

Chance Bates

☎ (970) 555-4186 ✉ chance.bates@example.com 💼 linkedin.com/example/chancebates 📍 Fort Collins, CO 80524

SUMMARY

Data Scientist VI with 9 years of experience delivering **statistical modeling**, **machine learning**, exploratory analysis, **data engineering**, and production data systems. Leads research across large data sets, evaluates new data sources, and converts complex findings into clear recommendations for technical and business stakeholders. Builds validated models, analytical tools, scalable ETL and ELT workflows, data quality controls, and maintainable production services using Python, Scala, R, SQL, Spark, Hive, Airflow, and **AWS Step Functions**. Guides junior data scientists through research design, model review, debugging, and production response. Balances delivery commitments with technical debt, system health, performance, cost, and roadmap priorities. Brings practical judgment, clear visualization, and collaborative communication that can support complex analytical initiatives and lasting business value.

EXPERIENCE

Principal Data Scientist

January 2023 - Present

Alpine Signal Analytics

Fort Collins, Colorado

Leads analytical strategy across customer, product, and operational data. Directs exploratory research, model development, production review, technical debt reduction, and mentorship for four junior and mid level data scientists.

- Led customer data exploration with **Spark** and **SQL**, qualifying new sources for production analytics.
- Built validated modeling workflows in **Python**, improving confidence in executive recommendations and research decisions.
- Translated complex findings into visual narratives for leaders, engineers, analysts, and senior decision makers.
- Guided four data scientists through model reviews, debugging, research design, and production response.
- Replaced fragile analytical jobs with maintainable components, balancing roadmap commitments and **technical debt**.
- Established cloud service review practices with data controls, release checkpoints, and incident follow up.

Senior Data Scientist

March 2020 - December 2022

Summit Ridge Intelligence

Denver, Colorado

Delivered **machine learning** and **data engineering** solutions across behavioral and transactional data. Shaped source discovery, warehouse workflows, pipeline performance, stakeholder recommendations, and production readiness while mentoring developing scientists.

- Developed machine learning solutions with Python, SQL, Spark, and **Hive** for production workflows.
- Profiled acquired data sources, defining transformations and analytical use cases with business stakeholders.
- Created **Airflow ETL** and ELT pipelines across ingestion, scoring, and quality validation stages.
- Optimized **Spark** joins and partitions, reducing runtime and cloud processing costs for large transformations.
- Presented statistical findings through concise visualizations, clarifying uncertainty and tradeoffs for varied audiences.
- Mentored data scientists on experiments, model evaluation, code structure, and production readiness.

Data Scientist

July 2017 - February 2020

Mesa Peak Data Labs

Colorado Springs, Colorado

Developed predictive and descriptive models, repeatable analytical data sets, distributed processing workflows, and operational dashboards. Supported code reviews, data quality, production services, and cloud troubleshooting under senior technical direction.

- Built operational models with **Python**, **R**, SQL, regression, and classification methods for planning.
- Prepared warehouse and event data through reproducible scripts, ETL jobs, and validation checks.
- Applied Spark and Hive when data volumes exceeded practical single node processing limits.
- Created dashboards linking analytical patterns with concrete operating decisions for business partners.
- Detected source and feature anomalies before scheduled analysis through targeted **data quality** checks.
- Supported code reviews and cloud troubleshooting for maintainable analytical services and production workflows.

Junior Data Scientist

June 2015 - June 2017

Clearwater Decision Science

Boulder, Colorado

Performed supervised exploratory research across customer and operations data. Built early models, reusable transformation scripts, validated data sets, and warehouse analyses while developing **independent research** and communication skills under senior review.

- Conducted exploratory research with SQL, Python, and statistical methods across customer operations data.
- Prepared structured modeling data by reconciling missing values, duplicate records, and inconsistent fields.
- Compared regression and classification results under senior review, documenting limitations before presentations.
- Wrote reusable **Python** scripts for recurring transformations and monthly reporting preparation.
- Produced charts and concise findings for mixed technical audiences, incorporating reviewer feedback.
- Traced warehouse and data mart fields, testing queries and documenting lineage for repeatable studies.

PROJECTS

Scalable Data Pipeline Modernization 📅 2024

Reworked analytical ingestion and transformation patterns with Airflow, Spark, SQL, and data quality controls. Created smaller delivery increments that improved maintainability, supported production operations, and gave stakeholders clearer visibility into pipeline health.

Distributed Model Performance Study 📅 2022

Compared Spark partitioning, query joins, and intermediate data designs across large data sets. Findings guided performance and cost decisions for machine learning workflows and helped teams select practical distributed processing patterns.

LEADERSHIP & AWARDS

- Recognized as a technical mentor for guiding four data scientists through research, reviews, debugging, and production response.
- Selected to shape analytical roadmaps by balancing technical debt, delivery tradeoffs, performance, cost, and long term system health.
- Earned leadership trust through clear visual explanations of statistical findings for varied technical and business audiences.

EDUCATION

Bachelor's Degree in Statistics

2015

Colorado State University GPA: 4.0

Fort Collins, Colorado

Coursework: Statistical Modeling, Machine Learning, Data Analysis, Probability

CERTIFICATIONS

- AWS Certified Data Engineer Associate 📅 2025
- AWS Certified Data Engineer Associate, Production Data Services 📅 2025

TECHNICAL SKILLS

- **Programming Languages:** Python, R, Scala
- **Query Languages:** SQL, Spark SQL, HiveQL
- **Distributed Processing:** Apache Spark, Apache Hive, Big Data
- **Workflow Orchestration:** Apache Airflow, AWS Step Functions, Scheduled Workflows
- **Data Engineering:** ETL, ELT, Data Pipelines
- **Data Architecture:** Data Warehouses, Data Marts, Data Modeling
- **Machine Learning:** Classification, Regression, Model Evaluation
- **Statistics:** Statistical Modeling, Advanced Statistics, Probability
- **Cloud Infrastructure:** AWS, Cloud Troubleshooting, Production Operations
- **Data Quality:** Validation Checks, Data Controls, Source Profiling
- **Software Development:** Code Reviews, Design Patterns, Query Optimization
- **Visualization:** Dashboards, Data Visualization, Analytical Presentations

SKILLS

- | | | | |
|----------|------------------------|-----------------------|----------------------|
| • Python | • Scala | • ETL and ELT | • Big Data |
| • SQL | • Machine Learning | • Data Pipelines | • Data Warehouses |
| • Spark | • Statistical Modeling | • Data Modeling | • Data Visualization |
| • Hive | • Advanced Statistics | • Distributed Systems | |

PROFESSIONAL AFFILIATIONS

- Association for Computing Machinery, professional member focused on computing and data practice.
- Data Management Association, professional member focused on data architecture, quality, and governance.

LANGUAGES

- English (Native)
- Spanish (Intermediate)

ADDITIONAL INFORMATION

Work Status : Authorized to work in United States. No sponsorship required.

REFERENCES

AVAILABLE ON REQUEST