Big Data: Storage, Analytics and Visualization

Kurra Lavanya (700728999)

Job 1: Data Analyst

Description: The Data Analyst position is situated within the Technology team at MetroNet. Primarily, this position will work designing service area boundaries for new markets, performing QA functions, and carrying out various system administrator duties. This person will work primarily with GIS software and data. They will also work with multiple internal departments as the Analytics Team responds to requests. This individual must possess close attention to detail and a willingness to learn. The Data Analyst position works in a team environment with other Data Analysts, Developers, and Data Processors to help MetroNet continue its phenomenal growth and support the day-to-day business needs of the company.

Requirement: Looking for tech-savvy, innovative thinkers to join our team and help us create what's next!

Skill set Required: Excellent interpersonal, analytical, organizational, prioritization, and problem-solving skills

Skills I have: SQL experience

Job 2: Data Scientist 1, People Analytics - remote

Description: Nordstrom People Analytics team focuses on using organizational research, data science, and data analysis techniques to uncover employee and organizational insights that contribute to Nordstrom's long-term success. The Data Scientist 1 will have hands-on experience with the development of data products, from data ingestion to deployment. They will focus on creating genuine business value through the alignment of technology, data, and business interests. Example projects include helping to identify the right people to recruit, staffing workers at the right time and place, and keeping employees happy and motivated to do their very best. In this role, excellent communication skills and a collaborative attitude will be crucial to effectively partnering with the aligned Technology and Data Science teams.

Requirement: Bachelor's Degree in quantitative field with scientific computing, with an advanced degree (MS or PhD) preferred

Skill set Required:

- Experience manipulating and modeling employee and staffing data, such as attrition, time punch, survey, and ambient unstructured data.
- Experience with ML (regression, classification, forecasting) and ML workflows (cross validation, hyper parameter tuning, hold out testing).
- A strong understanding of research design and methods

- Ability to translate technical work to non-technical audiences
- Excellent verbal and written communication skills and a team-oriented attitude

Skills I have: High proficiency in Python or R (Python preferred), with demonstrated proficiency in SQL. Additionally, experience in GIT is desirable.

Job 3: Data Scientist

Description: Position Requires a Top Secret (TS/SCI) Clearance with a Polygraph.

The customer requires support to perform cyber security-related analysis of data within the customer organization, to baseline systems and identify anomalies across datasets in different repositories.

The Contractor shall consolidate and maintain information in spreadsheets, databases, and other formats, as appropriate about cyber security-related selectors and other data of interest to the customer group.
The Contractor shall collaborate with other customer offices to understand the context of the selectors and activity associated with the data.
The Contractor shall analyze the customer's data to identify baseline patterns of activity and activity that deviates from these patterns.
The Contractor shall search for changes in patterns of activity across the customer group's data.
The Contractor shall search customer systems either directly or via API to identify anomalies or other actionable information related to the customer group's data.
The Contractor shall analyze patterns of activity related to data gathered from external systems to identify potential actionable issues.
The Contractor shall create reports and briefings to notify key customers and partners of potential issues, and recommend remediation.

Requirement: Degree or educational background in computer science, data analysis, data science, information science, statistics, applied mathematics, or a related discipline.

Skill set Required:

- Demonstrated experience performing technical analysis on data.
- Demonstrated experience with development and coding using Python.
- Demonstrated experience connecting information across multiple datasets such as database queries, or API calls.
- Demonstrated experience with networks, to include ability to understand network architecture diagrams, and related information.
- Demonstrated experience with cyber selector data, system logs, and network traffic.

Skills I have: I'm doing Big Data in this semester.

Job 4: **Data Engineer (Analytics)**

Description: Traditional recruiting is broken. If you're looking for a job, or looking to hire, you know it. Our client is fixing it! They're reengineering the entire recruiting process. The world has noticed and rewarded them with over $145 million in financing and a valuation of well over a billion dollars. In other words, they're a tech unicorn. Want in?

Well, they've asked us to provide them with a Data Engineer who can build their own pipelines but also branch out into a more classical analyst role - interfacing with business users, building and analyzing data sets, and presenting the results in a meaningful way. If you're good with data, this is a great role with a company that's going places.

Requirement: The work includes meeting with stakeholders from across the business to determine what they need, gathering the required data, building the necessary analyses, interpreting the results, and presenting your findings to management.

Skill set Required:

- Mandatory **SQL** (Snowflake Preferred)
- Mandatory **Python**
- Important **DBT** (or comparable **ETL**)
- Important **Orchestration** (Airflow, Prefect, Dagster etc.)
- Desirable **Visualization** (Looker or Tableau Preferred)
- Desirable SaaS Industry Marketing Metric Experience

Skills I have: SQL and Python

Job 5: Data Engineer

Description: Swag...everyone loves to receive it

Whether it be welcoming new hires into their dream job, thanking customers for their loyalty, celebrating a community milestone, or engaging event attendees, when done right, swag is the glue that bonds brands together with the people that matter to them. But no one enjoys putting this stuff together...

The $30 Billion branded swag supply chain is a complete mess, filled with too many products, too many intermediaries, too little quality control, and not enough technology.

Swag Up is disrupting the industry by streamlining the end-to-end supply chain and making it available through a single platform utilizing powerful APIs and interfaces (e.g., dashboard). As the fastest growing swag company, we have made great strides towards our mission of eliminating the friction in the swag creation and distribution process. BUT we've just begun to scratch the surface...

And that's where YOU come in...

Requirement: A Data Engineer and/or BI Developer role, ideally in an e-commerce environment

- Strong SQL skills and knowledge of data warehouse principles & visualization tools
- Experience with data manipulation & processing of large & complex data sources on cloud systems
- Experience with data pipeline tools such as dbt and Google Cloud Platform
- Experience with Python, R or another scripting language
- Experience in statistical models & analysis
- Excellent quantitative and analytical with ability to draw conclusions based on data
- Able to work in a fast-paced environment with quick turnaround
- Strong problem-solving skills and ability to provide creative solutions

Skill set Required: SQL, data pipeline tools, Python, R or another scripting language.

Skills I have: Python and R

Skills that I have acquired:

1.Good proficiency in Python

2. Data modelling

3.Learned Pig, Hive and Spark

4.Able to write and execute the queries.

5.Learned interactive mode and batch mode.

6.AWS Cloud computing.

7.Learned cluster creation