

## JOB SUMMARY

Covalent Solutions (Covalent) is seeking a mission driven, highly motivated and collaborative **Data Scientist** who is truly excited about unleashing the full potential of data assets by bringing technology and/or data and insights to the forefront of decision making via fit-for-purpose data analytics solutions to support the *All of Us* Research Program (AOU) at the National Institutes of Health (NIH). AOU was established to accelerate health research and medical breakthroughs to enable an era of precision medicine for all. The program seeks to achieve this mission through building relationships with one million or more participant partners, delivering the largest, richest biomedical dataset ever, and catalyzing a robust ecosystem of researchers eager to use the dataset to advance human health. By combining health-related information from a large pool of diverse participants, AOU will reach the scale and scope necessary to enable research on a wide range of diseases and health topics.

This is a newly created role. The Data Scientist will work within a growing team under the Senior Advisor to AOU's Chief Executive Officer (CEO). You will build sophisticated data pipelines and products to wrangle and analyze data, drive analytics innovation and experimentation to enable data-driven insights for complex business and scientific problems, thereby playing a key role to enable application of advanced analytics techniques across the organization.

This opportunity is full-time with Covalent and is on-site in Rockville, Maryland.

## DUTIES AND RESPONSIBILITIES include, but are not limited to:

- Design data analyses to answer hypothesis-based questions.
- Develop enterprise level code with strong markdown skills able to be replicated and understood by others.
- Perform code review of analyses done by others in both R and Python.
- Ensure consistency in analysis techniques and delivery against cross-functional needs.
- Ensure accuracy of results and analytic methods.
- Ensure data integrity and quality prior to analysis.
- Follow and implement solid research methodologies in study design, evaluation planning, and reproducibility of results.
- Evaluate and assess available relevant data sources to bring them together for a given project.
- Perform complex day-to-day ETL tasks such as data gathering, data cleaning, wrangling, coding or programming, business and analytics requirements gathering, and data analysis.
- Use creativity to solve data and analytical problems for the organization as well as be able to visualize complex data into an easy-to-understand format for interpretation by leaders across the organization.
- Promote and leverage advanced analytic tools and techniques where applicable.
- Become a trusted analytical partner and collaborate on the development and planning of analytic projects in response to AOU needs.
- Communicate status at the executive level as well as the execution level - being able to go deep at the drop of a hat, while also being able to hit the headlines for a quick summary.

- Collaborate with partners. Be exceptional at developing and nurturing relationships with key internal and external stakeholders in support of AOU's goals.

## QUALIFICATIONS

### Education:

- MSc/PhD in data science or other advanced degree in life sciences with post-doctoral or other training/work in Clinical Informatics or related field.

### Experience:

- Superior coding skills using common data science tools, including R, Python 3, and Jupyter notebooks in a cloud computing environment.
- Background participating directly in the scientific lifecycle including hypothesis testing and basic scientific methodology.
- Proven experience with data visualization and presentation techniques for products including presentations and papers. You can share several examples of real work you've delivered, and what you learned along the way.
- Consistent track record of delivering value through the use of complex data collected from healthcare and/or biomedical research settings.
- Demonstrated ability to build long-term relationships with partners at all levels, understand relevant scientific/business challenges at a deep level and translate into informatics activities to deliver defined value.
- Proven experience working with large cross-functional teams across multiple platforms including Microsoft and Google products.
- Strong ability to work with structured and unstructured data.
- Hands-on experience designing analyses to answer questions independently and using data to support conclusions.
- Experience must demonstrate increased independence and overall responsibility for more complex projects.
- Experience in data processing using SQL, Hive, Impala, Spark, or equivalent querying language a plus.
- Understanding and experience with advanced statistics and modern machine learning predictive techniques, including GLMs, decision trees, forests, boosted ensembles, neural networks, deep learning, and graph analytics a plus.
- NIH experience a plus.

## COMPETENCIES

### Professionalism:

- Ability to apply judgment to manage conflicting priorities and resources, driving work and results that are the highest value for the organization.
- Sharp eye for detail. Keep meticulous notes and be able to manage multiple projects in different phases of work at the same time. Context-switching does not bother you.

- Ability to identify issues, analyze and participate in the resolution of issues/problems.
- Motivated by professional rather than personal concerns.
- Show persistence when faced with difficult problems or challenges.
- Ability to stay calm under stress and uncertainty, influencing the same in your teammates.
- Naturally curious about things you don't understand and actively look for opportunities to grow and learn from those around you.

**Communication:**

- Actively participate in meetings and be willing to advocate for what you believe in. Provide thoughtful, constructive review of others' work, knowing that everyone is working towards the same objective.
- Superior written and verbal communication skills. You can always articulate the "why" behind your recommendations.
- Listen to others, correctly interpret messages from others and respond appropriately.
- Ask questions to clarify, and exhibit interest in having two-way communication.
- Tailor language, tone, style and format to match the audience.
- Demonstrate openness in sharing information and keeping people informed.
- Speak intelligently about a wide variety of topics but know when to say, "I don't know."

**Teamwork:**

- Ability to form trust-based relationships quickly and lead cross-functional teams through influence (without formal authority).
- Work collaboratively with colleagues to achieve organizational goals.
- Place team agenda before personal agenda.
- Support and act in accordance with the final group decision, even when such decisions may not entirely reflect your own position.
- Share credit for team accomplishments and accept joint responsibility for team shortcomings.

**Accountability:**

- Take ownership of all responsibilities and honor commitments.
- Deliver outputs for which you have responsibility within prescribed time, cost and quality standards.
- Operate in compliance with organizational regulations and rules.
- Take personal responsibility for your own shortcomings, where applicable.