Smit Patel

480-803-4141 • smitpatel30071@gmail.com• linkedin.com/in/smit3062 • github.com/smit30patel

SUMMARY

Graduate student in Data Science, Analytics & Engineering with strong experience in Python, R, and SQL, data-driven research, quality assurance, and statistical analysis. I have experience in managing large, multi-source datasets, building automated ETL pipelines, and translating complex quantitative results into clear, decision-ready insights. I have Strong interest in global and population health research, health systems performance, and applying data science to produce scalable public goods.

EDUCATION

M.S. Data Science, Analytics, & Engineering Graduating 05/2026
Arizona State University, Tempe, AZ 3.56 GPA
B.E. Information Technology 06/2023
Gujarat Technological University, India 8.20/10 GPA
Relevant Courses: Probabilities & Statistics, Data Driven Optimization, Data Processing, Big Data Analysis, Data Structures, Database Management, Artificial Intelligence, Data Science, Machine Learning

SKILLS, CERTIFICATIONS & ACHIEVEMENTS

Python, R, SQL, Statistics, Hypothesis Testing, Data Mining, Data Wrangling, Data Visualization, Exploratory Data Analysis, Database Management, Data Cleaning & Validation, Storytelling, A/B Testing, AWS, Azure, Tableau, Power BI, Airflow, Kafka, Hadoop, Spark, Snowflake, Docker, Matplotlib, Seaborn, analytical storytelling, publication-ready tables and figures, Regression and classification models, clustering, model diagnostics, performance monitoring, Predictive Modeling, Big Data Analysis, Data Analysis and Interpretation

Certifications

Quantium: Data Analyst Job Simulation on Forage 12/2025
IBM: Introduction to Data Engineering on Coursera 06/2022
IBM: Python for Data Science, AI & Development on Coursera 06/2022

Achievements

2 x Hackathon Winner (Opportunity hacks, Sun hacks) 2025
Authored and published literature review synthesizing peer-reviewed research on time series forecasting. 07/2023

PROFESSIONAL EXPERIENCE

Full Stack Intern | **Arth Info soft Pvt. Ltd | Ahmedabad, IN** 01/2023 – 05/2023

- Built and maintained reproducible data pipelines using Python and SQL, processing datasets with 100K+ rows and improving data processing efficiency by approximately 30%.
- Managed and optimized PostgreSQL and Snowflake databases, supporting 5+ recurring analytical requests per week from stakeholders.
- Automated ETL workflows and data QA checks using Airflow, Kafka, and Docker, reducing processing time by approx 40%.
- Supported analytics and machine learning workflows by integrating PyTorch model outputs, running diagnostics, and documenting assumptions to improve result interpretability.

PROJECTS

AI Powered Excursive form evaluation | ASU Capstone 01/2026 – Present

- Developed an end-to-end multimodal ML pipeline for a mobile exercise analysis application, applying Vision-Language Models (VLMs) and Action Quality Assessment (AQA) methods to evaluate exercise form from video data.
- Scraped and curated large-scale exercise video datasets from YouTube, performing data cleaning & validation, preprocessing, and feature extraction (camera angle, pose, temporal, and biomechanical features) to support model training and evaluation.
- Implemented motion tracking and rhythm-based center-of-mass analytics, integrating AI inference outputs into real-time, interpretable feedback for end users.

Retail Store Optimization Analysis | Quantium 11/2025 – 12/2025

- Developed and validated standardized performance indicators from 10,000+ transaction-level records using data cleaning and transformation protocols.
- Designed and executed a quasi-experimental A/B analysis comparing 2 intervention stores against control locations, identifying a 5–8% improvement in performance indicators.
- Created reproducible visualizations and summary tables that informed executive recommendations for rollout across additional locations.