**PROGRAMS**

* SAS
* Excel
  + M.O.S.
* Relational or MPP/distributed systems:
  + Oracle
  + Teradata
  + Vertica
  + Hadoop
* Java
* Python
* Scala
* ETL data pipelines
* Java
* Spark
* Hive
* Oozie
* Kafka
* Map Reduce Preferred
* R
* Python
* SQL
* Tableau
* Spark
* MapReduce
* HDFS
* Cassandra
* Kafka
* C
  + Objective-C
  + C++
* POSIX/synchronization primitives
* SQA
* Git
* CI systems
  + Jenkins
  + Xcode server
* scripting and automation (Python)
* QA automation
* NLP
  + image classification or deep learning Preferred
* Apache
* Hadoop
* Spark
  + or other MR frameworks
* multi-threaded programming, network, and POSIX/synchronization primitives
* cloud computing
  + AWS
  + Azure
    - or similar cloud platform

**EXPERIENCE**

* advance courses Preferred
  + data science
  + machine learning is
* 3+ years of industry experience working with large-scale and high-throughput systems
* 2-3 years of directly related experience required
* 3 or more years’ experience building large-scale distributed systems
* 2+ years of professional experience in C, Objective-C, or C++
* 5+ years of experience working on data science projects.
* Hands-on experience with real world, big data is required
* Knowledge of
  + regression
  + time series analysis
  + clustering
  + decision tree techniques
* background in data product engineering
* 8 years of experience in data analytics
* Master’s Degree **plus** 6 years of experience in data analytics
* PhD plus 4 years of experience
* At least 3 years’ coding experience with open source programming languages
  + R, Python, or Scala
* 3 years’ experience developing commercial data science solutions leveraging one or more of the following:
  + operations research
  + natural language processing
  + machine learning
  + deep learning
  + video/image analysis,
  + time-series analysis
* Master’s Degree **plus** 6 years of experience in data analytics
* PhD plus 4 years of experience
* Master’s Degree Preferred
  + - “STEM” Major (Science, Technology, Engineering, or Math)
* Ph.D. Preferred
  + - “STEM” Major (Science, Technology, Engineering, or Math)
* At least 1 year experience working with: Preferred
  + AWS
  + Azure
    - or similar cloud platform
* 5 years’ experience developing commercial solutions in Preferred
  + Python
  + Scala
  + R
* 5 years’ experience developing commercial data science solutions leveraging one or more of the following: Preferred
  + operations research
  + natural language processing
  + machine learning, deep learning
  + video/image analysis
  + time-series analysis.
* At least 3 years of experience mentoring Preferred
  + providing technical governance and oversight
* Strong understanding of:
  + statistics (and applications)
* strong working knowledge of
  + machine learning
  + data mining techniques

**EDUCATION**

* Bachelor’s Degree
* strong quantitative academic background
* BA/BS in a quantitative field
  + Statistics
  + Computer Science
  + Ops Research
* Masters Preferred
  + Statistics
* PhD Preferred
  + Statistics
* MS
  + Computer Science
  + related field
* BS
  + Computer Science
  + related field
  + Math
  + Computer Science
    - or equivalent experience

**DESCRIPTIVE / SOFT SKILLS**

* self-starter
  + Self-driven
* highly motivated
* ability to learn quick
* Team player
* long term vision
* will to take lead / Leader
* quality
* attention to detail \*\*\*
* time management / effectively manage your own time
* interpersonal, written, and verbal communication skills
  + ability to present own results
* Ability to analyze a problem
  + develop and engineer a full solution from concept to deployment to production
* curiosity
* support
  + team
  + other teams
* Innovative
* Business-Minded

**SKILLS**

* attention to detail \*\*\*
* analytical and communication skills
* big data and advanced programming languages Preferred
* statistical
* machine learning
* data mining
* data modeling
* analysis
* oral and written communication skills
* network
  + (ability to) drive collaboration
* Tuning
  + troubleshooting
  + scaling big data
* designing
* architecting
* tuning
* debugging
* collaboration
* communicate
  + driving multiple projects
* attention to detail
* cross functional systems optimization
* solid background in:
  + data science
  + data mining
  + machine learning
* (ability to) analyze
  + (extensive database)
* mine extensive database
* Technical
* Prototyping
* Implementing
  + Solutions
* designing and building dimensional data models
  + to improve:
    - accessibility
    - Efficiency
    - quality of data

**RANDOM and FREQUENTLY SEEN TOPICS**

* attention to detail \*\*\*
* designing and building
* dimensional data
  + models
* Models
* Data pipelines
* data engineering
* attention to detail
* cross functional systems optimization
* motivated
* engineering / engineer
* team work
  + collaboration
* software libraries
  + frameworks
* Analyze
* Driver
* Optimization
* Efficiency
* Evaluating
* exploration
* analysis
* database
  + database management
* analytical
* problem-solving
* attention to detail
* analyze (data)
* forecast
* data management
* data extraction
* data reporting
* acquire data
* business support
* manipulate
* transform
* report
* exploratory analysis
* modeling (decisions)
* data driven business decisions
* communications
* queries
* visualization

**RESUMES**

**words:**

* · data
* · attention to detail
* · analysis
* · problem solving
* · report
* · data manipulation / cleansing / processing
* · raw data
* · recommendations
* · processes
* · stat
* · quantitative/qualitative
* · extract(ion)
* · relational (database)
* · consistency/accuracy
* · (data) integrity
* · reporting
* · (id opportunities)
* · (process development)
* · mining
* · segmentation
* · architect
* · interpersonal skills /comm skills