Dat Nguyen (Dylan)

tiendatnguyen.kuas@gmail.com - (+358) 46 952 7847 - Helsinki, Finland
Portfolio: tiendatscorpy.github.io

SUMMARY

AWS-Certified Solution Architect and Data Analytics - Specialty. Team-oriented data engineer proven in critical projects. Patient problem-solver with ability to leverage cloud solutions and an appreciate for clean code, adept at data gathering, ETL pipelines, and databases.

TECHNICAL SKILLS

Python • PySpark • SQL • Docker • Flask • S3 • Git • Hadoop • Airflow • EMR • Lambda • DynamoDB
API Gateway • R • 4D Software • Kinesis • Unit Testing • Redshift • Glue • Athena • Splunk

WORK EXPERIENCE

Basware   **May 2021 – Present**
Data Engineer   *Helsinki, FI*
- Design, build and launch an AWS streaming pipeline that uses Glue ETL to consume messages from Kinesis data stream, apply business mappings to the data-frame and persist them into parquet files.
- Enhance visibility, monitoring and alerting throughout the stack with Splunk and Cloudwatch
  Tech Stack: AWS solutions, including Lambda, API Gateway, EMR, Redshift, Glue, Athena, DynamoDB, S3, Tableau; Splunk.

Trafore Oy   **May 2018 – May 2021**
Backend and Data Engineer   *Helsinki, FI*
- Develop and maintain a RDBMS for a Helsinki taxi management software, which is responsible for shifts and driver data for tens of taxi companies in Helsinki, Espoo, Rovaniemi, etc.
- Design and implement on-premise Airflow for optimal extraction, transformation, and loading (ETL) of data from a wide variety of data sources both internally and externally
- Implement an end-to-end data pipeline to sync salaries into Income Register (Tulorekisteri)
  Tech Stack: Python, 4D, PostgreSQL as backend engines | Apache Airflow as ETL pipeline orchestrator

PROFESSIONAL CERTIFICATES

- AWS Data Analytics – Specialty DAS-C01
- AWS Solution Architect – Associate SAA-C02

PROJECTS

Optimizing Public Transportation with Kafka Streaming
Stream public transit status using Kafka and the Kafka ecosystem to build a stream processing application that shows the status of trains in real-time.
Tech Stack: REST Proxy, Kafka Connect, KSQL, Apache Avro, Faust Python Stream Processing

Evaluate Human Balance with Spark Streaming
Produce a stream of data to a Kafka topic and made it available to the STEDI application to consume. This data is used in a new feature for the application which consist of a graph that shows fall risk (will they fall and become injured?) for recent assessments
Tech Stack: Spark Streaming, Redis, Base64, JSON

EDUCATION

Aalto University   **August 2021**
Master of Security and Cloud Computing   *Helsinki, FI*
Aalto University   **August 2018 – June 2021**
Bachelor of Data Science, GPA 4.43   *Helsinki, FI*