

Choudhury Durga Prasad

Pathapatnam, Andhra Pradesh, India



durgaprasadchowdury@gmail.com



+919502192674



[linkedin.com/in/durga-prasad-](https://www.linkedin.com/in/durga-prasad-choudhury-75022314b)

[choudhury-75022314b](https://www.linkedin.com/in/durga-prasad-choudhury-75022314b)

Summary

Results-oriented Data Engineer with over 4+ years of experience delivering scalable, cloud-native data solutions across AWS, Microsoft Azure, and Google Cloud Platform (GCP). Certified in all three platforms, I specialize in PySpark, DBT, Snowflake, Python, and SQL to design, build, and optimize data pipelines that drive real-time analytics and data-driven decisions. Proven track record in delivering real-time data solutions and actionable insights. Seeking to leverage expertise to drive impactful data-centric solutions.

Professional Experience



Data Engineer II / Analytical Data Engineer

Softborne Technology Solutions Pte. Ltd.

Sep 2023 – Present

- Collaborated with the product management team to identify and prioritize new feature opportunities, leading to timely implementation.
- Designed and optimized scalable data models in Reactor Data using Snowflake and Big Query, improving query performance and reducing processing time.
- Designed and maintained real-time data pipelines to ensure seamless data availability in dashboards and analytics.
- Utilized Domain Specific Language (DSL) for accurately mapping source data to destination schemas.
- Used Source Links to join and map multiple data sources into unified schemas.
- Tuned complex SQL queries in Snowflake and Big Query to optimize performance and manage large-scale retail datasets.
- Worked closely with cross-functional teams to implement tailored data solutions.
- Conducted thorough validation and reconciliation checks between source systems and data warehouse to ensure 99%+ data accuracy.
- Optimized query performance and data workflows to efficiently handle large-scale retail and e-commerce datasets.
- Conducted performance tuning and optimization of database queries.
- Developed dynamic dashboards to track KPIs such as sales performance and customer lifetime value.
- Created and maintained transformation logic using DBT, promoting modular, testable, and reusable data models.



Data Engineer I

KPI Partners

Jul 2022 – Jul 2023 (12 months)

- Created data pipeline using Apache Airflow, Python, and AWS to store weekly raw data.
- Designed a multi-user, multi-database app for processing thousands of records.
- Managed SQL and NoSQL data modelling using Big Query, Snowflake, and Amazon DynamoDB for scalable data collection and structured storage.
- Collaborated with cross-functional teams for pipeline design and business needs.
- Developed scalable data ingestion processes and optimized ETL workflows.
- Implemented data quality checks and monitoring mechanisms for data integrity.



Associate Professional Software Engineer

DXC Technology

Aug 2021 - Jun 2022 (11 months)

- Developed and maintained scalable data pipelines using Azure Data Factory and Databricks.
- Optimized data storage and retrieval using Azure Synapse Analytics and Delta Lake.
- Implemented data security best practices with Azure Key Vault and Role-Based Access Control (RBAC).
- Monitored and troubleshooted data pipeline failures using Azure Monitor and Log Analytics.
- Designed and implemented data models for efficient querying and reporting.
- Automated ETL processes to streamline data ingestion and transformation.
- Collaborated with data analysts and stakeholders to ensure data accuracy and consistency.
- Conducted performance tuning of SQL queries and Spark jobs for optimal execution.



Cyber Security Intern

Cyber Crime Cell, Gurugram Police

Jun 2021 - Jul 2021 (2 months)

- Provide first-level compliance monitoring and investigations. Assist with forensics analysis and fact gathering. Assist with vulnerability assessments and penetration testing for specific applications, services, networks, and servers as required.



Google Cloud Ready Facilitator

Google

Apr 2021 - Jun 2021 (3 months)

- The Google Cloud Ready Facilitator program will introduce you to Computing, Application Development, Big Data & Machine Learning using Google Cloud's training platform called Qwiklabs.
- I learned each of these topics using self-paced labs that provide you temporary credentials to Google Cloud Platform, so you can learn the cloud using the Real thing – no simulations



Project Trainee

DXC Technology

Feb 2021 - Apr 2021 (3 months)

- The project aims to analyze and identify the data quality issues and strategies to mitigate these Issues in the given data. We need to start with a PowerPoint presentation that outlines the approach that we will be taking.
- The client has agreed on a 3-week scope with the following 3 phases as follows - Data Exploration: Model Development

Projects:

- **Shopping Sales Project:** Developed an end-to-end data pipeline using AWS services and Airflow for reporting from Google Sheets API data.
- **Data Pipeline Automation:** Automated data extraction, transformation, and loading using Apache Airflow.
- **Real-time Data Streaming:** Implemented a real-time data ingestion system with Spark Streaming.
- **Gender and Age Detection System:** Built an image recognition model to predict gender and age.
- **Customer Segmentation System:** Utilized R and clustering techniques for customer segmentation.
- **Flight Delay Prediction Model:** Developed a model for predicting flight delays.

Education



Aditya Institute of Technology and Management

B. Tech degree, Electronics and Communication Engineering

2017 – 2021

CGPA: 6.8



Sri Chaitanya Junior College Higher

Secondary School Certificate 2015 - 2017



Ravindra Bharathi School, Pathapatnam

High School Diploma

2012 – 2015

CGPA: 8.2

Licenses & Certifications



Google Cloud Certified Professional Data Engineer - Google

Issued May 2024 - Expires May 2026

c4c5365d33e64ba58a5b0dd883f1f171



Microsoft Certified: Azure Data Engineer Associate - Microsoft

Issued Apr 2023 - Expires Apr 2025

I703-3125



Microsoft Certified: DevOps Engineer Expert - Microsoft

Issued Jun 2022 - Expires Jun 2025

I317-8196



Microsoft Certified: Azure Developer Associate - Microsoft

Issued May 2022 - Expires Jun 2025

I288-9769



Microsoft Azure Administrator Associate - Microsoft

Issued Apr 2022 - Expires Apr 2025

I233-3068



AWS Certified Data Engineer - Associate - Amazon Web Services (AWS)

Issued Jun 2025 - Expires Jun 2028

af8915b3b89141ddbbbc2f915b5db6a9



Microsoft Certified: Fabric Analytics Engineer Associate - Microsoft

X525F0-DE5E74 Issued Jan 2025 - Expires Jan 2026



Microsoft Certified: Azure Data Fundamentals - Microsoft

I510-9199

Skills

- Programming Languages: Python, SQL, PySpark, Scala
- Scripting Languages: UNIX, Powershell scripting
- Data Warehousing: Amazon Redshift, Google Big Query, Snowflake, Databricks, DBT, Azure Synapse
- ETL Tools: Apache Airflow,
- Big Data Technologies: Hadoop, Spark, Hive
- Database Systems: MySQL, Oracle
- Data Modeling: Dimensional Modeling, Star Schema, Snowflake Schema
- Data Integration: Extract, Transform, Load (ETL) processes
- Data Visualization: Tableau, Power BI, Sigma Computing
- Cloud Platforms: AWS, Azure, Google Cloud Platform
- Version Control: Git
- MS Office 365, Active Directory, MS Excel, MS Applications.
- Strong problem-solving and analytical skills
- Excellent communication and collaboration abilities

Languages

- Telugu
- English
- Hindi

Hobbies

- Reading Newspaper
- Listening Music
- Singing songs
- Researching Mysteries, Archaeology, and Mythology

Websites

GitHub: <https://github.com/durgaprasadpr718>

Credly: <https://www.credly.com/users/choudhury-durgaprasad>

Google Cloud Skills Boost: https://www.cloudskillsboost.google/public_profiles/c6999496-03ff-4cf6-88b2-be6fb55adac1

Portfolio: <https://durgaprasadchoudhury-portfolio.vercel.app>