

Seongju, Hwang (Alex)

Phone: +82-10-9742-2257

Email: zztisdudoo@gmail.com

LinkedIn: <https://www.linkedin.com/in/alex-hwang1998/>

GitHub: <https://github.com/svng-zu>

Summary

- Data Engineer with hands-on experience building and operating large-scale analytical data platforms in insurance and financial services. Specialized in Spark-based batch and streaming pipelines, workflow orchestration, and BI enablement. Strong interest in self-service BI system design, focusing on reliable, observable, and scalable data infrastructure.

Work Experience

(Jul 2024 – Present) **Data Engineer Hanwha General Insurance**, Seoul

- Operate and optimize Azure HDInsight-based Spark/Hive platforms with YARN-level resource tuning
- Designed and maintained **300+ Oozie workflows**, processing **~1B rows/day**, with robust retry and failure isolation
- Built Kafka-based real-time pipelines ingesting **~100K events/sec** from **~700K IoT** devices using Spark Structured Streaming
- Migrated Sqoop-based CDC pipelines to Spark JDBC ingestion, reducing latency from **~7h** → **~3h**
- Standardized batch/stream outputs into managed Hive tables for reliable BI and analytics consumption
- Built Power BI infrastructure end-to-end, including on-prem gateway and security coordination
- Led PoCs for **Apache Superset** and **Databricks**, evaluating performance and operational feasibility

(Apr 2024 – Jul 2024) **DW/BI Intern | Carrot Insurance (Hanwha General Insurance)**, Seoul

- Developed Hive SQL-based data marts aligned with business KPIs
- Migrated Tez-based workflows to Spark, reducing batch runtime **~2h** → **~30m**
- Built and maintained Power BI dashboards and refresh pipelines

Education

(2023) **Seoul National University of Science and Technology**, Seoul, Korea

Bachelor of Science in Industrial & Information Systems Engineering

Projects & Awards

(2023) **Cheonan City Data Visualization Idea Contest - 2nd Awards**

- Analyzed public transportation and mobility datasets to identify traffic accident risk patterns affecting elderly pedestrians.
- Derived key indicators such as high-frequency accident zones and senior pedestrian movement flows.
- Developed an interactive Power BI dashboard visualizing movement paths and facility accessibility.

Presented data-driven policy recommendations to municipal stakeholders, receiving high recognition for social impact.

(Jun 2023 – Dec 2023) **LG HelloVision DX Data School**

- Completed an intensive program covering Docker-based containerization and AWS cloud architectures.
- Gained hands-on experience with IAM, EC2, ECS, CloudFront, and CI/CD-oriented system design principles.

Technical Skills

Languages

- Python, SQL

Data Engineering & Platforms

- Apache Spark, Hive, Hadoop, Airflow, Oozie, Kafka, Docker

Cloud & Infrastructure

- Azure, AWS, Docker, Linux, Databricks

Databases

- MySQL, Tiberio, Azure Cosmos DB, Postgres SQL

BI & Analytics

- Power BI, Apache Superset, Tableau