

# Priyamvradha Parthasarathi

Dallas, Texas | (469) 465-3992 | [prriyamvradha@gmail.com](mailto:prriyamvradha@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## CERTIFICATIONS & TECHNICAL SKILLS

**Certifications:** Alteryx Designer Core, AZ - 900 Azure Fundamentals, SQL Fundamentals (ANSI) certified  
**Tools:** Ab Initio, Alteryx, Kleeene.AI, Databricks, Tableau, Advanced Excel, Excel VBA, JIRA, Git  
**Programming:** R, Python, SQL, Hadoop, Hive/Impala, PySpark, Unix/Linux, Data Manipulation Programming  
**Databases:** MySQL, Snowflake, Oracle, MongoDB, DynamoDB  
**Cloud:** AWS (Glue, Redshift, S3, Athena, Lambda), Azure (Azure Data Factory, Azure Blob Storage, AKS)

## PROFESSIONAL EXPERIENCE

**Data Specialist - Kaay Labs Inc,** United States

June 2024 – Present

- Cleaned and transformed raw data using **Pandas** and **SQL** to ensure data quality and integrity, preparing it for downstream **analysis** and **reporting**.
- Designed and deployed a scalable web service architecture that leverages **RESTful APIs** to connect AI models with web applications, providing real-time data processing and response capabilities.

**Data Engineer Intern - Kaay Labs Inc,** United States

June 2023 – August 2023

- Formulated and executed data schema for internal data warehouses and **SQL/NoSQL databases**, leading to a 25% increase in storage efficiency and a 15% improvement in data retrieval speed.
- Created reusable **PySpark** scripts and **Databricks** notebooks to automate data cleansing and preprocessing tasks.
- Utilized Hadoop ecosystem tools (**HDFS, Yarn, MapReduce**) and AWS services (**S3, DynamoDB**) to develop and maintain scalable data architectures.

**Senior Data Engineer - Larsen and Toubro Infotech LTI,** India

May 2019 – July 2022

- Developed **data solutions** for a banking client by orchestrating a smooth and effective data flow through data pipelines and **ETL workflows** resulting in a 30% reduction in overhead data processing time.
- Utilized advanced **Excel** and **Ab Initio** (ETL Tool) to build an automated reporting system which created daily and monthly reports for the US Federal Reserve system, leading to 90% reduction in manual report preparation time.
- Worked on importing data from **MySQL** DB to **HDFS** and vice-versa using **Sqoop** to configure Hive metastore with MySQL, which stores the metadata for Hive tables.
- Implemented scalable ETL workflows in **PySpark** on **Databricks** to efficiently transform, clean, and load data from diverse sources (e.g., JSON, CSV, Parquet) into a centralized data lake.
- Led a team of 20 members and **collaborated** on upgrading a web application (MHUB 4. x) and completed upgrade without impact within 2 months and received strong client appreciation for efforts.
- Leveraged **Agile methodology** to prioritize and manage project tasks, realizing a 25% uptick in team productivity.

## PROJECTS

**Food Sales Data Analysis using PySpark** | (PySpark, Databricks) | [GitHub](#)

June 2024

- Utilized **PySpark** on Databricks for **data cleaning** and KPI metric extraction from a food sales dataset, employing PySpark SQL and visualized actionable **business insights**.

**Google Play store Data Analysis using PySpark** | (PySpark, SQL, Databricks) | [GitHub](#)

May 2024

- Analyzed Google Play Store data using PySpark on Databricks, performing data cleaning, **KPI metric extraction** with **SQL**, and visualization to reveal **key trends** and insights.

**Truck Fleet Movement Reporting and Analytics** | (Cloudera, Tableau, Python) | [GitHub](#)

May 2023

- Worked on a Truck fleet movement project integrating **Hadoop** ecosystems, **Hive/Impala** with **Tableau**, and **regression analysis** to boost safety and efficiency in transportation.

**Alteryx Datathon Project – Streamlining IRS Reporting process** | (Alteryx, Power BI) | [GitHub](#)

February 2023

- Enhanced U.S. companies' IRS reporting with **Alteryx**, constructed **ETL pipelines** for data cleansing and aggregation, mitigated processing time by 30%, and elevated fiscal transparency through insightful reports.

**E-Commerce Database Management System** | (Data Modeling, MySQL) | [GitHub](#)

November 2022

- Created an amplified e-commerce **MySQL database model**, applying data normalization techniques for precise data analysis and elevated business insights to boost sales and profits by 70%.

## EDUCATION

**Master of Science, Business Analytics** - The University of Texas at Dallas | GPA 3.83 - Scholar with Distinction

May 2024

**Bachelor of Engineering, Electronics and Communication Engineering** - Anna University | GPA 3.30

April 2019