



## **Position: Data Engineer**

**Reports To: Chief Technology Officer**

**Job Type: Full Time Position**

### **Summary/Job Description**

Advocate Community Providers (ACP) supports community-based primary care physicians serving some of New York City's most vulnerable Medicaid patients. The Data team at ACP builds, buys, and maintains a suite of cutting edge population health, quality, and cost analytics systems. By leveraging data intelligently, we are able to connect physicians, patients, hospitals and ACP administrative staff in coordinating care.

ACP Data Engineers work on building ETL pipelines, analysis tools, and core data processing components in ACP's cloud-based data processing environment. We are heavily biased towards open source/\*nix tools, but will always prioritize using the best tool for the job. At ACP, data is our lifeblood, and the Data team is responsible for ensuring ACP's data is up-to-date, clean, secure, and readily available.

### **Key Job Responsibility Areas**

- Architecting and building python-based ETL jobs to connect with ACP's data partners
- Light web application development of purpose-built internal tools
- Technical guidance in support of our project management team when defining the scope of data integration projects
- Contribute to designs of new components in modeling and data pipelines
- Remain current on emerging open source data processing projects and tools

### **Core Competencies**

- Curious, informed and opinionated about data processing technologies (kafka vs. rabbitmq, Presto vs. SparkSQL, etc)
- Experience with structured and unstructured data storage and modeling
- Deep understanding of database and filesystem storage/access
- Experience with various data engineering architecture patterns – pub/sub vs. batch ETL, etc.
- Interest in Data Science and Data Analysis



## Qualifications

- Three or more years of experience working with \*nix-based, open source data processing tools
- Fluent in Python, Ruby, SQL, shell scripting, or whatever combination of the above gets the job done
- Two or more years of experience developing production ETL applications in a language of your choice
- Four or more years of experience in software development
- Three or more years of experience with SQL

## Preferred Education and Experience

- BS or MS in Computer Science
- Experience writing production Python
- Implementation experience with Aurora, Airflow, Luigi, etc.
- Experience with healthcare data formats (x12 EDI, HL7, etc)
- Experience implementing stream processing pipelines (Spark, Storm, etc)
- MapReduce/Hadoop ecosystem experience (Hive, HDFS/S3, Presto)

## **Additional responsibilities, tasks and duties may be assigned as necessary.**

*All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, gender, identity or expression, sexual orientation, national origin, marital status, genetics, disability, age, veteran status or any other legally protected status.*

*ACP is committed to providing service that is culturally and linguistically appropriate for our diverse partnership. We work to ensure that our philosophy of cultural and linguistic diversity is embraced in all levels of our organization. Culturally competent services are required from both our employees and our partners.*