

# BAO TRAN

| Dec 30, 2000 | Ho Chi Minh City, Vietnam

## WORK EXPERIENCE

---

<b>Data Engineer</b>	<b>TMA Solutions</b>	<b>Jun 2022 - Present</b>
----------------------	----------------------	---------------------------

- Built, migrated and transformed cyber education client's data from transactional sources (Postgres, RDS, DynamoDB) to lakehouses on cloud infrastructures for downstream analytics and ML use cases (S3, Delta, Databricks).
- Responsible for implementing backend data pipelines and features for B2B Customer Data Platform (CDP) product, including automating marketing process, unifying customer data across contact points, and modeling NoSQL and relational data for efficient downstream application fetching and analytics (DynamoDB, Athena, Glue, Lambda).
- Extracted, Loaded, and Transformed (ELT) terabytes of data with PySpark and SQL queries for analytics serving on Redshift, Databricks and Athena
- Provisioned and maintained data infrastructures with IAC tools (AWS CDK, Terraform).

---

	<b>TMA Solutions</b>	<b>Feb 2022 – May 2022</b>
--	----------------------	----------------------------

<b>Data Engineer, Intern</b>
------------------------------

- Developed ETL jobs with Spark on AWS Glue and EMR to process gigabytes of data from AWS S3 data lake and loaded them to Redshift Database.
- Scheduled Lambda functions to orchestrate ETL jobs from S3 to Glue jobs.
- Experimented a proof of concept method of serverless ETL jobs directly through Lambda functions.
- Wrote SQL scripts to query data in Redshift for analytics service platforms.

## EDUCATION

---

<b>Bachelor of Science, Economics</b>	<b>Quy Nhon University</b>	<b>Aug 2018 – Jun 2022</b>
---------------------------------------	----------------------------	----------------------------

- Undergraduate Coursework: Statistical analysis; Time-series analysis; Quantitative methods; Strategic management; Financial market; Economics theory.

## PERSONAL PROJECTS

---

### Projects

- **Academic research – Co-author** (08/2020 – 03/2021): Funded by the university – Team leader of 5 students and 1 professor to carry out quantitative academic research focusing on statistical correlations between tourists' perceived intentions, actual behaviors, and their impact on the environment.

- **Cryptor – ETL pipeline**: Built a small near real-time data pipeline to extract, transform and load cryptocurrency

Terraform.

**Stacks** | Glue, PySpark, Lambda, EventBridge, S3, EC2, Terraform, Docker containers, Apache Kafka

## CERTIFICATES

- 
- **Databricks Certified Associate Developer for Apache Spark 3.0**: Pyspark.
  - **AWS Certified Solutions Architect**: Professional (SAP – C02).
  - **AWS Certified Solutions Architect**: Associate (SAA – C03).
  - **AWS Certified Data Analytics**: Specialty (DAS – C01).
  - **AWS Certified Machine Learning**: Specialty (MLS – C01).
  - **IELTS - 7.5 Academic** (12/2020): Reading - 8.5 | Listening - 8.5.

## TECHNICAL SKILLS

- 
- **Programming languages and frameworks**: Python, Java, SQL, PySpark.
  - **Tools** Amazon Web Services, Databricks, Postgres, Apache Spark, Apache Kafka, Terraform, CI/CD, Networking, Bash shell.