

Dev Patel

(609)-598 3786 | pateldevp01@gmail.com | www.devpatel.com | [LinkedIn](#) | [Github](#) | New Jersey, USA

PROFESSIONAL SUMMARY

Seasoned Data Engineering and Analytics professional with over 3 years of experience delivering business-focused data solutions. Skilled in translating complex data challenges into clear and actionable insights by working hands-on with stakeholders to gather requirements and communicate technical concepts effectively. Having experience across the entire data analytics and information lifecycle, while building robust ETL pipelines and scalable, distributed data architectures. Proven success developing impactful solutions in the healthcare (revenue cycle management), AI sales prospecting, and retail domains. Highly proficient in writing and optimizing SQL queries and PySpark transformations for reliable and high-quality data pipelines.

TECHNICAL SKILLS

Languages: Python, SQL

Data Engineering and Analytics: Apache Spark (PySpark, Kafka Streaming), PostgreSQL, MySQL, Pandas, Numpy, ETL, Data Warehousing, Data Modeling, Excel (indexing, data manipulation, VLOOKUP)

Cloud & Tools: FastAPI, RestAPI, Postman, AWS (S3, EC2, EMR, Lambda, Athena), Azure (ADLS Gen2, ADF, Key Vault, Unity Catalog, SQL DB, CSV, Delta Lake, JSON, Parquet), Snowflake, Redis, DataBricks, Github, Jira, Confluence

Databases: Microsoft SQL Server, MySQL, PostgreSQL, Redis

EXPERIENCE

Data Analyst

03/2025 – Present

Aurobindo Pharma USA, Inc.

Dayton, NJ

- Ingested ICD-10, CPT, and NPI code data into Azure Data Lake Storage Gen2 using REST APIs, handling JSON and CSV formats for downstream processing and analytics.
- Developed and maintained 10+ ETL processes using Azure Data Factory to ensure accurate, reliable data transfers across diverse sources and systems.
- Built advanced transformations layer in ADF and Databricks with data cleaning, enrichment, quarantine logic (isolating 5% low-quality records), Common Data Model alignment, and SCD Type 2 handling, boosting data quality score by 40%.
- Designed and implemented a comprehensive audit logging system using Delta tables to capture 100% of pipeline activity for compliance and end-to-end data lineage tracking.
- Converted sequential ADF pipelines to parallel execution, reducing end-to-end processing time from 8 hours to under 2 hours and improving overall pipeline efficiency.

Data Engineer

11/2024 – 03/2025

Flow AI

Austin, TX (Remote)

- Led design and execution of ETL pipeline processing 100K+ email batches, reducing extraction latency by 40% and improving customer engagement analytics.
- Led a team of 3 engineers in migrating data from source systems to Amazon S3 and optimizing transformation logic with PySpark and SQL, achieving an overall 75% reduction in latency.

Software Engineer

07/2022 – 05/2024

Frodoe LLC

Cranbury, NJ

- Built a scalable real-time messaging platform using WebSockets & Redis, supporting 10K+ concurrent users.
- Integrated Apache Kafka for high-throughput data streaming, reducing query load by 35% and downtime by 50%, with SQL-backed fault-tolerant storage.

Software Engineer Trainee

03/2022 – 06/2022

Paul Mason Consulting

Vadodara, GJ

- Oversaw 150+ web portals streaming real-time data, ensuring optimal performance and data integrity.
- Created an end-to-end automated pipeline for extracting, monitoring and analyzing data using client's portals.

EDUCATION

Johns Hopkins University

Baltimore, MD (Remote)

MS, Computer Science

2024 - Present

Maharaja Sayajirao University of Baroda

Gujarat, IN

Bachelors of Computer Applications(Honors)

2019 - 2022

CERTIFICATIONS

Bloomberg Market Concepts | *Bloomberg*

08/2023 - Present

AWS Certified Machine Learning Specialty 2023 (MLS-C01) | *Udemy*

01/2023 - 05/2023