|  |  |
| --- | --- |
|  |  |

Data Engineer | Software Engineer

# Experience

|  |  |
| --- | --- |
| * **Bank of America** | * Dallas, TX (Hybrid) |
| * *AVP Data Engineer II | Permanent* | * *Sep. 2022 - Present* |

* + Designed, Developed, and Deployed an End to End Python Flask Data Platform used for tracking the MetaData and Linneage of 900+ reports totaling 1mil+ Data Points.
  + Restructured chaotic job scheduling system into a robust, centralized system written in Python | Luigi based orchestration.
  + Brought in to transform an archaic SQL only ETL process into a Hadoop based ELT process utilizing PySpark, Kafka, and the rest of the Apache Suite, processing over 20k Data Points per hour.
  + Developed CI/CD Pipelines, Jenkins, and Led design of new Git Repo Architecture
* Technologies: Python, SQL, Azure, Jenkins, Git, Hadoop, MapReduce, HDFS, Sqoop, Impala, CosmosDB

|  |  |
| --- | --- |
| * **LOOP** | * Dallas, TX (Remote) |
| * *Senior Data Engineer | Contract* | * *April 2022 - Sep. 2022* |

* + Hosted Airflow instance on Docker Container, via GCP, as central scheduler for routine jobs focused around Python, BigQuery, and GCS.
  + Designed data pipelines in GCP DataFlow and Fivetran for the ingestion of data from multiple 3rd party vendors.
  + Meticulously documented technical systems w/ Data catalogs, DBDocs & DBDiagram.
  + Integrated a 3rd party C# application into quote flow to dynamically insure our honest quote comparison between other providers; runs on more than 7.5k quotes a month.
* Technologies: Python, C#, Docker, GCP (Cloud Storage, Data Catalog, BigQuery), Airflow, ERD, API’s.

|  |  |
| --- | --- |
| * **Southwest Airlines** | * Dallas, TX (Remote) |
| * *Software Engineer (Data Platforms) | Internship* | * *Jan. 2022 - April 2022* |

* + Developed a Scala Framework to properly parse X12 EDI Documents.
  + Orchestrated and automated Scala ELT Pipeline, using Baker, to ingest 5k X12 EDI Invoices Daily into an On-Prem Postgres Data Warehouse.
  + Deployed and monitored pipeline using Docker, Jenkins, & TravisCI.
* Technologies: Scala, Baker, Postgres, Docker, Git, Jenkins.

|  |  |
| --- | --- |
| * **University of North Texas at Dallas** | * Dallas, TX (Remote) |
| * *Lead Analytics Engineer | Permanent* | * *Aug. 2020 - December. 2021* |

* + Oversaw a team of 5 engineers focused on Data Governance, Reporting, and Quality for the University.
  + Rewrote existing SAS pipelines in Scala, decreased time for State report generation by 35%.
* Technologies: Scala, SAS, SQL Server, Postgres, Azure, Git, API’s.

|  |  |
| --- | --- |
| * **Goldman Sachs** | * Dallas, TX (In Person) |
| * *Software Engineer (Data Platforms) | Internship* | * *Jun. 2021 - Aug. 2021* |

* + Developed Python Application to compare data lineage between application and cloud schema’s, reconciled 100% of mismatches.
  + Rewrote Enterprise Data Platform test cases to utilize pytest, testing coverage increased from  70% to 91%+.
  + Developed and Deployed a key Python Library (PAT) streamlining workflows with instant connections to 5 central data stores.
* Technologies: Python, JIRA, Confluence, API’s, Git, Docker, Kubernetes, Tableau. PyPi.

|  |  |
| --- | --- |
| * **Scout Energy Partners** | * Dallas, TX (Hybrid) |
| * *Software Engineer (Data Platforms)* | * *Jun. 2019 - Aug. 2020* |

* + Developed Scala ETL pipelines to retrieve more than 400k+ data points daily from Field SCADA Systems.
  + Led Development of a C# Blazor Data Platform to visualize and maintain data governance.
  + Developed and Deployed multiple Python & Scala Scripts & Applications, Orchestrated with Luigi & Baker.
* Technologies: C#, Python, Scala, Azure, Git, Github Actions, SQL Server, LaTeX

# Skills

|  |  |
| --- | --- |
| Orchestration: Airflow, Luigi, Prefect, Azure Data Factory | Databases: Postgres, CassandraDB, MongoDB |
| Python Frameworks: Flask, PySpark, pandas/numpy | Scala Frameworks: Spark, Kafka, Baker |
| C# Frameworks: Synapse, ADO.NET, Blazor | Cloud Platforms: GCP, AWS, Azure |
| Containerization: Docker, Kubernetes | Interest: FinTech, ClimateTech |

# Education

|  |  |
| --- | --- |
| * **University of North Texas** | * Denton, TX |
| * *M.S in Data Engineering* | * *Dec. 2022 – Dec. 2024 (Anticipated)* |

* + GPA: 4.0/4.0.
* Notable Courses: *Graph Theory, Empirical Analysis*

|  |  |
| --- | --- |
| * **University of North Texas at Dallas** | * Dallas, TX |
| * *B.B.A in Analytics* | * *Aug. 2020 – Dec. 2022 (Graduated)* |

* + GPA: 3.7/4.0.
  + Early graduation of 2 years.
* Notable Courses: *Statistical Modeling, Database Applications, Data Warehousing, Big Data Management.*

|  |  |
| --- | --- |
| * **Dallas Community College | *(Richland)*** | * Dallas, TX |
| * *A.A.S in Software Development* | * *Aug. 2016 – May. 2020 (Graduated)* |

* + GPA: 3.6/4.0.
  + Local Dallas County Community College District partnered with Emmett J. Conrad’s Global Collegiate Academy. Awards students with 60 college credit hours, and an Associate’s Degree in Applied Science.
* Notable Courses: *Operating Systems, C# Development, System Design, Java, Python, Networking.*

*Following sections items are clickable, unfortunately majority of our Robotics Team code is Private.*