

SUMMARY

You are a high-impact Data Analyst with over four years of experience delivering forecasting models, automated analytics pipelines, and executive-level dashboards across healthcare and IT consulting environments. You excel at transforming complex, fragmented datasets into clear, decision-ready insights that support strategic planning and operational excellence.

With strong expertise in Python, SQL, AWS, Apache Spark, and Databricks, you design and maintain scalable ETL workflows and analytics solutions that enable reliable, ML-driven insights. Your technical depth allows you to work efficiently with large, diverse data sources while ensuring data quality, performance, and usability.

You are highly effective at communicating insights through compelling data stories using Power BI and Tableau, supported by advanced statistical analysis. By bridging technical analytics with business needs, you consistently deliver measurable value and enable data-driven decision-making at the leadership level.

SKILLS

Methodologies

1. Scrum

- **What it is:** An agile framework for iterative development and continuous improvement.
- **Application:** Managing analytics and data projects in short, structured sprints.
- **Example:** "I worked in Scrum teams to deliver analytics features incrementally and adapt to changing requirements."

2. Kanban

- **What it is:** A visual workflow management method focused on continuous delivery.
- **Application:** Tracking tasks and optimizing workflow efficiency.
- **Example:** "I used Kanban boards to monitor data pipeline tasks and reduce bottlenecks."

3. Risk Management

- **What it is:** Identifying and mitigating project and operational risks.
- **Application:** Anticipating data quality, delivery, or dependency risks in projects.
- **Example:** "I identified data dependency risks early and planned mitigation strategies."

4. Stakeholder Management

- **What it is:** Coordinating with business and technical stakeholders.
- **Application:** Aligning analytics outcomes with business expectations.
- **Example:** "I regularly communicated insights and progress to stakeholders to ensure alignment."

Programming & Query Languages

1. Python

- **What it is:** A versatile programming language for data analysis and automation.
- **Application:** Data processing, ETL development, and analytics workflows.
- **Example:** "I used Python with Pandas to clean and transform large datasets."

2. R

- **What it is:** A statistical programming language for advanced analytics.
- **Application:** Statistical analysis, modeling, and visualization.
- **Example:** "I used R to perform hypothesis testing and analyze experiment results."

3. SQL

- **What it is:** A language for querying and managing relational databases.
- **Application:** Extracting, transforming, and analyzing structured data.
- **Example:** "I wrote SQL queries to generate monthly KPI reports."

4. SAS

- **What it is:** A statistical software suite for analytics and reporting.
- **Application:** Data analysis and enterprise reporting.
- **Example:** "I used SAS to analyze structured datasets and generate reports."

Analytics & Data Science Packages

1. NumPy

- **What it is:** A library for numerical computing in Python.
- **Application:** Performing mathematical and array-based operations.
- **Example:** "I used NumPy for efficient numerical computations."

2. Pandas

- **What it is:** A Python library for data manipulation and analysis.
- **Application:** Data cleaning and transformation.
- **Example:** "I used Pandas to handle missing values and reshape datasets."

3. Matplotlib

- **What it is:** A plotting library for data visualization.
- **Application:** Creating charts and graphs for analysis.
- **Example:** "I visualized trends using Matplotlib charts."

4. SciPy

- **What it is:** A library for scientific and statistical computing.
- **Application:** Advanced mathematical and statistical operations.
- **Example:** “I used SciPy for statistical computations.”

5. Seaborn

- **What it is:** A visualization library built on Matplotlib.
- **Application:** Creating statistical visualizations.
- **Example:** “I built correlation heatmaps using Seaborn.”

6. Ggplot2

- **What it is:** A data visualization package in R.
- **Application:** Creating layered statistical plots.
- **Example:** “I used Ggplot2 to visualize experiment results.”

Data Visualization & BI Tool

1. Tableau

- **What it is:** A BI tool for interactive dashboards.
- **Application:** Visualizing business KPIs.
- **Example:** “I built Tableau dashboards for executive reporting.”

2. Power BI

- **What it is:** Microsoft’s business intelligence platform.
- **Application:** Creating reports and dashboards.
- **Example:** “I developed Power BI dashboards for operational insights.”

3. Microsoft Excel

- **What it is:** A spreadsheet tool for analysis and reporting.
- **Application:** Data analysis and reporting.
- **Example:** “I used Excel for data analysis and reporting.”

4. QlikView

- **What it is:** A BI and analytics platform.
- **Application:** Interactive data exploration.
- **Example:** “I used QlikView to analyze business metrics.”

5. Google Analytics

- **What it is:** A web analytics tool.

- **Application:** Tracking and analyzing website performance.
- **Example:** “I analyzed user behavior using Google Analytics.”

Machine Learning & AI

1. Generative AI (LLM, RAG)

- **What it is:** AI models for text generation and retrieval-based tasks.
- **Application:** Building intelligent data-driven solutions.
- **Example:** “I worked with LLM-based systems using retrieval-augmented generation.”

2. Machine Learning

- **What it is:** Algorithms that learn patterns from data.
- **Application:** Predictive modeling and classification.
- **Example:** “I built ML models to identify data patterns.”

3. Deep Learning (Neural Networks)

- **What it is:** ML techniques using layered neural networks.
- **Application:** Handling complex data patterns.
- **Example:** “I applied neural networks for predictive tasks.”

4. MLOps (CI/CD)

- **What it is:** Practices for deploying and maintaining ML models.
- **Application:** Automating ML pipelines.
- **Example:** “I supported CI/CD workflows for ML models.”

5. NLP Frameworks

- **What it is:** Tools for processing human language.
- **Application:** Text analysis and language modeling.
- **Example:** “I applied NLP techniques for text analysis.”

6. TensorFlow

- **What it is:** A deep learning framework.
- **Application:** Building and training neural networks.
- **Example:** “I used TensorFlow to develop ML models.”

7. Recommender Systems

- **What it is:** Systems that suggest relevant items to users.
- **Application:** Personalization and recommendation use cases.

- **Example:** “I worked on recommendation logic for user data.”

8. Fraud Detection Models

- **What it is:** Models to identify anomalous or fraudulent behavior.
- **Application:** Risk and anomaly detection.
- **Example:** “I supported fraud detection analysis using ML models.”

Cloud Platforms

1. AWS

- **What it is:** A cloud computing platform.
- **Application:** Data storage, processing, and analytics.
- **Example:** “I used AWS services for data pipelines.”

2. Google Cloud Platform (GCP)

- **What it is:** Google’s cloud infrastructure.
- **Application:** Analytics and data processing.
- **Example:** “I used GCP tools for data analytics.”

3. Microsoft Azure

- **What it is:** Microsoft’s cloud platform.
- **Application:** Data integration and analytics.
- **Example:** “I supported data workflows on Azure.”

Big Data Technologies

1. Hadoop

- **What it is:** A framework for distributed data storage and processing.
- **Application:** Handling large-scale datasets.
- **Example:** “I worked with Hadoop for big data processing.”

2. Hive

- **What it is:** A data warehouse system on Hadoop.
- **Application:** Querying large datasets.
- **Example:** “I used Hive for analytical queries.”

3. Apache Spark

- **What it is:** A distributed data processing engine.
- **Application:** Large-scale data transformation.
- **Example:** “I processed big datasets using Spark.”

4. Databricks

- **What it is:** A unified analytics platform built on Spark.
- **Application:** Data engineering and analytics workflows.
- **Example:** “I used Databricks notebooks for data processing.”

ETL Tools

1. Informatica

- **What it is:** An enterprise ETL tool.
- **Application:** Data integration and transformation.
- **Example:** “I used Informatica for ETL workflows.”

2. Talend

- **What it is:** An open-source ETL and integration platform.
- **Application:** Data extraction and transformation.
- **Example:** “I built ETL jobs using Talend.”

3. Alteryx

- **What it is:** A data preparation and analytics platform.
- **Application:** Data blending and analytics.
- **Example:** “I used Alteryx to prepare analytics datasets.”

Reporting Tools

1. SAS BI

- **What it is:** A business intelligence platform.
- **Application:** Enterprise reporting.
- **Example:** “I generated reports using SAS BI.”

2. Cognos

- **What it is:** IBM’s BI and reporting tool.
- **Application:** Dashboarding and reporting.
- **Example:** “I supported reporting through Cognos.”

3. Business Objects

- **What it is:** SAP's BI and reporting suite.
- **Application:** Enterprise analytics.
- **Example:** "I worked with Business Objects reports."

Databases

1. MySQL

- **What it is:** A relational database system.
- **Application:** Data storage and querying.
- **Example:** "I queried MySQL databases for analytics."

2. SQL Server

- **What it is:** Microsoft's relational database platform.
- **Application:** Transactional and analytical workloads.
- **Example:** "I worked with SQL Server for reporting."

3. MongoDB

- **What it is:** A NoSQL document database.
- **Application:** Handling semi-structured data.
- **Example:** "I stored JSON data in MongoDB."

4. Snowflake

- **What it is:** A cloud data warehouse.
- **Application:** Scalable analytics.
- **Example:** "I used Snowflake for centralized analytics."

Other Professional Skills

1. Data Cleaning

- **What it is:** Preparing data for analysis.
- **Application:** Removing inconsistencies and errors.
- **Example:** "I cleaned raw datasets before analysis."

2. Data Wrangling

- **What it is:** Transforming raw data into usable formats.

- **Application:** Structuring data for analytics.
- **Example:** “I wrangled data from multiple sources.”

3. Critical Thinking

- **What it is:** Analytical problem-solving ability.
- **Application:** Identifying insights and solutions.
- **Example:** “I analyzed data trends to support decisions.”

4. Problem-Solving

- **What it is:** Identifying and resolving issues.
- **Application:** Improving data processes.
- **Example:** “I resolved data quality issues in pipelines.”

5. Communication

- **What it is:** Conveying insights clearly.
- **Application:** Explaining results to stakeholders.
- **Example:** “I communicated insights through dashboards.”

6. Presentation

- **What it is:** Structuring and delivering insights visually.
- **Application:** Executive reporting.
- **Example:** “I presented analytics findings to leadership.”

Project Management

1. JIRA

- **What it is:** An issue and project tracking tool.
- **Application:** Managing tasks and sprints.
- **Example:** “I tracked analytics tasks using JIRA.”

2. Trello

- **What it is:** A visual task management tool.
- **Application:** Workflow organization.
- **Example:** “I organized project tasks in Trello.”

3. MS Project

- **What it is:** A project planning tool.
- **Application:** Scheduling and planning.

- **Example:** “I used MS Project for project timelines.”

4. Budgeting

- **What it is:** Managing project costs.
- **Application:** Resource and cost planning.
- **Example:** “I supported budgeting for analytics projects.”

5. Timeline Management

- **What it is:** Managing project schedules.
- **Application:** Ensuring timely delivery.
- **Example:** “I tracked milestones to meet deadlines.”

6. Cross-functional Collaboration

- **What it is:** Working with diverse teams.
- **Application:** Coordinating across technical and business groups.
- **Example:** “I collaborated with engineers and stakeholders on data initiatives.”

EXPERIENCE

Cardinal Health, CA | Data Analyst | Oct 2023 – Current

1. Built automated demand-forecasting pipelines using Python, SQL, and AWS, improving medical product forecast accuracy by 18% and reducing stock-outs across 40+ distribution centers.

- **What it means:** Automated demand forecasting to help the organization plan inventory more accurately at large scale.
- **Simplified example:** Instead of manually estimating demand, forecasts are generated automatically using historical data.
- **Project Implementation:**
 - Queried historical demand data using SQL
 - Built forecasting logic in Python
 - Deployed scheduled pipelines on AWS
- **Project Interview Script:**
“At Cardinal Health, I built automated demand-forecasting pipelines using Python, SQL, and AWS. This improved forecast accuracy by 18% and reduced stock-outs across more than 40 distribution centers.”

2. Designed and deployed interactive Power BI dashboards for supply chain, regulatory, and finance leaders, cutting reporting turnaround time from days to under 2 hours.

- **What it means:** Enabled leadership teams to access critical metrics quickly without manual reporting delays.
- **Simplified example:** Leaders could open a dashboard and see updated KPIs instead of waiting days for reports.

- **Project Implementation:**
 - Modeled reporting datasets
 - Built interactive Power BI dashboards
 - Configured automated data refreshes
- **Project Interview Script:**
"I designed and deployed Power BI dashboards that reduced reporting turnaround time from days to under two hours for supply chain and finance leaders."

3. Conducted root-cause analysis on shipment delays and order discrepancies using Pandas and Spark, enabling logistics teams to reduce delays by 12%.

- **What it means:** Analyzed large operational datasets to identify why shipments were delayed or orders were incorrect.
- **Simplified example:** Found common patterns causing delays and shared fixes with logistics teams.
- **Project Implementation:**
 - Analyzed shipment and order data using Pandas and Spark
 - Identified trends and anomalies
 - Shared insights with logistics stakeholders
- **Project Interview Script:**
"I used Pandas and Spark to perform root-cause analysis on shipment delays, which helped logistics teams reduce delays by 12%."

4. Led cross-functional initiatives to standardize data quality rules, eliminating 30k+ recurring data errors across product, vendor, and transaction datasets.

- **What it means:** Improved data accuracy and trust by enforcing consistent data quality standards across systems.
- **Simplified example:** Incorrect or incomplete records were flagged before appearing in reports.
- **Project Implementation:**
 - Defined data validation rules
 - Applied checks across multiple datasets
 - Monitored and resolved recurring issues
- **Project Interview Script:**
"I led cross-functional initiatives to standardize data quality rules, eliminating over 30,000 recurring data errors across enterprise datasets."

5. Developed predictive risk-scoring models using Random Forest and XGBoost for high-value SKUs, supporting replenishment prioritization and avoiding multi-million-dollar shortages.

- **What it means:** Used machine learning to identify products at risk of stock-outs and prioritize replenishment decisions.
- **Simplified example:** High-risk products were flagged early so teams could restock before shortages occurred.
- **Project Implementation:**
 - Built and evaluated models using Random Forest and XGBoost
 - Generated risk scores for SKUs
 - Shared results with supply chain teams

- **Project Interview Script:**
"I developed predictive risk-scoring models using Random Forest and XGBoost to prioritize replenishment and help avoid high-value shortages."

6. Implemented automated ETL workflows using Informatica and Talend for regulatory and compliance reporting, reducing manual effort by 40%.

- **What it means:** Automated data pipelines to reduce manual work in compliance and regulatory reporting.
 - **Simplified example:** Reports were generated automatically instead of being manually compiled.
 - **Project Implementation:**
 - Built ETL workflows in Informatica and Talend
 - Scheduled automated data loads
 - Validated outputs for compliance use
 - **Project Interview Script:**
"I implemented automated ETL workflows using Informatica and Talend, reducing manual reporting effort by 40%."
-

Infinite Infolab, India | Data Analyst | Jan 2020 – Feb 2023

1. Delivered end-to-end analytics solutions including data modeling, ETL development, dashboarding, and ML insights for enterprise clients across operations, sales, and finance.

- **What it means:** Owned analytics projects from raw data ingestion to final business insights.
- **Simplified example:** Turned raw client data into dashboards and predictive outputs.
- **Project Implementation:**
 - Performed data modeling and ETL development
 - Built dashboards and analytical reports
 - Delivered ML-driven insights
- **Project Interview Script:**
"At Infinite Infolab, I delivered end-to-end analytics solutions covering data modeling, ETL, dashboards, and ML insights for enterprise clients."

2. Built scalable ETL pipelines using Alteryx, Talend, and SQL databases, reducing data refresh times by 50–70% and stabilizing daily reporting.

- **What it means:** Improved the speed and reliability of enterprise reporting systems.
- **Simplified example:** Reports refreshed faster and consistently without failures.
- **Project Implementation:**
 - Built and optimized ETL pipelines
 - Tuned SQL queries and workflows
 - Stabilized daily reporting jobs
- **Project Interview Script:**
"I built scalable ETL pipelines using Alteryx, Talend, and SQL that reduced data refresh times by 50–70% and stabilized daily reporting."

3. Designed predictive churn, fraud detection, and recommendation models using Python, R, and TensorFlow, improving client KPIs by 10–30%.

- **What it means:** Applied machine learning models to improve customer and operational performance.
- **Simplified example:** Identified customers likely to churn or suspicious transactions.
- **Project Implementation:**
 - Built models using Python, R, and TensorFlow
 - Evaluated model performance
 - Delivered insights to client teams
- **Project Interview Script:**
"I designed churn, fraud detection, and recommendation models that improved client KPIs by 10–30%."

4. Developed executive-ready Tableau and Power BI dashboards consolidating multi-source data into unified business views.

- **What it means:** Enabled leadership to track performance from a single, trusted dashboard.
- **Simplified example:** Executives viewed all key metrics in one place instead of multiple reports.
- **Project Implementation:**
 - Integrated multiple data sources
 - Designed executive dashboards
 - Automated refresh cycles
- **Project Interview Script:**
"I developed executive-ready Tableau and Power BI dashboards that consolidated multi-source data into unified business views."

5. Led data cleaning and standardization initiatives using Pandas, SQL, and regex pipelines, significantly improving data reliability.

- **What it means:** Ensured data accuracy and consistency across analytics workflows.
- **Simplified example:** Clean data led to more reliable reports and models.
- **Project Implementation:**
 - Cleaned and standardized datasets
 - Applied validation checks
 - Verified data before reporting
- **Project Interview Script:**
"I led data cleaning and standardization initiatives that significantly improved data reliability."

6. Collaborated with engineering teams to deploy ML workflows in MLOps environments, reducing model downtime and accelerating release cycles.

- **What it means:** Supported smooth deployment and maintenance of machine learning models.
- **Simplified example:** Models moved from development to production with fewer issues.
- **Project Implementation:**
 - Coordinated with engineering teams
 - Supported ML workflow deployment

- Monitored post-release performance
- **Project Interview Script:**
"I collaborated with engineering teams to deploy ML workflows in MLOps environments, reducing downtime and speeding up release cycles."