

Yogita Darade

Data Engineer

(832) 965-7295 | Houston, Texas | yogitap.darade@gmail.com | [GitHub](#) | [LinkedIn](#) | [Kaggle](#)

Performance-driven professional with 10+ years of experience in Software Industry, with 1 year in data engineering, and data science. Seeking opportunities to build data intensive application, tackle challenging architectural and business problems using Data Engineering and Data Science.

SKILLS

Languages : Python (NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn), PySpark, SQL, PL/SQL, T-SQL
Microsoft Azure : DataBricks, Key vault, Data Lake, Synapse, Data Factory, Logic Apps, Functions, Devops, Azure ML
Database\Data Warehouse: Oracle, MS-SQL server, Azure Synapse
Web Scrapping : BeautifulSoup, Selenium
Frameworks and Tools : Git, Jupyter, Google Colab, VSCode, PowerBI
Machine Learning : Regression, Classification, Clustering

WORK EXPERIENCE

Data Engineer Intern (Cubic Corporation)

November 2021 – Current

“Data Management and Analytics Platform (DMAP) is an eco-system that enables a blueprint for democratizing data and leverage the principle of domain-oriented data ownership in form of information and analytical modules to empower successful decision-making and initiatives through actionable insights, transform operations, open revenue channels, and achieve competitive advantages”.

- Conceptualized and developed a data ingestion framework (real-time, batch and micro-batch) for the data lake using modular PySpark functions. This has been rolled out as a key feature for the DMAP product and the ability to scale and support 20+ transit agencies with minimal changes reducing ~90% of development efforts.
- Drafted document for Dimensional modeling (Kimball), source to target mapping and business matrix for the data warehouse.
- Feasibility analysis between the data lake house and data warehouse.
- Reviewed the high- level customer’s business and contract requirements.
- Analyzed and documented high-level approach for General Transit Feed Specification (GTFS), data ingestion and processing.
- Designed the data model for facts and dimensions in the data warehouse.
- Authored design documents translating functional and business requirements.
- Designed and developed the data pipeline (ELT) integrating millions of raw records from 5+ data sources using Azure Functions, Azure Data bricks, Data Factory, Azure Synapse Analytics, and SQL stored procedures.
- Designed the scheduling and orchestration of the data pipeline using Logic Apps.
- Designed and developed the dashboard for service quality analytics and ridership in Power BI.
- Conducted unit, integration, and system testing.
- Volumetric analysis and cost assessment of Azure components and services.
- Automated deployment process by building a CI/CD pipeline using Azure DevOps.
- Configuration and Release Management for Azure Subscription (Dev and QA).
- Conducted show and tell sessions for the onshore and offshore team in US, UK, AU and India.
- Consulting with customers for workshops on business requirement and sprint demos.
- Engaged with business, SME and IT stakeholders for project planning, costs, estimates and sprint.

Data Science Intern (Women’s in Data)

August 2021-October 2021

- Collected data from various sites and web scraped data from Glassdoor for the City of Rancho Cordova.
- Analyzed and visualized data using Tableau to provide insights on thriving Industries, occupations, job opportunities, and current skills requirements.

- The purpose of these insights is to help the Economic Department of the city to build a skilled workforce that spurs business competitiveness and economic growth.
- Used Python to analyze data to find outliers, inconsistent formats, and other data quality items that would impact the data visualizations.

Career break

2015-2021

- Relocated to USA and didn't have permit to work

Associate Consultant (Tata Consultancy Services, India)

2013– 2014

- Led and directed a team of 14 people on critical tasks in running Risk Management (RMS), Admin (AWS), and IPO modules.
- Implemented and supported 8 of India's renowned clients in Trading and Brokerage systems
- Authored development plans, priorities of work and distributed the same among team members.
- Conducted training and peer-mentoring programs for 40+ new hires reducing onboarding time by 40%.

Assistant Consultant (Tata Consultancy Services, India)

2010 – 2013

- Supervised a team of 10 people on critical tasks in running Risk Management (RMS), Admin (AWS)
- Compiled user data to understand technical issues, and modified software to fix bugs and enhance the interface, improving user satisfaction by ~30%
- Managed a team of 10 people and coordinated with the client towards the successful launch of a new product feature of Systematic Investment Plans (SIPs) in the trading applications.
- Collaborated and Communicated with various Stakeholders, the Back office, and Third parties.
- Designed database models for Online Trading Applications without impacting existing product.
- Migrated a product from VB to Vb.net and Oracle 9i to Oracle 10g database.

IT Analyst (Tata Consultancy Services, India)

2008 – 2010

- Coordinated with cross- functional team of 8 people to implement new Advance Trading feature in product.
- Reviewed test cases, and test scenarios to ensure the completeness & coverage of testing to increase efficiency by ~25%.
- Finalized the Business document with the clients, designed screens and workflows as per the requirement.
- Carried integrated testing across all modules reducing client's test run times by 50 percent..

Assistant System Engineer (Tata Consultancy Services, India)

2003 – 2008

- Devised complex SQL queries to move and transform data from operational data sources to reports.
- Tuned performance of applications by ~20%, analyzed solutions ,and coded fixes for software problems.
- Created 30 + screen for Windows Application in Vb and Vb.net.

EDUCATION

University of Texas, Austin, Texas

September 2021

Post-Graduation in Data Science and Business analytics

S.N.D.T University, Mumbai, India

Bachelor of Technology in Computer Science and Technology

CERTIFICATES

- Microsoft Certified Azure Data Engineer
- Microsoft Certified Azure Fundamentals
- Microsoft Certified Azure Data Fundamentals
- Oracle PL/SQL Developer Certified Associate

ACHIEVEMENTS

- Promoted twice ahead of schedule.
- Ranked 1st in my Post Graduation cohort of 80 people.
- Ranked in Top 4% in 30 Day ML Kaggle Competition.