

SKILLS / TOOLS

Data engineering | Software development | Data cleaning | Database design | Version control | Machine learning

AWS: [Certified Solutions Architect – Associate](#) | Python | SQL | pandas | Spark / PySpark | AWS Glue | Step Functions
S3 | Athena | QuickSight | CloudWatch | IAM | Git | Airtable | Airflow | MySQL | Java | Scala | Power BI | Tableau

EDUCATION

The University of Texas at Austin, M.S. in Data Science (Aug. 2021 – May 2025)

- GPA: 3.97

Columbia University in the City of New York, B.S. in Mechanical Engineering (Sep. 2013 – May 2017)

- GPA: 4.05 departmental / 3.87 cumulative

EXPERIENCE

Data Engineer, Amazon [Contractor via Security Industry Specialists] (Jul. 2023 – Present)

- Architected and implemented a data platform on AWS to enable analytics for Amazon's Emergency Services Unit
- Created pipelines using AWS Glue and Step Functions to ingest data via APIs from SaaS sources such as Airtable
- Utilized Glue Catalog to create a data lake on S3 and used Athena to load these datasets into QuickSight dashboards
- Used AWS and Airtable to streamline operations via automation and implement mechanisms for data governance

Data Engineer, Amazon (Oct. 2022 – Mar. 2023)

- Supported Amazon Flex Analytics by creating and maintaining datasets for analysis and machine learning use cases
- Updated production databases by using AWS, PySpark, and Airflow to create and monitor scheduled data pipelines
- Created Redshift tables using an ETL workflow and altered existing tables to satisfy evolving business requirements
- Resolved urgent data availability and quality issues in production datasets in response to alarms and internal tickets

Data Engineer, Microsoft [Contractor via Design Laboratory] (Aug. 2021 – Oct. 2022)

- Helped Microsoft's MSAI Search Relevance team improve the quality of search results across Microsoft products
- Processed large datasets of search queries with Python and Spark to extract insights and engineer ML features
- Developed tooling and data pipelines on the Azure Machine Learning platform to automate engineering workflows
- Created Power BI dashboards to visualize key metrics, enable exploratory data analysis, and guide engineering effort

Teaching (Oct. 2018 – Apr. 2020)

- Established a sole proprietorship business providing in-person and online lessons in the Seattle area
- Taught Python and math to middle and high school students via a personalized curriculum

Mechanical Engineer, ASML (May 2017 – Feb. 2018)

- Developed mechanical designs that improved the performance of ASML's photolithography machines
- Designed and conducted statistical experiments and analyzed experimental data to guide design decisions

Researcher, Columbia University Creative Machines Lab (Sep. 2015 – May 2017)

- Created Eva, an open-source facially expressive robotic face, to facilitate artificial intelligence research in Python
- Video demonstration and published, peer-reviewed paper: <https://www.zanwarfaraj.com/eva.html>

PORTFOLIO

Designing and Implementing a MySQL Database on AWS for Financial Data (<https://zanwarfaraj.github.io/database/>)

- Designed a relational database for data about US stocks by creating an ER diagram and performing normalization
- Created a Python script to clean and transfer financial data from Interactive Brokers into the MySQL database