

# RUPESH JADHAV

Data Engineer | Cognizant Pvt. Ltd.



+91 9284248602



rupeshjadhav.works4@gmail.com

## OBJECTIVE

Dedicated and results-oriented professional with experience in building scalable data pipelines, designing efficient data architectures, and delivering actionable insights. Pursuing a challenging role to leverage expertise in ETL development, Big Data technologies, and cloud platforms to optimize data-driven decision-making processes and contribute to organizational success.

## EDUCATION

April 2022

**B.Sc. in Computer Science** from **Pune University, Maharashtra**

## CORE COMPETENCIES

- Data Analysis & Visualization
- Database Management
- Data Modeling
- Version Control
- Data Migration
- Data Integration
- Data Warehousing
- Business Intelligence
- Data Cleaning & Preprocessing
- Predictive Analytics
- Performance Optimization
- Data Governance
- Data Quality Assurance
- Statistical Analysis

## Certification

- Azure Databricks and Spark for Data Engineers (PySpark / SQL) from Udemy

## SOFT SKILLS

- Collaborator
- Communicator
- Numerical Competence
- Problem Solver
- Time Management
- Team-oriented

## PROFILE SUMMARY

- Possess **over 3 years** of experience in **data engineering and analytics**, focusing on the intricate processes of **data extraction, transformation, and migration**, particularly within the technology industry.
- Currently working as a Data Engineer, **collaborating with cross-functional teams, understanding business requirements, and delivering data-driven solutions** that align with organizational goals.
- Cultivated a diverse skill set that includes **data processing, cloud computing**, and various analytical methodologies, particularly focusing on the application of **Python and SQL** for effective **data manipulation and analysis**.
- Engineered **a series of interactive dashboards and comprehensive reports** through **Power BI**, proficiently conveying insights and trends to stakeholders, which in turn promoted informed decision-making.
- Possess extensive expertise in AWS services such as **Glue and Lambda**, facilitating the efficient execution of data transformation and migration tasks within cloud environments.
- Proficient in **designing, developing, and maintaining robust data pipelines and ETL processes** to enable efficient data integration and analysis.
- Strong background in **transforming raw data into actionable insights** using advanced analytics, statistical techniques, and data visualization tools.
- Knowledge of **data quality, security, and compliance standards** to ensure accurate and reliable reporting.
- Committed to **staying updated with emerging technologies and industry best practices** to drive innovation in data engineering and analytics.

## TECHNICAL SKILLS

- **Programming Languages:** SQL, Python
- **Data Science Libraries:** Pandas, Matplotlib, Numpy, Seaborn
- **Data Analysis and Tools:** Jupyter Notebook, Anaconda, Microsoft Excel
- **Cloud Services and Platforms:** AWS (S3, Lambda, Glue, Athena, Redshift), Microsoft Azure (ADF, Databricks, DMS, Azure Database)
- **Data Visualization Tools:** Power BI, Tableau
- **Big Data Technologies:** Apache Spark (PySpark, Spark SQL)
- **Databases:** MySQL, SQL Server
- **Project Management and Collaboration Tools:** Azure DevOps, Azure Boards, Agile, Scrum, Jira
- **Operating Systems and Networking:** Windows, Linux

## ACHIEVEMENTS

- Reduced report generation time by 60% by optimizing SQL queries.
- Built Python scripts using Pandas to process millions of records.
- Optimized PySpark jobs by implementing caching and broadcast joins, reducing ETL runtime by 40%.

## ORGANIZATIONAL PROJECTS

- **Xerox X-Data Factory | Xerox, USA | Industrial | Current Project**
  - Xerox Holdings Corporation is a global provider of print and digital document solutions in over 160 countries.
- **Macy's Analytics Solution | Macy's, USA | E-commerce | Previous Project**
  - Macy's, a US-based high-end department store chain, operates a leading eCommerce platform offering fashion, accessories, beauty products, and home goods.
- **23 Digitizing Grid | UnitedHealth Group (UNH), USA | Healthcare | Previous Project**
  - UnitedHealth Group is the largest health insurance provider, offering healthcare products and services aimed at improving health outcomes globally.

## WORK EXPERIENCE

Jan 2022 – Present | Data Engineer | Cognizant Pvt. Ltd., Pune

Key Result Areas:

- Developing scripts for collecting, cleaning, and preprocessing large datasets using Python libraries like Pandas and NumPy.
- Conducting exploratory data analysis (EDA) to identify trends, patterns, and anomalies in datasets.
- Designing and creating interactive dashboards and reports to communicate actionable insights using tools like Power BI and Excel.
- Collaborating with cross-functional teams to implement data-driven solutions, addressing complex business challenges.
- Creating optimized data visualization solutions for presenting findings to clients, ensuring clarity and relevance.
- Extracting and transforming data from diverse sources, including SQL Server and spreadsheets, into data lakes (e.g., S3) using AWS Glue and cloud services.
- Monitoring AWS Lambda events via CloudWatch, implementing changes to enhance functionality and resolve issues.
- Applying advanced transformations in AWS Glue, including schema modifications, filtering, joins, aggregations, and managing missing data.
- Utilizing Spark and Python for writing transformation scripts, ensuring efficient data processing and delivery.
- Delivering a Proof of Concept (POC) for an automated data pipeline, showcasing optimized design strategies.
- Designing and implementing ETL pipelines in Azure Databricks, performing data extraction, transformation, and loading from various databases and files.
- Leveraging AWS and Azure services for scalable and reliable data migration and processing workflow.
- Analyzing enterprise databases to improve product development, marketing strategies, and operational efficiencies.
- Transforming raw data into actionable insights using Python (Pandas) and visualizing findings with tools like Matplotlib.
- Recommending workflow optimizations by exploring alternative technologies and approaches.
- Participating in Scrum meetings to track progress, discuss milestones, and resolve challenges.
- Supporting project planning, strategy development, and code/ design analysis to ensure smooth delivery.
- Attending project review meetings to discuss milestones and ensure alignment with requirements.
- Collaborating with architects to design scalable, high-performance data architectures that meet business needs.
- Integrating third-party APIs for data ingestion and processing to expand data sources and enrich analysis.
- Building and maintaining a data catalog to enhance discoverability and usability of data assets across the organization.
- Implementing version control strategies for datasets and data models to ensure reproducibility and traceability.
- Optimizing query performance for large datasets and ensuring minimal latency for data access.
- Working closely with software engineers to integrate data solutions into applications and services.
- Conducting in-depth exploratory analyses to uncover patterns, trends, and anomalies in data.
- Continuously improving dashboards and reporting tools for enhanced usability and efficiency.
- Developing segmentation models to identify target audiences for marketing and business strategies.

## PERSONAL DETAILS

Address : Pune, MH

Languages Known : English, Marathi, and Hindi