DATA SCIENTIST SENIOR

The PNC Financial Services Group, Inc. seeks a Data Scientist Senior in Pittsburgh, PA with the ability to work from home with appropriate telecommuting systems for up to two days per week, with a minimum of three days per week in the office. Within the Retail Lending Credit & Operational Intelligence department, the position will be responsible for performing analytical tasks on vast amounts of structured and unstructured data to extract actionable business insights. Specific duties include: (i) performing quantitative financial analysis using statistics, exploratory data analysis, machine learning models and time series analysis to enhance credit risk management and operational efficiency; (ii) engaging in gathering, processing, and mining of large and complex quantitative financial datasets persisted in Hadoop and cloud datastores; (iii) building automated data workflows to consolidate and summarize data for weekly/monthly reporting to recommend optimal actions to management; (iv) developing quantitative algorithms using advanced mathematical and statistical techniques like machine learning, deep learning and natural language processing to drive decision making at PNC; (v) running quantitative financial analytical experiments in a methodical manner to find opportunities for product and process optimization; (vi) assisting in the presentation of business insights to management using visualization technologies and data storytelling; and (vii) partnering with Data Architects, Data Engineers, and Visualization Experts to develop data-driven quantitative financial solutions for the business.

Master's degree in Analytics, Economics, Engineering, Statistics, Mathematics, or Information Systems Management plus 3 years of experience in data querying with SQL and programming in Python for developing applications for a large organization (>10,000 employees) is required. Must have experience with: (i) handling, manipulation and analysis of large datasets (multiple gigabytes of data) persisted on Hadoop, Neo4j and cloud data platforms; (ii) using statistical analysis techniques including logistic regression, time series analysis, and hypothesis testing; (iii) utilizing data query tools including SQL, R, and Python to manipulate, analyze and interpret data; (iv) writing and implementing code to clean and transform large tabular and text datasets (typically exceeding 1 million rows of data) to a consistent format usable in data visualization and developing machine learning models; (v) enriching in-house data by integrating with external data sources (Bureau and third party vendor data) to extract and analyze trends for risk management; (vi) designing rich data visualizations to communicate complex ideas to customers or company leaders; (vii) programmatically extracting data from a Hadoop database and transforming the data into a presentable form using python plotting libraries (Matplotlib and Seaborn) and Tableau visualizations; and (viii) building machine learning models to drive business insights and improve financial outcomes.

40 hours/week, 8:00 a.m. – 5:00 p.m. Interested individuals apply online at www.pnc.com using keyword R188602. PNC provides equal employment opportunity to qualified persons regardless of race, color, sex, religion, national origin, age, sexual orientation, gender identity, disability, veteran status, or other categories protected by law.