# **Name:**

# **Email:**

# **Phone:**

# **Sr. Big Data Analyst (Big Data & Analytics)**

## **Professional Summary**

* Over 10 years of experience as a Technical Business Analyst working on Big Data and analytics projects, bridging business requirements with technical solutions in complex enterprise environments.
* Expertise in gathering, analyzing, and documenting requirements for large-scale data pipelines, analytics workflows, and system integrations, ensuring measurable technical outcomes.
* Skilled in translating business needs into functional and technical specifications for Big Data environments using Hadoop, Hive, Spark, Cassandra, and BigInsights.
* Designed and validated data flow diagrams, process workflows, and source-to-target mapping documents, ensuring clarity in system integration and data lineage.
* Collaborated with engineering and QA teams to review test plans, support UAT cycles, and validate solutions against documented requirements.
* Experienced in regulatory and compliance documentation and maintaining traceability for technical solutions, audits, and validation purposes.
* Supported cross-functional teams to align technical deliverables with business objectives, translating technical outputs into actionable insights for stakeholders.
* Led initiatives to optimize data ingestion, transformation, and reporting pipelines, enhancing accuracy, scalability, and efficiency of analytics solutions.
* Developed requirement traceability matrices (RTMs) to ensure end-to-end alignment of business needs, technical implementation, and QA validation.
* Provided training, guidance, and documentation for junior analysts and cross-functional team members on Big Data analytics processes.
* Participated in data governance, quality, and validation activities to ensure integrity and reliability of enterprise data assets.
* Proficient in SQL, Python, Shell scripting, ETL tools, Tableau, Power BI, and cloud-based Big Data services (AWS Glue, EMR, Redshift, S3).
* Collaborated in Agile/Scrum environments, including backlog grooming, sprint planning, and status reporting using Jira and Confluence.
* Delivered end-to-end analytics solutions, integrating multiple data sources and ensuring technical solutions meet business expectations.
* Recognized for ability to translate complex technical challenges into actionable business recommendations and measurable outcomes.

## **Technical Skills**

| **Category** | **Skills** |
| --- | --- |
| **Big Data Ecosystems** | Hadoop (HDFS, Clusters), Spark, Hive, Cassandra, Hortonworks, BigInsights |
| **Cloud Computing (AWS)** | AWS Glue, AWS EMR, AWS Redshift (Data Warehousing), AWS S3 (Storage) |
| **Databases & Languages** | SQL (Proficient in validation, queries, logic), Python (Scripts, validation), Shell Scripting (Proficient) |
| **ETL & Data Integration** | ETL Tools (General proficiency), Informatica (Explicitly mentioned), Data Ingestion / Transformation |
| **BI & Visualization** | Tableau, Power BI |
| **Methodologies & Tools** | Agile / Scrum, Jira, Confluence, Requirement Traceability Matrices (RTMs), Source-to-Target Mapping |
| **Core Technical Practices** | Data Flow Diagrams, Process Flow Documentation, Data Lineage, Metadata Management, Data Quality / Data Validation |

## **Professional Experience**

### **Client:**

### **Sr. Big Data Analyst / Technical Business Analyst**

**Responsibilities:**

* Partnered with stakeholders to analyze business needs and translate them into technical requirements for Big Data ingestion and analytics projects using Hadoop and Spark environments.
* Designed comprehensive data flow diagrams, process flow documentation, and integration impact assessments to provide visibility into system dependencies.
* Collaborated with engineering teams to implement HDFS, Hive, Spark, and Hortonworks-based pipelines, ensuring scalability and accuracy of data ingestion processes.
* Reviewed test plans, assisted in UAT execution, and validated technical deliverables against functional requirements, ensuring alignment with stakeholder expectations.
* Developed and maintained traceability matrices to track requirements through design, development, and testing phases.
* Participated in cross-functional meetings with QA, product, and solution engineering teams to resolve design and validation issues.
* Led workshops to define data mapping rules, validation criteria, and performance metrics for analytical reporting.
* Coordinated with teams to standardize metadata and capture lineage for all major datasets.
* Validated data ingestion and transformation logic using SQL and Python scripts, ensuring accuracy of datasets for analytics.
* Supported compliance documentation efforts, including capturing data lineage and system validation evidence.
* Delivered requirement walkthroughs and stakeholder sign-offs, ensuring consensus on analytics deliverables.
* Assisted in optimizing ETL processes and resolving performance bottlenecks in large-scale Hadoop clusters.
* Contributed to process improvements by identifying automation opportunities in data validation and reporting.
* Collaborated on dashboards and reporting solutions with Tableau and Power BI, validating metrics and calculation logic.

**Environment:** Hadoop, Hive, Spark, Hortonworks, Cassandra, BigInsights, AWS Glue, EMR, Redshift, S3, SQL, Python, Tableau, Power BI, Jira, Confluence, Agile

### **Client:**

### **Big Data Business Analyst**

**Responsibilities:**

* Facilitated requirement-gathering sessions for complex Big Data ingestion, transformation, and analytics initiatives, converting business needs into technical specifications.
* Defined functional and non-functional requirements, source-to-target mappings, and data lineage documentation to guide engineering teams.
* Partnered with QA teams to review test plans, perform data validation, and support UAT cycles ensuring deliverables matched business expectations.
* Developed system impact assessments, identifying dependencies across multiple pipelines and integration points.
* Collaborated with engineers to implement Informatica, Spark, and SQL-based ETL processes, ensuring reliable data transformation.
* Conducted workshops with stakeholders to define KPIs, metrics, and reporting logic for enterprise-level analytics dashboards.
* Created requirement traceability matrices to maintain alignment between business needs and technical implementation.
* Supported migration of pipelines from on-prem Hadoop clusters to cloud-based Redshift architecture, validating data integrity post-migration.
* Documented detailed business and functional requirements for analytics and reporting solutions.
* Validated dashboards and reports with business teams, ensuring accuracy and completeness of analytical outputs.
* Assisted in data profiling, data quality assessments, and validation of large datasets prior to production deployment.
* Participated in sprint planning, backlog management, and Agile ceremonies to track project progress and deliverables.
* Provided input on process improvements, data governance practices, and metadata standardization.
* Delivered post-production validation and recommendations for enhancement of analytical pipelines.

**Environment:** Hadoop, Hive, Spark, Informatica, Redshift, SQL, Python, Tableau, Power BI, Agile/Scrum, Confluence

### **Client:**

### **Technical Business Analyst**

**Responsibilities:**

* Gathered requirements for enterprise analytics and reporting initiatives, translating them into detailed BRDs, FRDs, and system specifications.
* Partnered with engineering teams to design BigInsights-based data ingestion and transformation pipelines, ensuring accuracy and performance.
* Conducted test plan reviews, UAT support, and validation of dashboards, ensuring alignment with documented requirements.
* Created data flow and process flow diagrams for analytical projects to illustrate data lineage and transformation logic.
* Collaborated with stakeholders to define KPIs and metrics, ensuring that reports reflected accurate business insights.
* Conducted data validation and reconciliation activities using SQL and Python, identifying discrepancies and recommending corrective actions.
* Developed and maintained traceability matrices linking business requirements to system functionality and testing artifacts.
* Assisted in migration and integration projects to modernize analytics platforms.
* Participated in sprint planning, backlog refinement, and Agile ceremonies to ensure timely delivery of requirements.
* Provided documentation and training to stakeholders and team members on analytics processes and workflows.
* Supported process optimization initiatives, improving efficiency and reducing errors in data processing pipelines.
* Collaborated with reporting teams to validate metrics in Tableau and Power BI dashboards.
* Assisted in implementation of data governance practices, ensuring consistency and reliability of datasets.
* Provided input for regulatory and compliance validation, ensuring audit readiness and data traceability.

**Environment:** BigInsights, Hadoop, Hive, Spark, SQL, Python, Tableau, Power BI, Jira, Confluence, Agile

### **Client:**

### **Big Data System Analyst**

**Responsibilities:**

* Gathered and analyzed requirements for data ingestion and transformation pipelines using Hortonworks and Spark technologies.
* Documented business rules, data mapping, and integration impact analyses for complex analytical workflows.
* Partnered with QA and engineering teams to validate ETL processes, dashboards, and reporting outputs.
* Participated in UAT cycles, test case reviews, and validation of system functionality against business requirements.
* Supported creation of traceability matrices and metadata documentation to ensure alignment across multiple projects.
* Collaborated with stakeholders to define data quality metrics, validation rules, and acceptance criteria for deliverables.
* Assisted in migration to cloud-based data storage and analytics solutions, ensuring minimal disruption to existing pipelines.
* Participated in Agile ceremonies to monitor project progress, manage backlogs, and coordinate cross-team activities.
* Created process and data flow documentation to enhance understanding of system interactions and dependencies.
* Provided input for release readiness, validation, and deployment of Big Data solutions.
* Supported training of junior analysts on technical documentation and best practices.
* Collaborated with cross-functional teams to optimize ETL workflows and improve data validation procedures.
* Ensured end-to-end alignment of business requirements, technical implementation, and QA verification.
* Delivered stakeholder presentations and sign-offs for completed analytics deliverables.

**Environment:** Hortonworks, Hive, Spark, SQL, Python, Tableau, Agile, Jira, Confluence

### **Client:**

### **Business Analyst (Data Analytics & Reporting)**

**Responsibilities:**

* Collected and documented requirements for data warehouse and analytics projects, converting business needs into actionable technical specifications.
* Developed functional specifications, source-to-target mappings, and validation matrices to guide ETL development.
* Supported QA testing, UAT cycles, and post-deployment validation to ensure deliverables matched business expectations.
* Created data flow and process flow diagrams to illustrate end-to-end analytics pipelines and system interactions.
* Collaborated with reporting teams to define metrics, KPIs, and calculation logic for dashboards and reports.
* Performed data validation, reconciliation, and quality assurance using SQL and Python scripts.
* Assisted in migration and integration of analytics pipelines across multiple systems and platforms.
* Developed documentation for system changes, requirements traceability, and release readiness.
* Participated in Agile ceremonies, sprint planning, and backlog refinement to ensure timely delivery of analytics solutions.
* Assisted in identifying process improvements and automation opportunities in ETL and reporting workflows.
* Supported training and onboarding of new team members on analytics and documentation practices.
* Delivered post-production validation and recommendations for pipeline optimization.
* Collaborated with cross-functional teams to resolve data discrepancies and technical issues.
* Provided stakeholder presentations and sign-offs for completed analytics projects.

**Environment:** SQL, Python, ETL tools, Tableau, Power BI, Agile, Confluence