

# SMITA CHAUDHARI

+1 (602) 813-4841 | [smitamchaudhari@gmail.com](mailto:smitamchaudhari@gmail.com) | [linkedin-smita-chaudhari](https://www.linkedin.com/in/smita-chaudhari)

## SUMMARY

---

Technology professional with experience in data engineering, analytics, and cloud solutions (AWS, Azure). Competent in Python, SQL, Hadoop, Apache Spark and building ETL pipelines. Expertise in data modeling, automation, and solving complex data challenges in industries.

## WORK EXPERIENCE

---

### Arizona State University

Jul 2024 – Present

*Data Engineer*

*Tempe, AZ*

- Designed and implemented a serverless data lake, scalable AWS ELT pipeline using DMS, S3, Lambda, Glue, Redshift, and Step Functions to process transactional and event data, reducing vendor costs by \$45K and enhancing Artificial Intelligence driven business operations.
- Implemented data quality validation and pipeline monitoring using CloudWatch, EventBridge ensuring near real-time Tableau reporting and achieving 99.9% data compliance and availability SLA.
- Partnered with clients and cross-functional teams to establish business needs and provide tailored Tableau solutions improving data accessibility and reducing manual data analysis by 7 hours.

### IBM

Feb 2021 – Jul 2022

*Data Engineer (Application Consultant)*

- Engineered Snowflake data warehousing solutions using Wherescape Red, SQL server, SSMS optimizing ETL workflows and data processing time by 3 hours.
- Optimized ETL and data pipelines, enhancing Apache Airflow reliability and automating data processing, reducing failures and boosting data access, performance by 30%.
- Mentored a team of 5 data Science engineers, increasing operational performance and generating 17 reports to translate complex data insights for non-technical stakeholders.
- Led AWS Cloud & Snowflake migration, reducing downtime by 20% and enhancing scalability, data automation, data modeling, and team productivity.

### Cognizant Technology Solutions

Aug 2018 – Feb 2021

*Big Data Engineer (Programmer Analyst)*

- Pioneered Big Data solutions using Apache Spark, Hadoop, and Hive, enabling real-time/batch analytics and designing 30+ dashboards via Qlik Sense, improving data-driven decision-making.
- Reduced operational errors by 75% by implementing unit testing and configuring Control-M jobs, while integrating JIRA for project management, Agile, SDLC and cutting manual debug time.
- Accelerated corporate decision-making by transforming massive datasets with Hive and Spark, increasing data accuracy across over 10 million records.
- Orchestrated CI/CD pipeline development in Jenkins, partnering with the DevOps team to automate deployment workflows, detect bugs early, and enhance release efficiency.

### Tata Communications

May 2017 – Aug 2018

*Associate Engineer (Data Analyst)*

- Architected ETL workflows integrating structured and unstructured data, while automating reporting workflows with SQL, Excel, and Tableau, reducing manual effort by 40%.
- Conducted data analysis on 1M+ records, optimizing query performance and indexing in MySQL, reducing processing time for large datasets and enabling real-time KPI tracking.
- Applied statistical strategy and machine learning models (Python, TensorFlow) to improve customer segmentation, uncover trends, and resolve 90% of system discrepancies before deployment.

## EDUCATION

---

### Arizona State University

May 2024

*Master of Science in Information Technology*

*Tempe, AZ*

### MIT Pune University

Sep 2017

*Bachelor of Engineering: Electronics and Telecommunication*

*Pune, India*

## TECHNICAL SKILLS

---

**Programming Languages:** Python, SQL, SparkSQL, PySpark, Shell Scripting

**Big Data Technologies:** Apache Spark, Hadoop, Hive, HDFS, Sqoop, Kafka, Snowflake, Databricks, IntelliJ

**Cloud Platforms:** AWS (DMS, Glue, Redshift, RDS, S3, IAM, VPC, CloudWatch, Lambda), Azure (Azure Data Factory, Azure Synapse, Azure Data Lake), Databricks, AI

**Data Orchestration:** Apache Airflow, AWS Step Function

**Data Visualization:** Tableau, Power BI, Excel

**DevOps & Version Control:** Git, Jenkins, Jira, Confluence

**Database Management:** SQL Server, MySQL, Azure SQL, NoSQL, SSIS, SQL Server Management Studio

**Low-code/No-code Platforms:** SnapLogic (used to simplify data pipelines at ASU)

## CERTIFICATES

---

- **AZ-900: Microsoft Azure Fundamentals**
- **DP-200: Azure Data Engineer Associate**
- **IBM Certified Advocate – Cloud v1**
- **AWS Certified Cloud Practitioner Certification**

## PROJECTS

---

### **Dashboard for Udemy Courses Analysis — Python, Tableau**

- Constructed analytical models using Python Bokeh, achieving 95% data accuracy and uncovering key trends in Udemy's course performance.
- Designed Tableau dashboards, detecting pricing anomalies and driving data-driven decisions, increasing user engagement by 20%.
- Automated data extraction and transformation, enhancing reporting efficiency, metric tracking, and business intelligence insight.

## LEADERSHIP ROLES

---

- **Ambassador at ASU** – (Multicultural Communities of Excellence) Led team, organized and coordinated college events authorities and delegated roles to help make the events a success.
- **National Service Scheme (NSS)** – Administered in social service projects, teaching technical knowledge to orphan students.
- **Director of Talent Management at ASU – Indian Student Association** – Leading recruitment, onboarding, and team engagement initiatives for a diverse student organization.
- **Technical Member – Rewriting the Code Community.**