

# Summary

---

Data/AI Engineer focused on building analytics and AI systems from scratch for US and Middle East clients. I design and operate cloud data platforms on AWS (Redshift, S3, Glue, Lambda, Athena) and build data products and APIs using Python, SQL, Java, and dbt to power reporting, risk scoring, and market analytics. I work end-to-end from requirements and data modeling to quality frameworks and stakeholder communication, with domain experience in insurance, cannabis analytics, COVID-19 risk modeling, and Dubai real-estate and rental analytics.

## Skills

---

<b>Data Platforms</b>	AWS Redshift, S3, Glue, Lambda, Athena, RDS; GCP data services
<b>Data Engineering</b>	ETL/ELT, dbt, data modeling, dimensional design, data quality frameworks, APIs, CI/CD, orchestration
<b>Programming</b>	Python, SQL, Java, R, Bash
<b>Tools</b>	Git, Docker, Linux, BI/Reporting tools, Excel (advanced macros)
<b>Domains</b>	Insurance, Cannabis analytics, Legal data, Real-estate/rental markets, COVID-19 risk modeling

## Experience

---

### ISHIR / SageSure

Remote

SENIOR DATA ENGINEER (CONSULTANT TO SAGESURE)

Apr 2024 – Jan 2026

- Owned Redshift-based analytics platform for US insurance, building scalable ELT pipelines (dbt, S3, Lambda) and API integrations for high-volume policy and claims data.
- Translated business initiatives into data assets and dimensional models, improving query performance and report delivery times for analytics teams.
- Implemented data validation suites and tests using data APIs, reducing data quality incidents and rework in downstream reports.
- Partnered with BI and product stakeholders to troubleshoot data issues, refine KPIs, and harden the semantic layer for self-service analytics.
- Recognized with Rising Star and Team Excellence awards for delivery quality and collaboration.

### CannaSpyglass

Remote

DATA ENGINEER

Jun 2021 – Mar 2024

- Built data pipelines for cannabis wholesalers, transporters, growers, dispensaries, and Google Maps data across AWS and GCP.
- Scraped and ingested PDFs, spreadsheets, and e-commerce sources using Scrapy, Requests, and Selenium on EC2, Glue, and Lambda, landing curated datasets in RDS and S3.
- Automated data quality checks and troubleshooting workflows, ensuring reliable inputs for BI dashboards and regulatory reporting.
- Collaborated with analytics teams to enhance BI tools and create reusable data assets for market and operational insights.

### AI Squared

Remote

DATA SCIENTIST

May 2020 – May 2021

- Developed Java Spring Boot APIs and statistical models for a COVID-19 risk assessment platform used by over 2M users.
- Created Java bindings for R packages, enabling production-grade statistical workflows and predictive models.
- Designed data funneling strategies on AWS (EC2, Glue, S3, Athena, Lambda, RDS) to support advanced analytics and reporting.
- Worked closely with stakeholders to define KPIs, run deep-dive analysis, and recommend improvements across multiple projects.

### Casepoint

India

ASSISTANT PROJECT MANAGER

2016 – 2019

- Managed and delivered data-heavy projects over \$1M revenue, coordinating with US-based senior stakeholders.
- Designed data strategies including pipeline design and SQL/NoSQL data stores to support legal analytics and reporting.
- Automated reporting and analysis with Excel macros, improving turnaround time and reducing manual effort.
- Drove Agile/Scrum ceremonies, tracked risks, and implemented mitigation plans for data processing workflows.

## Education

---

### Dr. Babasaheb Ambedkar Open University, Ahmedabad

India

BACHELOR OF COMPUTER APPLICATIONS

2016

- Coursework in data structures, databases, and software engineering.

## Certifications

---

2025 IBM Data Engineering Professional Certificate, Coursera

2025 dbt Fundamentals, dbt Labs

## Awards & Recognition

---

Year	<b>Rising Star Award</b> , Recognized for delivery quality and collaboration on SageSure data platform.
Year	<b>Team Excellence Award</b> , Acknowledged for cross-functional teamwork and impact on analytics outcomes.