

Hiring software engineer as easy as calling a taxi.

Vadym U.

Data Engineer

SUMMARY

- Data Engineer with solid data pipelines, DWH, data lake architecture, development and optimization expertise on cloud platforms including Azure, GCP, and AWS.
- Snowflake strong - advanced level
- with a proven track record of automating ETL processes with multiple tools managing large-scale data warehousing, and enabling business intelligence through sophisticated analytics solutions.
- Strong Python, Spark, Kafka, skills,
- Experience creating datastore and DB architectures, ETL routines, data management and performance optimization
- MSSQL, MySQL, Postgres.
- In multiple projects, Vadym analysed and Improved operational efficiency, reduced data-related infrastructure costs, and delivered seamless data integration and transformation across systems.

SKILLS

Main Technical Skills	Python, Kafka, Apache Spark, Snowflake
Data Analysis and Visualization Technologies	Databricks, Tableau
Databases & Management Systems / ORM	Data Warehousing, dbt, Firebase Realtime Database, Microsoft SQL Server, MySQL, PostgreSQL
Cloud Platforms, Services & Computing	AWS, Azure, GCP (Google Cloud Platform)
Amazon Web Services	AWS S3
Azure Cloud Services	Azure, Databricks
Google Cloud Platform	GCP BigQuery
Web/App Servers, Middleware	Apache HTTP Server
Virtualization, Containers and Orchestration	Docker

Scripting and Command Line Interfaces	Shell Scripts
Other Technical Skills	Apache superset, Data Validation, Snowflake Data Warehouse, Talend ETL, Trino

WORK EXPERIENCE

Data Engineer, Cloud Data Pipelines and Analytics

Duration: 2022 June – Current

Summary: Engineered solutions for data streaming from Kafka to Spark and managed data workflows within cloud platforms such as Google Cloud and Azure.

Responsibilities:

- Managed and optimized Snowflake data warehousing and orchestrated data workflows into Amazon S3 using the Data Build Tool.
- Developed data pipelines on Google Cloud and Azure, ensuring seamless data ingestion and transformation.
- Engineered solutions for data streaming from Kafka to Spark, processed and stored the data in Delta tables, registered them in Trino for efficient querying and integrated them with Apache Superset for BI analytics and visualization.

Achievements:

- Improved data processing efficiency through optimizing Spark jobs and Snowflake configurations.
- Automated manual data orchestration tasks using Data Build Tool, reducing operational overhead.
- Successfully migrated key data pipelines to cloud platforms, reducing operational costs.

Technologies: Python, Kafka, Spark, Snowflake, GCP, Azure, AWS, Docker, DBT, Trino, Apache Superset

Data Engineer, Retail Data Analytics Project

Duration: 2021 January – 2022 May

Summary: Developed and implemented efficient data pipelines for a retail project using various Microsoft Azure services.

Responsibilities:

- Managed and optimized MS SQL Server databases, including modifying stored procedures and scheduled jobs.
- Supported data analytics by providing custom queries and comprehensive datasets for business insights.



- Developed ETL processes to integrate data from multiple sources, ensuring accuracy and consistency.
- Conducted performance tuning and query optimization for improved database performance.
- Automated routine tasks with SQL scripts, reducing manual effort and errors.
- Implemented data validation checks to maintain high data quality and integrity.

Technologies: Python, Kafka, Spark, Azure, MS SQL Server

EDUCATION

Bachelor's degree in political science

Studied at Pereyaslav-Khmelnytskyi State Pedagogical University

2010 - 2014

