

TRIET LE

ben.le.data@gmail.com • (+84)-825-717-449

• LinkedIn: [://trietbenle](https://www.linkedin.com/in/trietbenle) • Ho Chi Minh City, Vietnam

SKILLS

- Language: Python, Scala, JavaScript, SQL, Bash, HTML, CSS
- Technology: BigQuery, DuckDB, dbt, Pandas, Airflow, Spark, GCP, Docker, Ansible, Vagrant, Terraform
- Skills: Cloud, ETL, Data Warehouse, Business Intelligence, Data Modeling, RDBMS, Machine Learning

EXPERIENCE

Holistics

Ho Chi Minh, Vietnam

Data Engineer

Aug 2022 - Present

- Redesign a legacy **Data Warehouse** architecture with various modeling paradigms(Dimensional Modeling, Event Schema) using **dbt**, enabling scalable support for advanced business analytics use cases.
- Develop and orchestrate **~100 ETL data pipelines** with **Prefect** from **~20** data sources to a central Data Warehouse, providing robust analytics insights and capabilities.
- Implement a comprehensive **Data Catalog** in **dbdocs** with over **100** tables to streamline data accessibility, resulting in a **40%** increase in weekly active SQL Data Warehouse users.
- Design and implement a data usage monitoring system with **Holistics** to pinpoint expensive query run on production. The system ultimately reduces redundant **BigQuery** usage by **70%** compared to its peak.
- Setup **Snowplow event tracking** to measure product analytics, offering visibility into key success metrics and help Product team make data-informed decisions to accelerate iterative development.
- Build a **data observability system** in **Grafana** for **~1500** data products (dashboards, data models and data pipelines). This system has enabled prompt responses to data pipeline failures with **60%** decrease in failure rates and reducing the mean time to recovery
- Collaborate with the DevOps team to automate **GitHub Actions CI/CD** pipelines for internal data projects, reducing manual deployments and enabling faster, error-free releases.

Technical Writer

Aug 2023 - Jan 2024

- Managed a team of **3** technical writers and data analysts, producing **20+** best practice articles and tutorials. These articles enhanced customers understand the product's capabilities which increases product adoptions.
- Researched and documented BI use cases to enhance user understanding and adoption of analytics solutions.

Troy University

Troy, Alabama

Computer Science Tutor and Lab Assistant

Jul 2021 - May 2022

- Tutored college students in Python, C++, Java with more than 200+ hours of coding lessons
- Assisted instructor in the daily assessment of student progress and inspected student assignment

PROJECTS

Data Clinic Bot

Mar 2023

- Developed a RAG Slack ChatBot using Langchain with GPT 4.0 and Pinecone Vector Database
- The chatbot answers high-level data inquiries and suggest relevant Holistics reports

Anime Recommendation

Mar 2022

- Developed a web-based anime recommendation tool using Flask, CSS, and HTML and cosine similarity algorithm
- Performed data cleaning and feature engineering using Scipy, Numpy, Pandas, and Sklearn

EDUCATION

Troy University, Troy, AL

Graduation Date: May 2022

B.S Computer Science

GPA: 3.5/4.0

CERTIFICATES

- [Data Engineer with Python \(DataCamp\)](#)
- [DataTalks Data Engineer Zoomcamp](#)
- [Big Data with Spark and Hadoop \(IBM\)](#)
- [Python for Data Science and Machine Learning Bootcamp \(Udemy\)](#)
- [SQL Advanced \(HackerRank\)](#)