Muhammad Arkan Nibrastama

Bandar Lampung, Indonesia | m.arkannibrastama64@gmail.com | +62 821 8456 7507
LinkedIn | Portfolio

SUMMARY

Curious and communicative with 2 years of experience in data engineering solutions. Driven by the challenge of transforming complex data into actionable business solutions. Automating everything to empower human insight. I am seeking a Data Engineer role where I can apply my skills in building scalable data pipelines and robust data governance frameworks. My goal is to build high-impact data solutions while evolving my skills to meet new challenges as the company grows.

EXPERIENCE

Ditjen Diktiristek   **Central Jakarta, Indonesia**
Data Analyst • Internship   *May, 2024 - July, 2024*

- Boosted operational efficiency by 50% by delivering 8 actionable recommendations derived from analyzing over 15,000 user data duplications in BigQuery.
- Developed a comprehensive monitoring dashboard using Looker Studio to provide supervisors with ongoing oversight of data quality and integrity.

Database Engineer • Internship   *February, 2024 - April, 2024*

- Reduced HR reporting time by 50% by architecting and building a centralized PostgreSQL data warehouse.
- Built a scalable ETL pipeline to process and integrate over 300,000 employee data records and ensure high data accuracy for reporting.
- Enhance data security by implementing data governance principles to protect personally identifiable information (PII) and ensure regulatory compliance.

Data Engineering with Google Cloud Platform Book   **Singapore (Remote)**
Technical Reviewer & Content Contributor • Contract   *July, 2023 - September, 2023*

- Provided in-depth technical feedback across 12 chapters to enhance clarity, accuracy, and practical relevance for entry-level data engineers.
- Updated and validated 8 project-based exercises to reflect the industry best practices, latest GCP features and updates across the Google Cloud Platform (GCP) ecosystem.
- Validated and tested code solutions using 10+ core GCP services, including BigQuery, Dataflow, Dataproc, Cloud Composer, Pub/Sub, Cloud Functions, and Looker Studio.
- Ensured all technical content was accurate, up-to-date, and aligned with current industry standards for GCP-based data pipelines, ETL, orchestration, and analytics workflows.

Bank SMBC Indonesia   **Central Jakarta, Indonesia (Remote)**
Data Engineer • Internship   *December, 2022 - January, 2023*

- Helped reduce customer churn by 30% by transforming and analyzing over 2,000 customer data records to identify key churn drivers using T-SQL.
- Delivered actionable insights to business stakeholders through compelling data visualizations using PowerBI and directly influencing retention strategies.
- Provided data-driven solutions to 15 unique business problems to enhance the organization's decision-making capabilities

.

PROJECT

Automotive Manufacture Data Pipeline ([Show Project](#))
Reduced potential machine downtime by building a real-time streaming pipeline with Flink, ClickHouse, and Superset, while enhancing sales forecasting through a unified Data Lakehouse on MinIO, Spark, Iceberg, Nessie, and Trino.

People Performance Data Pipeline ([Show Project](#))
Optimized HR analytics by reducing data processing time by 75% through a centralized data in Delta lake with ETL pipeline that validated over 10,000 records using PySpark and Great Expectations in Azure Databricks.

E-Commerce Data Pipeline ([Show Project](#))
Reduced manual data processing by 80% for Shopify owners by building an automated ETL pipeline using PySpark, Airflow, and Docker including data validation using PyTest into a BigQuery data warehouse.

CDC and SCD Approach to Enhancing Expedition Operations ([Show Project](#))
Improved delivery process efficiency by implementing a real-time data tracking system with CDC (Capture Data Changes) and SCD (Slowly Changing Dimension) Type 2 methodologies using Debezium, Kafka, and BigQuery.

Spotify Data Pipeline ([Show Project](#))
Enhanced data analyst productivity by fully automating the ingestion of over 300 Spotify data records through an end-to-end pipeline, managed as code with Terraform and orchestrated using Google Cloud Composer and Dataflow.

EDUCATION

Computer Science   **2021 - 2025**
Universitas Lampung, Indonesia
GPA : 3.76
- **Assistant Lecturer for Database Course** (August, 2023 **-** December, 2023)
- **Related Course**: Database, Distribution Data Processing, Cloud Computing, Big Data and Data Warehouse.

SKILL

Communication, Collaboration, Problem-solving, Attention to detail, ETL (Extract, Transform, Load), Data Analysis, CI/CD, Slowly Changing Dimension (SCD), Capture Data Change (CDC), API, Python, SQL, Java, C++, DBMS (MySQL, PostgreSQL , Microsoft SQL Server, Clickhouse), Trino, Data Visualization (Power BI, Looker Studio, Apache Superset), Apache Airflow, Apache Kafka, Apache Beam, Apache Spark, PySpark, Apache Flink, Apache Iceberg, Project Nessie, Amazon Web Service (S3, EC2, Glue), Google Cloud Platform (Google Cloud Storage, Google Cloud Function, Google Cloud Run, Google Compute Engine, Dataproc, Dataflow, Cloud Composer, BigQuery), Microsoft Azure (Databricks, Azure Key Vault, Delta Lake), Terraform, Debezium, Docker.