

Who We Are

HemaSource is an innovative logistics solution partner that combines analytics and high quality medical product supply to specialty ambulatory clinics. Our mission is to drive customer operational efficiencies and related cost savings in high volume ambulatory care clinics. The HemaSource proprietary data analytics software tool and our superior customer service have enabled us to become the preferred customer product & solutions provider to the ambulatory healthcare market.

Project Background

HemaSource has recently invested in building a best-in-class data platforms from the ground up, leveraging a modern, AWS-based data stack (Python, Spark, Glue, Athena). We already have a rich dataset (thousands of tables) on medical supplies use throughout the entire supply chain lifecycle, with more data sources to be integrated in the future. At present, we use this data internally for business metrics, but will be expanding the reach throughout the business to customers, dashboards, and machine learning products.

Project Description

Students will have the opportunity to building end-to-end data products, including:

- Building ETL pipelines to ingest third-party data sources into our Data Lake.
- Modeling new data entities and building ETL pipelines into a star schema Data Warehouse.
- Building analytic products for stakeholders and customers. This could include dashboards and data visualizations, BI tables, and machine learning products.

The problem space for this work will be decided based on the student interest and company needs. HemaSource has many areas to apply analytics, including:

- Minimizing freight costs
- Increasing warehouse operational efficiency
- Analyzing product margins
- Forecasting future plasma donations

The analytic deliverable(s) and project plan will be finalized based on student preferences and experience before the session starts.

Requirements

We are looking for a team of 2-4 engineering students. Students should have an interest in data engineering and/or data science. At least beginner-level exposure to Python, Data Frames, and SQL is required. Some exposure to data modeling techniques, big data processing tools, and cloud computing is helpful but not required. The team will work under the guidance of an experienced Data Architect/Data Scientist.

This will be a great opportunity for students to get hands-on experience with the full data lifecycle using industry standard workflows and tools. Moreover, since this is a new initiative within a young company, there is plenty of opportunity for impact with data science products and analysis.

Our office is in the heart of Olde Town Arvada, one block from the RTD G Line Station. Given COVID-19, students will be working remotely during the field session. We will schedule weekly virtual collaboration and sprint planning sessions.

Intellectual Property

HemaSource data is for internal use only. HemaSource will retain ownership of the code and artifacts produced, but students will be able to reference it and present their work in conferences, meetups, white papers, etc.

Contact

- Kayla Andersen, Principal Data Scientist and Data Architect: kandersen@hemasource.com
- Joe Malory, VP Data Analytics and Customer Service: jmalory@hemasource.com