

# Ida Bagus Gede Purwa Manik Adiputra

Data Engineer | ETL Pipelines, Data Analytics, Cloud Platforms

☎ 081287430970 @ ida.adiputra@outlook.com 🔗 [Linkedin](#) 📁 [Portfolio](#) 📍 Kabupaten Tangerang, Indonesia



## SUMMARY

Results oriented Computer Science graduate with a strong foundation in data engineering and business intelligence. Skilled in building and automating modern data infrastructure using GCP, AWS, Azure, Terraform, and Docker. Experienced in designing scalable real-time and batch data pipelines with Apache Spark, Apache Kafka, dbt, and Airflow. Demonstrated ability to drive operational efficiency achieving a 20% improvement at PT Bank Central Asia through internal process optimization. Proficient in transforming data into actionable insights using tools like Looker Studio and Tableau across various analytics projects.

## TECHNICAL SKILLS

- **Cloud & Data Platforms:** Azure Synapse, Azure Data Factory, Azure Databricks, GCP (BigQuery, GCS, Dataproc, Composer), AWS (EKS, S3, Redshift, MWAA, ECR)
- **Data Engineering & Warehousing:** ETL / ELT Pipelines, Apache Spark, Apache Airflow, PostgreSQL, Snowflake, Apache Kafka, dbt
- **Programming & Analytics:** Python, Java, SQL, PySpark, Statistical Analysis, Data Modeling
- **Visualization & BI Tools:** Power BI, Tableau, Looker Studio
- **DevOps & Infrastructure:** Terraform, Docker, Git

## PROFESSIONAL EXPERIENCE

### Business Analyst Intern

#### PT Bank Central Asia Tbk

📅 02/2024 - 02/2025

- Analyzed business requirements through comprehensive stakeholder interviews to optimize operational processes and enhance decision making capabilities across remittance and banknotes divisions
- Designed and implemented solutions for internal website optimization that delivered measurable operational efficiency improvements of 20%
- Collaborated with cross functional teams including IT, operations, and business units to develop and align strategic business solutions with technical implementations

## TECHNICAL PROJECTS

### End to End E-Commerce Data Platform on Google Cloud

- Architected and deployed a scalable, end to end ELT data platform on GCP using Terraform for infrastructure as code, ensuring a fully automated and reproducible environment
- Orchestrated the entire data workflow using Cloud Composer (Airflow), automating the ingestion of raw data from GCS, transformation with dbt, and loading into a BigQuery data warehouse
- Implemented a modern data modeling workflow using dbt, building modular staging and mart tables with integrated data quality tests to ensure the integrity and reliability of the final analytics ready data
- Developed and delivered an interactive Looker Studio dashboard to provide key business insights into profitability, sales trends, and top performing products, completing the full data lifecycle

### Real Time Football Data Engineering Pipeline

- Deployed and managed all cloud resources including GCE for Kafka, a Dataproc cluster, and BigQuery datasets using Terraform, ensuring a fully reproducible and automated environment
- Developed a containerized Python producer (Docker) to stream live API data into Apache Kafka, which was then consumed and transformed in real time by a distributed PySpark streaming application
- Utilized Apache Airflow (Cloud Composer) to schedule, execute, and monitor the Spark processing job, creating a resilient and automated data workflow
- Loaded the processed, structured data into BigQuery for scalable analytics and delivered a live updating dashboard in Looker Studio to visualize key in game metrics

### Champions League Data Platform on AWS EKS

- Architected and deployed a production grade data platform on AWS, using Terraform to provision a secure VPC, EKS cluster, S3 data lake, MWAA environment, and Redshift data warehouse
- Engineered a complete CI/CD pipeline with GitHub Actions to automatically build, scan, and deploy four containerized data microservices to Kubernetes using Kustomize for configuration management
- Orchestrated the end to end ETL workflow with Apache Airflow (MWAA), dynamically running containerized PySpark jobs on EKS to process data through a multi-layered S3 data lake (Bronze, Silver, Gold)

## EDUCATION

### Bachelor of Computer Science

#### Bina Nusantara University

📅 09/2021 - Present

GPA

3.62 / 4.00

## CERTIFICATIONS

Apache Airflow: The Hands-On Guide (Udemy)

Associate Data Engineer in Snowflake (DataCamp)

Azure Databricks & Spark for Data Engineers: Hands-on Project (Udemy)

GitHub Foundations (GitHub)

## LANGUAGES

English

Proficient

Indonesia

Native