Jade Zhang

Levittown, PA | 267-667-6951 | szhangjd21@gmail.com | Portfolio | LinkedIn | Interactive Resume | GitHub

Objective

Results-driven data architect/analyst with healthcare experience, leveraging expertise in ETL, data integration, and analytics. Demonstrated ability to build scalable pipelines and optimize data systems to improve patient outcomes and operational efficiency.

Work Experience

NewYork-Presbyterian Hospital | New York, NY

Data Architect/Analyst | *Python, R, SQL, Visual Basics, ETL, SSIS, Git, FHIR, RESTful APIs, dbt, EHR* 04/2024 – Present

- Led and executed a data integration project from scratch using **Python, SQL, SSIS, dbt, and RESTful APIs** to ingest and transform **BCDA** (Beneficiary Claims Data API) data in **FHIR** (Fast Healthcare Interoperability Resources) framework, and merge it with **CCLF** (Claim and Claim Line Feed), significantly improving data completeness and usability.
- Built and maintained **ETLs** for commercial claims and **Epic Clarity EHR** extracts using **SSIS, Python, R, and SQL**, resolving critical issues in legacy pipelines, mining and transforming data from Epic Clarity, and integrating structured data into custom-designed tables within a customized Caboodle database.
- Identified flaws in existing commercial **supplemental files** and developed an improved SQL-based solution for multiple entities, increasing performance by **10–25%** on key HEDIS measures.
- Contributed to the CMS reporting process by supporting data preparation, validation, and submission for **eCQM** and **GPRO** reporting methods.

Statistical Data Analyst | *Python, SQL, Tableau, Power BI, Epic, Git* 09/2023 – 04/2024

- Implemented a robust pipeline in **Python** using **regular expressions and NLP** to extract over **95%** of unstructured text from doctor notes for the depression screening measure, boosting performance by more than **10%**.
- Developed and maintained **SQL** logic for the Kidney Health Evaluation and Adolescent Immunizations HEDIS measures, ensuring timely updates to reflect the most accurate and recent data.
- Designed three interactive dashboards in **Tableau** and **Power BI** to track and visualize commercial performance metrics.
- Integrated **GitHub** for version control and led the initiative to onboard the team to Git best practices.
- Collaborate with team members to support patient chart reviews, interviews, and code development and reviews.

Fulton Bank | Lancaster, PA 05/2023 – 08/2023

Data Science Intern | *Python, PowerShell, Time Series Forecasting, Optimization, Power BI, SQL*

- Crafted a keyword-based search utility using **Python** and **PowerShell** Script to extract corresponding reports, tables, and data sources within **~1,200 Power BI and SQL reports** on the server, significantly improving search efficiency for over **60** team members.
- Co-developed a forecasting and optimization pipeline using **machine learning** in **Python** to predict cash balances for 204 branches, enabling a utility projected to save **$5.5–10M** annually; prototype won **first place** in an internal datathon and was moving toward production.

Drexel University | Philadelphia, PA 08/2021 – 01/2023

Decision Science Graduate Researcher | *Python, MATLAB, Time Series Forecasting, Machine Learning Clustering,*

- Proposed and implemented a **bisecting hierarchical clustering algorithm** for time series data to forecast college enrollment, improving accuracy by **15%** over conventional methods.

Certificates

AWS Solutions Architect (2025) | HL7 FHIR Implementer (2025) | HIPAA Compliance Certificate (2025) |
Epic Caboodle Data Model (2023) | Epic Clarity Data Model (2023) | Epic Cogito (2023)

Areas of Expertise

Core Skills: ETL, Data Integration, Healthcare Data (Epic, FHIR, HIPAA, EHR), Time Series Forecasting, Machine Learning, NLP, Data Visualization, HEDIS Measures, Collaboration & Communication

Programming & Tools: Python, R, SQL, PowerShell, SSIS, Tableau, Power BI, Git, RESTful API, AWS, dbt

Education

Master of Science, Economics and Computer Science 09/2023
Drexel University | Philadelphia, PA | Drexel Merit Scholarship

Master of Science, Business Analytics 05/2021
Clark University | Worcester, MA | Clark SOM Dean's Scholarship

Bachelor of Science, Applied Mathematics 05/2019
University of Maryland, College Park | College Park, MD

Personal Projects

Wids Datathon Competition - Survival Rate Prediction | *Machine Learning, Ensembled Methods* | Code

- Predicted patient survival rates using ML Ensembled Methods and finished in the top 10% of participants.

Real or not? NLP with Disaster Tweets | *NLP, Deep Learning, PySpark* | Code

- Determined if the Tweets were emergency announcements and obtained accuracy > 85% and AUC > 0.89.