SATHVIK PUTTA

+1(475) 837 8069 ◊ puttasathvik16@gmail.com ◊ LinkedIn ◊ GitHub ◊ Portfolio
– New Haven, CT (Open to Relocation)

PROFILE SUMMARY

Data Analyst with hands-on experience in data engineering and cloud technologies. Interned at TRIX.edu, Hyderabad, and worked as a Microsoft Azure & Power BI Analyst at ICT Academy, Chennai. Proficient in designing ETL pipelines, performing database migrations, and developing data-driven solutions using Azure, Databricks, and Power BI. Earned a master's degree in data science from the USA and was inducted into the Upsilon Pi Epsilon (UPE) Honor Society for academic excellence in computing. Recognized for driving innovation, optimizing processes, and delivering impactful data solutions.

EDUCATION

Master of Science in Data Science and Engineering Jan 2024 – March 2025
Sacred Heart University, Connecticut, USA – Honor Society 2025
Relevant Coursework: Big Data Analytics, Advance Database Design, Data Science Architecture, Deep learning, Machine Learning, Text Based Analysis, Applied Statistics.

Bachelor of Technology in Computer Science and Engineering Aug 2019 - May 2023
Saveetha School Of Engineering, India
Relevant Coursework: Python, Java, R-Programming, SQL, Data Structures, Data Visualization, Artificial Intelligence, Computer Networks, Operating systems, Cloud Computing, Predictive Analysis

SKILLS

- **Programming Language**: Python, C++, C, JavaScript, HTML/CSS, SQL
- **Cloud & Big Data:** Microsoft Azure, Databricks, Azure DevOps
- **Frameworks:** Flask, Django, FastAPI, ReactJS, ExpressJS, Node.js, React Native
- **ETL & Data Engineering:** Building ETL pipelines, Data transformation, Data Modeling.
- **Database & Data Movements:** SQL, MongoDB, MySQL, SAP.
- **Visualization & Analytics:** Tableau, Power BI, Matplotlib, Data Architecture.

WORK EXPERIENCE

Data Engineering Intern Mar 2021 – Dec 2021
TRIX.edu Hyderabad, India

- Built ETL pipelines with Azure Data Factory & Databricks as part of a data engineering project across multiple data sources by writing Spark-based transformations and managing pipeline orchestration, resulting in efficient and reliable data integration.
- Migrated MySQL to Azure SQL in a team of four over six months using Python and by effectively communicating with team members, leading to enhanced performance and scalability.
- Processed data from Excel, SQL Server, and web APIs and created Power BI dashboards using DAX and Power Query, collaborating closely with stakeholders to define metrics and KPIs, which led to actionable academic insights and improved decision-making.
- Optimized pipeline performance ensured data security, and documented workflows as part of enterprise data governance by applying performance tuning, implementing RBAC policies, and creating structured documentation, ensuring compliance with organizational standards and reducing troubleshooting time by 30%.

Microsoft Azure and PowerBI Analyst Oct 2022 – Feb 2023
ICT Academy Chennai, India

- Worked on cloud-based data management using Microsoft Azure and Databricks with a focus on ETL pipeline development and data architecture by leveraging Spark, Delta Lake, and best practices in modular pipeline design, resulting in scalable and maintainable data workflows.
- Applied skills in cloud computing, data migration, and business intelligence using Azure Data Factory, Power BI, and SQL by automating data workflows, transforming datasets, and designing dashboards for KPI tracking, leading to improved operational visibility and data-driven decisions.
- Worked with Azure, Databricks, and Power BI for data integration, transformation, and reporting by developing SQL scripts, writing PySpark code, and building interactive dashboards, resulting in faster reporting cycles and improved data accessibility for stakeholders.
- Developed and optimized data pipelines for structured and unstructured data formats (CSV, JSON, SQL) using Azure Data Factory and Databricks with built-in security protocols and parallel processing, ensuring scalable ingestion, transformation, and compliance with data privacy requirements.

PROJECTS

Northeastern United States traffic accident trends: a geospatial and statistical analysis using python IEEE | GitHub

- Analyzed traffic accident trends in the Northeastern US using Python, applying geospatial mapping with geopandas and statistical modeling with scipy and pandas, which uncovered high-risk zones and seasonal accident patterns.
- Processed and analyzed traffic accident datasets using Python libraries, filtering and aggregating data to identify patterns and correlations among speeding, alcohol involvement, and insurance rates, helping inform targeted policy recommendations.
- Used Python libraries including Pandas, Scikit-learn, Matplotlib, and Seaborn to clean and model traffic accident data and visualize trends across demographics and regions, leading to clear identification of contributing risk factors and visual reports for stakeholders.
- Presented research findings at ICICT 2025 after conducting exploratory and inferential analysis of accident data and synthesizing insights into an academic paper and presentation, which offered actionable recommendations for road safety improvements and behavioral interventions.

Enterprise Data Migration and Analytics Pipeline | ICT Academy

- Developed a data migration and analytics pipeline using Azure, Databricks, and Power BI by designing modular ETL flows and integrating data across sources, resulting in a unified, real-time analytics platform.
- Migrated legacy databases to Azure SQL and implemented real-time ETL pipelines using Azure Data Factory and PySpark, followed by dashboard creation in Power BI, improving system performance and enabling timely decision-making.
- Used Azure Data Factory for data integration, Databricks & PySpark for transformation logic, and Power BI to design interactive visual dashboards, resulting in accurate reporting and efficient data pipeline orchestration.
- Enabled seamless migration, built a scalable architecture, and delivered insights by orchestrating secure ETL flows, implementing monitoring tools, and optimizing performance at each stage, leading to a 40% increase in data processing speed and timely reporting.

Exploring Sentiment Towards a 10-Year Minimum Occupation Period in Prime Location Housing (PLH) Model GitHub

- Developed a complete data migration and analytics pipeline using Azure, Databricks, and Power BI, integrating legacy systems into a unified platform to enable real-time business intelligence.
- Migrated legacy databases to Azure SQL and implemented real-time ETL pipelines using Azure Data Factory and PySpark, significantly improving data availability and system performance.
- Engineered data workflows with Azure Data Factory for ingestion, Databricks and PySpark for transformation, and Power BI for visualization, enabling interactive dashboards and automated reporting.
- Built a scalable, secure architecture and optimized ETL performance, resulting in a 40% improvement in data processing speed and delivery of actionable insights to stakeholders in real time.

AWARDS

- Graduated with distinction from MS in CS-IT (Data Science), Sacred Heart University, USA; inducted into Upsilon Pi Epsilon Honor Society (2025) for academic excellence in computing.
- Awarded Academic Topper of 2023 at SIMATS – Saveetha School of Engineering, recognized on the university stage with a merit-based cash prize for academic excellence.
- Recognized as Academic Student Topper of 2022 at SIMATS – Saveetha School of Engineering; honored on the university stage for outstanding academic performance.