

Vinayak Sivaprasad

Github: vinayak-byte

LinkedIn: Vinayak Sivaprasad

Email: vinayakbiju@gmail.com

Mobile: +91-813-7914-013

PROFESSIONAL SUMMARY

A Data Engineer with almost 3 years of experience at Ernst & Young. B.Tech. in Computer Science Technology with a solid foundation in data engineering principles. Skilled in designing and optimizing data pipelines for efficient processing and analytics. Known for independent delivery, collaboration across teams, and reducing operational inefficiencies.

EXPERIENCE

- Data Engineer – Ernst & Young Services LLP** Jan 2023 – Present
Kerala, India
 - ETL Pipelines:** Designed and implemented robust ETL pipelines using Azure Data Factory, Databricks, and SQL Server, reducing processing time by 30% (15 hrs/week).
 - PySpark Automation:** Optimized PySpark-based Python scripts, reducing manual effort by 25% and enhancing data quality by 20%.
 - Validation Automation:** Automated data validation processes, achieving 80% reduction in turnaround time and improving accuracy.
 - Modeling and Tuning:** Conducted data modeling and tuning, reducing storage cost by 25% and improving query response by 20%.
 - Standards and Practices:** Established best practices across teams, leading to 40% improved processing efficiency and 25% error reduction.
 - Warehouse Migration:** Migrated enterprise data warehouse to Azure Databricks and Azure SQL, optimized schemas and improved query performance, reducing storage costs by 40%.

EDUCATION

- Adi Shankara Institute of Engineering and Technology** Kerala, India
Bachelor of Technology - Computer Science and Engineering; SGPA: 7.61 Aug 2018 – Aug 2022

SKILLS SUMMARY

- Languages:** Python, SQL, C, PySpark
- Tools & Tech:** Azure Data Factory, Databricks, ADLS, Azure Data Studio, Synapse, SQL Server, Power BI, GIT, version control
- Platforms:** Windows, Linux
- Soft Skills:** Leadership, Time Management, Communication, Teamwork

PROJECTS

- Net Revenue Management Accelerator, UK:** Built a solution to analyze pricing, promotion, and product mix performance using Azure cloud tools.
 - Orchestrated PySpark-based data transformations within Databricks notebooks, refining raw data into Silver layers and analytical-ready Gold layers under the Medallion architecture, enhancing data quality by 60%.
 - Designed and implemented Delta Lake Tables/Columns; ensured schema alignment and efficient partitioning.
 - Performed data validation using Great Expectations and unit testing with Nutter framework.
 - Created and orchestrated Azure Data Factory pipelines to load data from ADLS into Azure SQL Data Warehouse.
 - Managed full-cycle data preparation and quality assurance across dev/test/prod environments.
- Retail Sector (Internal):** Created reusable pipelines to ingest CSV/XLSX data into Synapse using ADF + Databricks; handled schema mapping, views, and data QA.
- Utility sector, Scotland:** Processed Excel data in Databricks with PySpark transformations in databricks and created structured datasets for Power BI consumption.
- HR Analytics (Internal):** Migrated data from dedicated to serverless Synapse SQL pools, enabling SCD validation, logical and physical data modeling, and config-driven ETL loads, resulting in **40%** storage optimization and **50%** cost reduction. Designed and implemented ADF pipelines to ingest data from on-premises SQL Server and Oracle databases, perform delta comparisons against the latest data, and write incremental loads into Synapse.
- Formula 1 – Power BI Dashboard POC:** Built interactive dashboard showcasing race metrics from open datasets.
 - Designed visuals, calculated rankings and stats using DAX, and modeled data for race/event breakdowns.

CERTIFICATIONS & AWARDS

- Databricks Certified Data Engineer Associate:**
- Microsoft Azure Fundamentals (AZ-900):**
- Microsoft Azure Data Fundamentals (DP-900):**
- Microsoft Azure Data Engineer Associate (DP-203):**
- EY Spot Award for Delivery Excellence:**
- Client Exceptional Contributor Award:**

INTERNSHIPS

- Swinburne University, Malaysia:** International immersion program (Sept 2019)
- Cisco ThinkQbator:** Project developer internship (Jan 2019)