

Dhruvil Panchal

West Palm Beach, FL

+1-407-494-1791 | dpanchal.dp.2005@gmail.com | <https://www.linkedin.com/in/dhruvilpanchal/>
<https://github.com/DhruvilPanchal205?tab=repositories> |

SUMMARY

Experienced Data Engineer with 5+ years building scalable, reliable data pipelines and cloud architectures using Python, Airflow, Spark, AWS, and Snowflake. Skilled in automating ETL/ELT workflows, optimizing big data processes, and enabling data-driven insights. Proven ability to collaborate across teams and deliver solutions that boost efficiency and reduce costs. Passionate about clean, maintainable code and continuous improvement.

SKILLS

- **Programming Languages:** C#, Python, R, JavaScript
- **Databases & Data Modeling:** SQL, MySQL, RDBMS, PostgreSQL, SQL Server, Data architecture, NoSQL
- **Python API & Web Framework:** PyQt5, Django, PySpark
- **Cloud Platforms:** AWS, Azure, Snowflake, Databricks, Kubernetes
- **Machine Learning:** Numpy, Scikit, Pandas, Linear Regression, Statistical Analysis
- **Big Data:** Hadoop, Spark, dbt
- **Other:** Jupyter Notebook, GitHub, JIRA, Confluence, Trello, MS Office, Apache Airflow, CI/CD, AI/ML Knowledge, Terraform

EXPERIENCE

ThirthaSoft, LLC

Jul 2023 - Present

Data Engineer I

Marietta, GA

- Designed and implemented a scalable data ingestion factory processing 8,000+ daily historical records using Apache Airflow DAGs, ensuring reliable orchestration, fault tolerance, and traceable data workflows.
- Built modular ELT pipelines using Python, YAML, and Spark, enabling ingestion from structured and semi-structured sources (CSV, JSON, XML, REST APIs), while reducing onboarding time by 60%.
- Deployed pipelines to Snowflake using CI/CD practices with GitHub Actions, Docker, and AWS, automating version-controlled releases and improving deployment efficiency by 40%.
- Architected cloud-native data layers with data lakes and Snowflake warehouses, embedding partitioning, schema enforcement, and role-based access controls (RBAC) to ensure compliance and scalable analytics delivery.

ITHENA

Jun 2022 - Aug 2022

Data Engineer Intern

Glen Allen, VA

- Engineered secure ETL workflows using Python and AWS services (S3, Lambda, Athena) to process 100K+ daily IAM logs, enabling real-time RBAC role reconciliation and anomaly detection.
- Consolidated over 500K RBAC records into a centralized AWS Lake Formation data lake, enhancing data transparency, accessibility, and audit readiness across distributed systems.
- Supported financial and security compliance by mapping RBAC data flows, identifying 1,200+ access control issues, and collaborating with internal teams to ensure 100% audit compliance.
- Architected a scalable RBAC data mart supporting 15+ business units, reducing role conflict incidents by 30% and enabling secure, cross-team access analytics for governance and reporting.

BrainyBeam Technologies Pvt. Ltd.

May 2019 - Jun 2021

Jr. Data Engineer

Ahmedabad, IND

- Designed and developed complex SQL queries, triggers, and stored procedures to support ETL workflows and data validation processes across enterprise systems managing over 1 million records.
- Optimized query performance by 30% through indexing strategies, execution plan analysis, and database refactoring in SQL Server / PostgreSQL environments.
- Collaborated with cross-functional teams to implement data pipelines using Python and ETL tools like Talend and SSIS (SQL Server Integration Services).
- Maintained and improved data models (star/snowflake schemas) for reporting layers in Tableau and Power BI, enabling faster business decision-making.

EDUCATION

Cleveland State University

May 2023

MS, Information Systems

PROJECTS

US Finance Complaints Analysis for Risk and Product Strategy

- Developed and deployed machine learning models (e.g., Logistic Regression, Random Forest) to classify consumer finance complaints by issue type and risk category, achieving 85%+ accuracy and enabling proactive product and compliance strategy improvements.

Enhancing Adidas US Sales Strategy Through Product Insights

- Performed customer segmentation using K-Means clustering on Adidas US sales data with Pandas, NumPy, and Matplotlib, uncovering key product performance trends and generating actionable insights to inform regional sales and marketing strategies.

CERTIFICATIONS

- **Hands-On Essentials: Data Warehousing Workshop: Snowflake®**