

Senior Data Engineer

Sunghwan Ki (Danny)

Phone : 205-734-9654
E-Mail : kish1919@gmail.com
Portfolio : <https://kish1919.github.io/showcase/>
Linked-In : [linkedin.com/in/danny-ki/](https://www.linkedin.com/in/danny-ki/)
Git Hub : github.com/kish1919
Address : Fairfax, VA

Professional Summary

Experienced Data Engineer with over 6+ years of expertise in data pipeline automation, ETL development, and data warehousing. Adept at designing scalable data architectures and optimizing data pipelines for high-performance analytics. Skilled in Kafka streaming, Oracle Exadata, Hadoop ecosystem, Python, SQL, and cloud platforms (AWS). Proven ability to drive data-driven decision-making, enhance data integrity, and collaborate with cross-functional teams in fast-paced enterprise environments

Experience

PNC Bank | Senior Big Data Engineer

Pittsburgh, PA Jul. 2019 - Present

- Designed and implemented Kafka streaming data pipelines, ingesting real-time data from multiple sources into Oracle Exadata and subsequently into Hadoop platforms.
- Built data ingestion frameworks using PySpark and shell script, reducing manual intervention and improving efficiency.
- Automated ETL workflows for real-time data processing and batch processing, optimizing performance and reducing latency.
- Monitored Kafka real-time data streams, ensuring data integrity and troubleshooting issues to minimize downtime.
- Managed and optimized Oracle Exadata performance, ensuring efficient storage and retrieval of large-scale datasets. Implemented data partitioning and indexing strategies to improve query performance.
- Led the implementation of data governance and quality assurance processes, ensuring high data integrity.
- Collaborated with business analysts and data scientists to enhance data quality, governance, and analytics frameworks.

Hyundai Powertech America | Data Analyst

West Point, GA Jul. 2018 - Jul. 2019

- Designed ETL processes for integrating production data into analytics platforms, enabling better decision-making.
- Developed a machine learning model to optimize manufacturing processes, improving efficiency and reducing costs by 6%.
- Conducted data regression analysis, correlating irregular part usage with production conditions, leading to enhanced operational insights.
- Analyzed machine recall data to identify failure patterns, improving predictive maintenance strategies.

Donwon Autopart Technology Georgia LLC | Sales and Analytics Manager

Hogansville, GA Jul. 2015 - Dec. 2017

- Analyzed vast amounts of pricing-related data, improving pricing strategies and revenue forecasting.
- Led the adoption of 2D barcode logistics management, reducing mislabeled shipments and late deliveries by 94%.

Education

University of Wisconsin-La Crosse | Master of Science - Data Science

Jan. 2019 – May. 2024

- GPA: 3.58 / 4.0
- Relevant Coursework: Data Warehousing, Machine Learning, Data Mining, Prescriptive Analytics

Myoungji University | Bachelor of Business Administration

Mar. 2000 - Feb. 2009

- GPA: 3.6 / 4.5

Key Skills

- Programming Language: Python, PySpark, SQL, Linux
- Big Data Platform: AWS, Spark, Hadoop, Redshift
- Application: Jenkins, Bitbucket, Control-M, CA7
- Database: AWS RDS, MySQL, Oracle
- Analysis : AWS EMR, Hive, Impala
- Toolkits: Git, IntelliJ, JIRA

Certificates

- AWS Certified Cloud Practitioner (Issue date : Aug. 2024)
- AWS Certified Developer – Associate (Issue date : Jul. 2021)

Additional Training

- Udemy : Snowflake - The complete masterclass (Nikolai Schuler / Dec.2021)
- Udemy : Data Analyst using Sqoop, Hive and Impala (Mohamed F./ Nov.2021)
- Coursera : Data Structures (*Alexander S. Kulikov* / Jul. 2018)
- Coursera : Java Programming (*Susan H Rodgers* / Jul. 2018)
- Coursera : Probability and data (*Mine* / Aug. 2017)
- Coursera : Machine learning (Andrew Ng / Jul. 2017)
- Coursera : Machine learning foundation (*Carlos, Emily* / Jul. 2017)
- Coursera : Python programming (*Charles Severance* / Jul. 2017)