Sia Seko Mbatia

Data Ops & Analytics Engineering
LinkedIn | **Personal Site** | **GitHub**

Senior data analytics professional with experience as a data ops and pipeline management lead; including data cleaning, wrangling, analysis, visualization, and storytelling. Interested in solving challenging data product and engineering problems with industry leaders.

Skills:

Human/People:

Empathetic and collaborative leadership, growth mindset, always in learning mode , attentive communicator, bilingual

Technical:

Web scraping, automation, ETL/ELT, data modeling, integration, cleaning & wrangling (API, XML, JSON, csv, etc.), DDL/DML, requirements gathering, process mapping, user acceptance testing, visualization and storytelling, documentation, troubleshooting

Tools:

Highly Skilled: Python, Bash, SQL (Vertica, MS SQL, Oracle, PostgreSQL, MySQL), JIRA/Confluence, Git, Google Data Studio

Skilled: Java, Ruby/Rails, BigQuery, Linux

Introductory Experience: Airflow, dbt, Google Cloud Platform, Snowflake

Education:

- **University of Nevada, Las Vegas** – M.S. Data Analytics & Applied Economics
- **Concordia College, Moorhead** - B.A Mathematics, B.A. Creative Writing

Certifications/Licenses:

- **Correlation One,** Endowed Dana Farber Cancer Institute Data Science Fellow, Honors
- **Anaplan** Certified Model Builder

Organizations:

- **Deepnote** Ambassador
- **BlackTIDES**, Strategy & Content

Work Experience:

COVID Black (Jan 2022 - present)
Database Consultant
- Winning five+ figure contracts by demonstrating database design competencies
- Creating a guide for the history of COVID-19 in the Black diaspora through unstructured data wrangling and management with Python and Excel.

BlueLabs Analytics (Oct 2020 - Nov 2021)
Lead, Data Operations & Pipeline Management
- Drove 5 million hc.gov insurance enrollments after improving ETL data pipeline used for email and SMS outreach for a 3+million per year government contract.
- Decreased daily data ingestion workflow runtime by one hour/day by improving automation of Python and Ruby scripts for 110+ million transformed records
- Promoted to Senior Analyst within 3-months and subsequently to Data Operations Lead

Candidates.ai (Sep 2020 - Mar 2021)
Data Engineering Intern
- Removed 100% of redundant data processing and ingestion fromS3 buckets by creating an automated change log using Python scripts in PostgreSQL
- Saved 65% of executive search firms' sourcing time by performing ETL on Form 990 XML files to narrow down searches and target domain specific candidates.

Black Women's Business Collective (Jun 2020 - Sep 2020)
Data Consultant
- Built a web-based database application to catalog & query 2500 Black women owned businesses with MySQL and Google App scripts.

Plus Relocation Services (Sep 2015 - Jun 2018)
Global Compensation Data & Process Analyst
- Reduced 33% of weekly manual file processing time by automating/re-evaluating reporting procedures with Microsoft SQL and SPSS
- Increased sales by $70,000+/year by templating cost projection calculations, allowing team to increase client intake by over 50% for two years straight.
- Saved over $150,000 annually for two years taking on project manager and product owner duties for a non-technical compensation team of four.
- Became team's SME for data infrastructure and automation solutions in championing the creation of a proprietary expense management (EM) system.

Correlation One (Apr 2021 - present)
Data Engineering Curriculum Developer, Advisor, Teaching Assistant
- Facilitate and support over 1200+ students' learning in a 16-week intensive, immersive data science program
- Partner with leadership/TAs to develop curriculums, matching employer demands, interviewing, and boosting fellow performance with 94% employment rate in 2021