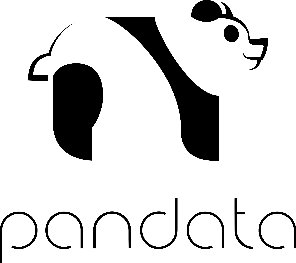
****

**Data Engineer**

Pandata is a Cleveland-based group of data scientists who help organizations derive value from data. We’re rapidly becoming the go-to resource for data science solutions in Northeast Ohio with services spanning critical functions from sales and marketing to operations and security.

**What you will do**

* Collaborate with our data scientists to acquire, curate, and clean data from a wide variety of both public and private sources
* Collaborate with data scientists to scale predictive models and machine learning algorithms on large data sets
* Select and integrate big data tools and frameworks to support client solutions
* Develop and maintain ETL workflows
* Provision and maintain databases and cloud computing environments
* Continue to grow your knowledge of data engineering tools and processes
* Represent Pandata at various community and regional events

**Your background**

* Undergraduate degree in software engineering (computer science, software engineering, systems engineering …) or equivalent certification or experience
* Prior experience implementing ETL processes from a variety of sources (Databases, JSON, XML, HTML, Text, etc..)
* Prior experience provisioning and maintaining database systems
* Demonstrated creativity and problem-solving aptitude
* **Preferred:** experience with Google Cloud Platform or Amazon Web Services
* **Preferred:** experience with Hadoop and HDFS
* **Preferred:** experience with Big Data querying tools (Pig, Hive, and Impala)

**Special note:** Willing to consider paid apprenticeship for the right candidate

**Benefits**

* 401k plan
* Health, dental, and vision
* Unlimited Coursera certifications