

Job Description for the role: Data Engineer

Interested? Send your resume to tech-hiring@bloccpower.io and say "Data Engineer"

About BlocPower

BlocPower is a clean energy startup creating Smarter, Healthier, Greener Buildings for All by reducing the barriers to money-saving, quality-of-life-improving green building retrofits. We provide engineering, financing, and project implementation services for clients in frontline communities across the country. These communities are underserved by traditional energy services companies because they are considered too small, too costly, or too risky. Our portfolio of projects include houses of worship, schools, non-profits, small businesses, and multifamily buildings. Through our work, we save our clients money, reduce greenhouse gas emissions, improve health, and create local employment opportunities.

At BlocPower, we value our mission. We are trusted advisors that get things done for our customers by using data to make the right decisions. We support and expect excellence from our team members. We treat both our customers and ourselves with affection and respect.

Job Overview

BlocPower is looking for a friendly Data Engineer to join our growing team to be responsible for expanding our data pipeline architecture. You should have a depth of experience building data pipelines to draw from, and see yourself adding to as much as benefitting from a supportive team environment. The Data Engineer will help keep data consistent across projects and teams. This may look at times like helping our engineers to perform energy audits using our proprietary models, writing software that will help us analyze big data that comes into our platform via sensors installed across our projects, or optimizing the project pipelines of our construction and sales teams.

The ideal candidate is someone who is self-directed, believes in our mission, and is an excellent written communicator.

Job Responsibilities

- Assemble large, complex, structured/unstructured data sets from various public/non-public data sources into the raw-zone of our data lake.
- Build the infrastructure required for optimal extraction, transformation, and loading of data from a wide variety of data sources using SQL and AWS 'big data' technologies.
- Design and implement internal process improvements: automating manual processes, optimizing data delivery, re-designing infrastructure for greater scalability, etc.
- Work with stakeholders including the Executive, Product, Data and Design teams to assist with data-related technical issues and support their data infrastructure needs.

- Keep our data and associated architecture secure.

Desired Skills

- Write good SQL, and work with a variety of databases (eg. AWS RDS, MySQL, PostgreSQL, and MongoDB)
- Stand up and maintain an ETL pipeline with little to no supervision, using tools such as AWS Glue, Informatica, or Talend
- Coherently organize a data lake, making it easy to collaborate on data
- Manipulate big data, including high-velocity streaming data using tools such as Spark, Kafka and Kinesis
- Strong project management and organizational skills

Nice to Have

- Experience with AWS cloud services: EC2, EMR, RDS, Redshift, Glue, Sagemaker
- Microservice architectures
- Docker, Kubernetes, Docker-Compose
- Git and Jenkins
- Experience working with IoT/Sensors

Education/Experience

- BA/BS or equivalent combination of work experience and education preferably in degree/course work/experience in computer science/data engineering
- 2+ years of experience as a Data Engineer