to 6% through targeted D2C outreach.
- Owned lead tracking workflows in HubSpot CRM for 1000+ leads, implementing segmentation and controlled email A/B tests to surface engagement drivers and inform B2B outreach strategy.

Analytics Consultant

BOARD | Boston, MA

Jun 2024 – Dec 2024

- Designed analytics-ready datasets via API-based ingestion and ETL for a healthcare nonprofit, building data models for complex budgeting workflows, reducing manual data prep by 40% and enabling self-service planning.
- Translated FP&A requirements into data models, governed KPIs and metrics, variance analyses, and budgeting workflows with role-based access, delivering a self-serve planning application used by finance teams across 3 enterprise clients.
- Led regression testing and validation during a platform upgrade (Board 12.6 to 14.1), identifying 10+ discrepancies and establishing reusable data validation practices across teams.

Analytics Engineer

TERADATA | Mumbai, India

Oct 2020 – Jun 2023

- Designed and enhanced scalable ETL frameworks using Python scripts to integrate 100M+ healthcare, behavioral, and first-party datasets, supporting enterprise analytics initiatives during the pandemic.
- Led migration of production ELT pipelines processing 1B+ records from Teradata to Snowflake for a global media client, ensuring reporting continuity and improving pipeline reliability.
- Optimized 200+ SQL transformations using modular, dbt-style patterns orchestrated with Airflow and Python, improving query performance by 40% and ensuring pipeline observability.
- Automated cross-platform data validation in Python, reducing manual QA from 4 hours to 5 minutes (98% efficiency gain), speeding data quality checks and UAT while improving downstream data trust.
- Improved overall data processing time by 44% (90 to 55 minutes) through SQL tuning and cloud storage optimization in AWS, strengthening SLAs for executive reporting.
- Created and managed 5+ Tableau dashboards tracking 15+ KPIs, proactively resolving data inconsistencies and pipeline issues to support ongoing stakeholder decisions.

PROJECT EXPERIENCE

Netflix Recommendation Pipeline – Shipped an end-to-end ELT pipeline on GCP, applying dbt-style modeling to Netflix API data, orchestrating workflows with Prefect, and deploying a Streamlit UI for personalized recommendations powered by KNN and LLM.

U.S. Commercial Aviation Analytics – Analyzed 5.8M+ flight records using BigQuery and PySpark, translating business questions into optimized SQL models and Tableau dashboards to surface delay and cancellation drivers.