



## **Employer's Website Posting**

**TITLE:** **APPLIED DATA SCIENTIST**

**DUTIES:** Responsible for Machine Learning (ML) and Data Science (DS) Practitioner involving constructing machine learning models from start to finish including data collection, normalization, and standardization, data pipeline construction, model selection and hyperparameter tuning, and in production, out of sample, working ML systems that can add new data into ML model. Create apps to present ML models or DS projects and host them in the cloud or locally. Create pipelines to query and retrieve and update data for existing applications to keep them up to date. Responsible for the classification and regression of algorithms for ML, including boosting, bagging, random forests, and deep learning neural net configurations, unsupervised learning algorithms including k-means, anomaly detection, and hierarchical clustering. Develop visualization packages including ggplot, seaborn, matplotlib, bokeh, plotly, and dash. Develop Deep Learning packages including TensorFlow and Keras. Responsible for programming Natural Language Processing (NLP) packages such as NLTK and Data Scraping packages such as Scrapy and BeautifulSoup. Responsible for data extraction, transformation and loading, including standardization and normalization, identification of outliers, log-transformation. Utilize programming Python language including packages such as numpy, scipy, pandas, scikit-learn, XGBoost, LightGBM, and CatBoost. Provide SQL Database understanding including JOINS, MERGE and Schema layouts. Create advanced Excel spreadsheets including VBA macro programming, Pivot Tables, SUMIFS, AVERAGEIFS, OFFSET functions. Develop spinning up remote servers, containers and containers management systems such as Docker and Kubernetes using AWS and Azure cloud architecture. Develop company's webpage for use with Javascript, CSS, HTML, RESTful APIs.

**REQUIREMENTS:** Master's Degree in Mathematics, Data Science or closely related fields and three (3) months of experience in the job or as a Data Scientist or closely related occupation responsible for creating visualization packages, building classification models, mining data from SQL server, developing Python scripts, and providing feedbacks on data quality issues to improve data solutions. Must have experience utilizing various Machine Learning models, Python libraries, and SQL Data Visualization tools/packages, including Matplotlib, Seaborn, Keras, JOINS, MERGE, SQL, and others. Must have experience using advanced Natural Language Processing (NLP) techniques/architecture/tools.

**Hours/Week:** 40hrs // (Mon–Fri) // 9:00am – 6:00pm

**Job Location:** 165 Broadway, New York, NY 10006

**CONTACT:** Andrew De Luca, HR  
Transatlantic Reinsurance Company  
165 Broadway, New York, NY 10006  
Tel/No: 212-365-2242  
E-mail: [Adeluca@transre.com](mailto:Adeluca@transre.com)

---

---