TRAN QUANG VINH

DATA ENGINEER

I hold a Software Engineering degree from BTEC FPT College and specialize in Data Engineering. With expertise in SQL, Python, Spark, and cloud platforms (Azure, AWS), I design scalable data pipelines that transform complex data into actionable insights, driving business growth and efficiency.

CAREER OBJECTIVE

As an aspiring Data Engineer with strong foundations in cloud and data systems, I aim to help build scalable pipelines, optimize workflows, and deliver business insights. I am passionate about solving real-world data challenges and continuously growing through technologies like Spark, Kafka, Azure, and AWS.

CONTACT

- 235371vinhtran@gmail.com
- +84 961321461
- Binh Thanh District, HCMC.

EDUCATION

BTEC FPT College
 Degree: Software Engineer
 Awarded: August 2024

TOEIC: 750

SKILLS

- Python Language, SQL Query
- Apache Spark (DataBricks in Cloud) for Transformation
- Cloud Knowledge: AWS, Azure
- Extra: Hadoop, Kafka, Docker

EXPERIENCE

STOCK MARKET ELT PIPELINE PROJECT
(**Aug – Sep 2024)**

Source: GitHub, Document
Description:
- Designed and implemented a full-scale ELT data pipeline to integrate financial and stock market data from Sec-api.io, Alpha Vantage, and Polygon into a centralized data warehouse.
- Analyzed API structures, defined ingestion strategies, and evaluated data latency and transformation requirements.
- Designed relational and dimensional data models (Galaxy Schema) to support analytical reporting.
- Built modular Airflow DAGs for orchestration, automated ingestion, transformation, and loading routines.
- Created technical documentation including data lineage, architecture diagrams, and process flows.

KAFKA SPARK STREAMING MODELING

Source: GitHub, Document
Description:
- Built a real-time streaming pipeline using Kafka, Spark, MinIO, and PostgreSQL for sentiment data analytics.
- Designed real-time data schema and Spark validation layers to ensure data quality and consistency.
- Developed Dockerized architecture integrating Airflow for orchestration and automated stream processing.
- Implemented monitoring logic with Prometheus & Grafana for system observability.
- Authored comprehensive technical documentation and enabled REST API access via FastAPI.