Responsibilities for Data Scientist

* Work with stakeholders throughout the organization to identify opportunities for leveraging company data to drive business solutions.
* Mine and analyze data from company databases to drive optimization and improvement of product development, marketing techniques and business strategies.
* Assess the effectiveness and accuracy of new data sources and data gathering techniques.
* Develop custom data models and algorithms to apply to data sets.
* Use predictive modeling to increase and optimize customer experiences, revenue generation, ad targeting and other business outcomes.
* Develop company A/B testing framework and test model quality.
* Coordinate with different functional teams to implement models and monitor outcomes.
* Develop processes and tools to monitor and analyze model performance and data accuracy.

Qualifications for Data Scientist

* Strong problem solving skills with an emphasis on product development.
* Experience using statistical computer languages (R, Python, SLQ, etc.) to manipulate data and draw insights from large data sets.
* Experience working with and creating data architectures.
* Knowledge of a variety of machine learning techniques (clustering, decision tree learning, artificial neural networks, etc.) and their real-world advantages/drawbacks.
* Knowledge of advanced statistical techniques and concepts (regression, properties of distributions, statistical tests and proper usage, etc.) and experience with applications.
* Excellent written and verbal communication skills for coordinating across teams.
* A drive to learn and master new technologies and techniques.
* We’re looking for someone with 5-7 years of experience manipulating data sets and building statistical models, has a Master’s or PHD in Statistics, Mathematics, Computer Science or another quantitative field, and is familiar with the following software/tools:
  + Coding knowledge and experience with several languages: C, C++, Java,
  + JavaScript, etc.
  + Knowledge and experience in statistical and data mining techniques: GLM/Regression, Random Forest, Boosting, Trees, text mining, social network analysis, etc.
  + Experience querying databases and using statistical computer languages: R, Python, SLQ, etc.
  + Experience using web services: Redshift, S3, Spark, DigitalOcean, etc.
  + Experience creating and using advanced machine learning algorithms and statistics: regression, simulation, scenario analysis, modeling, clustering, decision trees, neural networks, etc.
  + Experience analyzing data from 3rd party providers: Google Analytics, Site Catalyst, Coremetrics, Adwords, Crimson Hexagon, Facebook Insights, etc.
  + Experience with distributed data/computing tools: Map/Reduce, Hadoop, Hive, Spark, Gurobi, MySQL, etc.
  + Experience visualizing/presenting data for stakeholders using: Periscope, Business Objects, D3, ggplot, etc.