

Priyanka Yellesha

+1 312-973-9629 | priyankayellesha511@gmail.com | LinkedIn | GitHub | Portfolio

EXPERIENCE

Data Engineer Epsilon

Aug 2022 – Aug 2023
Bangalore, India

- Designed and maintained **data pipelines** using **AWS (S3, EC2)**, **Azure Data Factory**, **Databricks**, **Azure SQL Database**, and **PySpark**, ensuring accurate and scalable datasets for analytics.
- Automated reporting with **SQL stored procedures** and **Shell scripting**, reducing manual workload by **80–90%** and streamlining recurring processes.
- Migrated legacy **Excel reports to Power BI**, transforming raw **CSV data** from **AWS S3** into modeled datasets, cutting reporting turnaround by **35–40%**.
- Built and optimized interactive **Power BI dashboards**, improving refresh performance by **60%** and enabling faster decision-making.
- Developed real-time **event-driven pipelines** using **Azure Event Hubs** and **Stream Analytics**, containerizing workloads with **Docker and Kubernetes**.
- Led and mentored a team of **3 data engineers**, driving adoption of **DevOps practices** that reduced deployment time by **20%**.
- Conducted **data validation** and **anomaly detection** in **Databricks**, improving data accuracy, compliance, and trust in reporting.
- Optimized **SQL queries** and ETL workflows through indexing and partitioning, boosting pipeline performance by **40%** and supporting predictive analytics use cases.

Research Assistant

Feb 2021 – Aug 2021
Trichy, India

National Institute of Technology, Trichy

- Researched and implemented advanced **cooling strategies**, for **high-performance computing (HPC)** clusters to address rising **energy costs and semiconductor shortages** impacting **CPU sustainability**.
- Analysed **CPU thermal behavior** and its effect on execution time by monitoring **core temperatures** in a controlled **Ubuntu Linux** experimental setup using open-source **lm-sensors** and **affinity scheduling (puset)**.
- Designed experiments with **C (clock.t library)** to benchmark algorithm execution time across varied **CPU core temperatures**.
- Evaluated innovative cooling methods like **submersion cooling**, **rear-door exchangers**, and **direct-to-chip**, improving **heat transfer efficiency**.
- Demonstrated that **micro-channel liquid cooling** provided up to **1000x greater efficiency**, reducing **power consumption** and extending hardware life.
- Proposed **temperature-aware workload scheduling**, improving execution times, reducing thermal stress, and supporting **sustainable HPC operations**.
- Validated that higher **CPU temperatures** degrade **performance** and implemented **temperature-aware workload scheduling**, achieving **performance optimization**, improved **cooling efficiency**, and supporting **sustainable computing practices** in **HPC environments**.

Jr Data Engineer

April 2019 – Jul 2019
Bangalore, India

Environ

- Built a **SQL-based bibliographic data storage system**, streamlining storage and retrieval for research teams and improving productivity.
- Implemented **schema designs**, optimized **SQL queries**, and applied **indexing strategies**, improving query performance by **25%**.
- Developed and automated **ETL workflows** using **Python** and **Shell scripting**, reducing manual effort by **60%**.
- Designed and tested **data validation checks** with sample datasets, ensuring **data accuracy**, reliability, and integrity.
- Created **Power BI dashboards** to visualize bibliographic trends, enabling faster insights and supporting research decision-making.

PROJECTS

Sentinel – Enterprise Windows Activity Monitoring System

Azure, AD, Docker, PostgreSQL, MySQL, Ubuntu, TLS, Azure Blob Storage

- Built a **Windows-based monitoring system** to track login/logout times, app usage, and idle states, improving productivity and IT governance.
- Deployed a secure **cloud infrastructure on Microsoft Azure** with VNet, NSGs, TLS, and backup policies for compliance.
- Designed and optimized a **PostgreSQL database schema** with automated backups and retention policies.
- Integrated **Active Directory** for authentication and **RBAC** to secure and streamline user access.
- Containerized backend services with **Docker** and delivered a **web dashboard** for role-based activity visualization.

Retail Sales Analytics Platform

Azure Synapse, Snowflake, ADF, Databricks, PySpark, SQL, Power BI

- Ingested multi-region **POS data** into a centralized **Azure Synapse/Snowflake warehouse** using **Azure Data Factory** and **Databricks**.
- Designed and optimized **ETL pipelines** with **PySpark/SQL**, ensuring clean, standardized, and accurate datasets for analytics.
- Built predictive **sales forecasting models** on **Databricks** with **PySpark**, improving forecast accuracy by **25%** and supporting inventory and demand planning.
- Developed interactive **Power BI dashboards**, reducing manual reporting effort by **70%** and enabling real-time insights.
- Implemented **RBAC security** and aligned KPIs with business teams, cutting ad-hoc reporting requests by **40%**.

SKILLS

Programming & Scripting: Python, SQL, PySpark, Bash.

Data Engineering: ETL Pipelines, Data Modeling, Data Warehousing, Delta Lake, Data Governance.

Databases & Warehousing: MySQL, PostgreSQL, Oracle, MongoDB, Azure Synapse, Snowflake, BigQuery.

Cloud Platforms: Microsoft Azure, AWS, Google Cloud Platform (GCP).

Big Data & Streaming: Databricks, Apache Spark, Kafka, Azure Event Hubs, AWS Kinesis.

DevOps & Containerization: Docker, Kubernetes, Jenkins, Git, CI/CD.

BI & Analytics: Power BI, Tableau, SAS, Excel (Advanced), Pandas, NumPy.

Tools: GitHub, Bitbucket, Jira, Postman, VS Code.

EDUCATION

Elmhurst University

Elmhurst, IL

Master's in Computer Information Technology

National Institute of Technology, Trichy

Trichy, India

Bachelor of Technology in Mechanical Engineering

Certifications

- Microsoft Certified: Associate in Azure Data Engineer.
- Databricks Certified Data Engineer Associate.
- Microsoft Certified: Data Analyst Associate.
- AWS Certified Data Analytics – Specialty
- Microsoft Certified: Power BI Data Analyst.