Gorman Consulting (https://gormanconsulting.org) is a firm that provides specialized measurement and evaluation services, primarily in the health sector within low- and middle-income countries. We are seeking a consultant data engineer who will help set up and maintain a database and accompanying extract, transform, load (ETL) infrastructure for COVID-19 vaccine clinical trial results for our client. The ideal candidate is meticulously attentive to detail and proficient in Python/SQL with 5+ years of relevant experience. Candidates that also bring experience with data cleaning/analysis in Stata/R or data visualization experience with Tableau will be prioritized with the potential for a broader set of responsibilities relative to those below.

Summary

- Start date: November 1, 2021
- Hours: Full-time or part-time, depending on availability and breadth of experience.
- Duration: One year with high likelihood of extension
- Location: Remote, though candidates in a US time zone are strongly preferred
- Compensation: A competitive consulting rate that is commensurate with relevant experience

Responsibilities:

- Support building the infrastructure required for optimal extraction, transformation, and loading (ETL) of clinical trial data that we receive in batches from partner laboratories.
- Use Airflow to automate ETL processes.
- Design the schema for our staging and production databases.
- Deploy our data pipeline on Heroku and Amazon Web Services.
- Support the team to define and employ quality checks of incoming clinical trial data; work with partner laboratories to identify and resolve data issues.
- Manage the SSH File Transfer Protocol for receiving incoming data.
- Generally, contribute to all aspects of designing, developing, testing, and maintaining our data pipeline system while creating and managing clear documentation for internal/external use.

Requirements:

- 5+ years of relevant experience as a data engineer building components of a data warehouse and data pipeline including setting up and maintaining relational databases (MySQL, MS SQL, PostgreSQL), automating ETL workflows and using ETL software such as Airflow or Luigi.
- Proficient in Python (both functional and object-oriented programming) and in SQL.
- Experience with cloud computing, preferably Heroku and AWS.
- Excellent organization/interpersonal skills as well as strong written and verbal English communication abilities. Ability to organize/discuss complex technical information in a clear and concise way.
- Meticulous attention to detail. For example, please use the subject line "Application for Data Engineer: *Your Last Name*" in your email application.
- Experience with sensitive data or protected health information under GDPR and/or HIPAA frameworks.
- Self-starter who will thrive in a fast-paced, start-up environment by taking ownership and initiative.
- While not required, the following are preferred: experience with data visualization software such as Tableau; data analytic experience using Stata or R; experience with clinical trial or immunological data.

To apply, send your CV to jobs@gormanconsulting.org. Only shortlisted candidates will be contacted. Candidates must have legal authorization to work within the country they are based.