

DHANESH SARPALE

Senior Data Engineer | Full Stack Data Engineering

Dublin, Ireland | +353 892690765 | dhanesh.sarpale@gmail.com | Stamp 4 (Applied)

PROFESSIONAL SUMMARY

Senior Data Engineer with 7+ years of experience designing and building scalable data platforms across banking, healthcare, and global technology organizations including Bank of Ireland, Meta, and Amgen. Expert in cloud-native data engineering, distributed data processing (PySpark), and building robust ETL/ELT pipelines on Azure, AWS, and Cloudera. Proven track record in mentoring engineering teams, driving end-to-end data initiatives, optimizing data workflows, and enabling data-driven decision-making with a strong focus on business impact, scalability, and engineering excellence.

CORE TECHNICAL SKILLS

Programming Languages: Python, SQL, PySpark

Cloud Platforms: Microsoft Azure, AWS, Cloudera

Cloud Storage: Azure Data Lake Storage Gen2 (ADLSv2), Azure Blob Storage, AWS S3, Amazon Redshift, On-Prem Cloudera

Data Processing & Pipelines: Azure Data Factory (ADF), Azure Databricks (ADB), Cloudera Data Science Workbench, PySpark, FastAPI

Databases: Oracle, Presto, Azure SQL DB, Teradata, Delta Lake

ETL / Orchestration: Azure Data Factory, Oozie Workflow

BI & Visualization: Power BI, Tableau

CI/CD & DevOps: Jenkins, Git, Azure DevOps, Docker

Testing Frameworks: Pytest, Zephyr

Project Management: Jira, Confluence, Agile / Scrum

PROFESSIONAL EXPERIENCE

Data Engineer III | Bank of Ireland

Sept 2024 - Present | Dublin, Ireland

- Lead and mentor a team of data engineers, driving design, development, and delivery of scalable data solutions within enterprise risk systems.
- Design and build enterprise-scale data lake pipelines ingesting data from Teradata, Oracle, and flat files; transform data using PySpark and deliver curated datasets for reporting and analytics.
- Develop and optimize ETL/ELT workflows, improving data processing performance, reliability, and scalability across critical banking systems.
- Drive data engineering initiatives within Enterprise Risk platforms, enabling credit risk score generation and supporting regulatory and business reporting requirements.
- Build and maintain API-based data services using FastAPI for near real-time (online) data processing and seamless downstream integration.
- Establish and enforce data quality and testing frameworks using Pytest and Zephyr, ensuring pipeline reliability and consistency.
- Implement and manage CI/CD pipelines using Jenkins and Git, automating deployments and improving release cycles.
- Orchestrate complex workflows using Oozie, ensuring robust scheduling and dependency management.
- Collaborate with cross-functional stakeholders (risk, analytics, engineering) to deliver high-impact, production-grade data solutions within Agile/Scrum frameworks.

Data Engineer | Amgen

June 2023 - Aug 2024 | Dublin, Ireland

- Orchestrated data extraction and transformation from diverse sources (Legacy Systems, Workspace One) using Azure Data Factory for ingestion and Azure Databricks for filtering and aggregation.

- Established a centralized Delta Lake in Databricks as a foundational repository for downstream data consumption, ensuring consistency and integrity across analytics applications.
- Integrated MLflow with Delta Lake for efficient tracking and management of machine learning models.
- Implemented robust CI/CD practices through Azure DevOps, automating deployment of data pipelines and minimizing manual errors.
- Led creation of 3 Power BI dashboards providing insightful analysis and visualization of acquisition work processes across European sites post Amgen's acquisition of Horizon.

Data Engineer | Meta (Facebook)

Jan 2023 – June 2023 | Dublin, Ireland

- Collaborated with Analytics Manager to deliver impactful Tableau dashboards, transforming complex data into actionable insights for executive leadership.
- Created 13+ custom views and 85+ subscriptions, delivering weekly, monthly, and quarterly performance metrics to senior stakeholders.
- Implemented performance tuning and optimization techniques on Tableau workbooks, achieving efficiency improvements of up to 45%.
- Monitored and administered Tableau Server including user/group management, security, permissions, and resource optimization.

Cloud Data Engineer | Aledha Technologies LLP

July 2018 – Sept 2022 | India

- Managed batch and real-time data ingestion using Azure Data Factory and Azure Event Hubs.
- Developed and maintained ETL/ELT processes, reducing data processing time by 20% using Azure SQL Database, ADLSv2, and Azure Blob Storage.
- Orchestrated data workflows using Azure Data Factory for seamless scheduling and coordination of data processing tasks.
- Implemented comprehensive testing strategies to ensure accuracy, reliability, and quality of data pipelines.
- Implemented and optimized CI/CD pipelines in Azure DevOps, automating deployment from development to production.
- Developed 8+ interactive Power BI dashboards, improving decision-making efficiency by 10%; presented periodic performance reports for strategic and tactical initiatives.
- Participated in all Agile development activities including requirements gathering, solution design, report development, testing, and release.

EDUCATION

MSc in Business Analytics | Dublin Business School, Ireland | Sept 2022

PG Diploma in Data Science | Symbiosis International University | Sept 2021

CERTIFICATIONS

- Azure Data Engineer Associate (DP-203) — Microsoft
- Cloudera Data Engineer — Cloudera
- Tableau Desktop Specialist — Tableau
- Data Science and Big Data Analytics — Dell EMC
- Python Programming — Microsoft

PUBLICATIONS

Snake Species Recognition using Machine Learning (TensorFlow) — Published in Elsevier