



The Pharma IT Data Sciences group is recruiting an outstanding intern interested in designing, implementing, and delivering data mining solutions that have direct impact on patient health. The Data Sciences group spans across the entire value chain at Janssen from discovery to post launch and commercial, and its core mission is to leverage data to create insights that impact health care. This uniquely positioned organization will enable talented interns to work with various business units to help identify viable data mining opportunities and deliver novel, end-to-end data analytical solutions.

Internship Description

The candidate will utilize data science techniques to acquire, augment, and analyze diverse data sets, providing business solutions for Janssen. The role requires both a broad knowledge of data wrangling, data integration methods, ontology management, and dashboard/visualization creation. The ideal candidate will have grounded knowledge in the field of data sciences and creativity to invent, customize, and work in a multidisciplinary environment to drive business solutions. The candidate will be part of a dynamic, accomplished team that will support the Data Sciences organization.

Position Responsibilities

The candidate will work with colleagues in Data Sciences and partner with key businesses across Janssen to help in applying advanced analytics to answer business questions related to projects in one of the key focus areas of the group (Machine Learning, Real World Evidence, Big Data Analytics or Semantic Integration & Business Insights). She will be a full member of a project team to help connect with other data scientists and business partners to gather requirements around a critical business need, and also act in a primary role as a trusted primary IT partner. She will collaborate with subject matter experts from the business and use their feedback to create and iterate on data solutions.

Required Skills

- Familiarity with large datasets, handling of biomedical datasets and understanding of data analysis workflows is required.
- Proficiency in one or more programming languages such as Python, Perl, Java, C++, etc. is required.
- Coursework or background in biology, chemistry, biochemistry, molecular biology and other life sciences is strongly preferred.
- Strong understanding of Statistics as applied to data analysis and experience with statistical package such as R is preferred.
- Experience with data analysis, visualization and workflow software is preferred.
- Experience with accessing and working with relational databases using SQL is a plus
- Experience with web development (HTML, Javascript, CSS, PHP, etc...) is a plus.