Arnold Huete

Backend Software Engineer at X-team | AWS | Azure | 6 x GCP | English | Spanish

E-mail: ahueteg@uni.pe Linkedin: https://www.linkedin.com/in/arnoldhuete/
Web page: https://ahueteg.com/ Github: https://github.com/ArnoldHueteG
GCP Credentials: https://www.credential.net/profile/arnoldhuete961102/wallet
Last ML Competition
Av Cesar Vallejo 257, Dpto 1304 – Lince, Lima Peru - Phone: +51942603349

Summary

Helping companies get the most out of cloud services. Beginning my career as a Business Analyst and simultaneously pursuing software development as a side hustle, I honed my skills in understanding both business needs and technical solutions. Over the years, I smoothly transitioned into roles such as Data Engineer and Solutions Architect for Data, demonstrating expertise in designing robust data solutions. Currently, as a dedicated backend engineer, I am eagerly poised to expand my skill set and evolve into a well-rounded full-stack developer.

Experience

Backend Software Engineer – X-team (Jan 2023 – Present)

DistrictTaco

- Designed and implemented a robust data infrastructure on GCP utilizing tools such as Airflow, BigQuery, dbt, Dataflow, and Cloud Run.
- Architected and deployed application infrastructure on Kubernetes.
- Containerized backend services into Docker images, optimizing for scalability and fault tolerance with Kubernetes deployment manifests.
- Switched Node.js API services to Python with FastAPI.
- Led the creation and upkeep of FastAPI services for mobile apps and company websites.
- Customized Helm charts for various open-source applications, including Superset and Airflow.

Data Architect/Engineer – Globant (Apr 2021 – Dec 2022)

PSO Google Cloud – Univision

- Led a team in designing and implementing a data quality project that involved creating data pipelines, transformation processes, and connectors to extract and load data from Braze through their REST API. Conducted comprehensive analysis and optimization of user data to identify and eliminate duplicates and improve overall data quality.

PSO Google Cloud – Banyan

- Developed scalable and efficient dataflow pipelines using Apache Beam and Airflow to facilitate seamless receipt data movement from GCS to Kafka topics for over 10 diverse clients with varying data sharing formats and schemas.

PSO Google Cloud– Autonomic AI

- Automated infrastructure migration from AWS to GCP using Python, Google's SDK, and database replication technology. Developed a robust solution that streamlined the migration process, minimized downtime, and ensured continuity of service for end-users.

Trainer for Big Data Analytics program at "Fundación Telefonica" (May 2021 – Dec 2022)

Trainer for Data Engineering couse at UTEC (Jul 2022 – Dec 2022)

Solutions Architect – Rimac Seguros y Reaseguros (Feb 2020 – Apr 2021) (1 year 2 months)
- Changed the **AWS** architecture in Data Ingestion, replacing DMS by Attunity (Qlik Replicate now) to replicate data and Glue PySpark by Redshift to deduplicate data from CDC. we saved 7,500$ per month (77% reduction) and reduced update times from 4 hours to 1 hour.
- Defined data flows to integrate data between AWS and GCP services. Data Lake's Migration.
- Reduced costs in s3 service applying lifecycle policies by 55%.

Big Data Engineer Senior – Everis Perú (Sep 2019 – Feb 2020) (6 months)
- Developed data ingestion pipelines using Databricks (Scala & Python) to move and transform data from blob storage to Cosmos DB in **Azure** cloud platform. We used Azure Data Factory to orchestrate data movement. Built reusable pyspark functions.

Big Data Engineer – Telefónica del Perú (Aug 2017 – Jul 2019) (1 year 10 months)
- Deployed Big Data projects from Telefonica Global (Projects about CAPEX Optimization, Payment Scoring, Terminal Recommender, Social Network among others).
- Built dashboards which contains information about network experience, service use and client value (Developed in Tableau).
- Developed a library in python to orchestrate the execution of BigData services. Built reusable components.

Software Developer (Self-Employment) (May 2016 – Jul 2019) (3 years 2 months)

Cerro Colorado Agro Company
- Developed software for different Modules (Sales, Purchase, Inventory, Cash Flow, Payroll, Expenses)
- Developed dashboards on Power BI Service.
- Developed profitability model.

Data Analyst – Banbif (Banco Interamericano de Finanzas) (Sep 2016 – Aug 2017) (11 months)
- Developed optimization model for emailing marketing to avoid saturation.
- Developed propensity models for credit card/passive products acceptance.
- Built dashboards for tracking sales performance, campaign effectiveness and business's prospection.

Business Analyst – Genomma Lab Peru (May 2015 – Feb 2016) (10 months)
- Built a datamart focus on replenishment with sellout and inventory outlets of customers data.
- Developed forecasts for sellout products used for creating annual budget and purchasing planning.
- Built sales dashboards to both channel of distribution (sellout through to Retailers and Wholesalers).

Education:

Universidad Nacional de Ingeniería (2010 – 2015) Bachelor of Science in System Engineering

Acknowledgments:
- [1st place Hackathon Rimac 2022](#) : Winner of the competition along with two colleagues and recognized as the best Machine Learning Engineer of the event.

GCP Certifications:
- Professional Data Engineer ✔ 12/01/2022. Associate Cloud Engineer ✔ 18/01/2022

- Professional Cloud Architect ✔ 24/01/2022. Cloud Digital Leader ✔ 28/01/2022
- Professional Cloud Developer Engineer ✔ 03/02/2022. Professional Cloud DevOPs Engineer ✔ 18/02/2022

Software and Programming Languages
- Power BI, Tableau, QGIS, Excel.
- Python, R, Java, Scala, Javascript.

Databases / Big Data tools
- Microsoft SQL Server 2012 (MCSA), Oracle 11g (PL/SQL), Teradata Database,Netezza , Postgres / Redshift.
- Hortonworks Data Platform (Hadoop, Hive, Sqoop, Spark, PySpark, Hue, Kafka)

DevOps
- Kubernetes, Docker, Git, CircleCI, Github Actions.