

## **Doraswami Hanumanthu**

Azure Data Architect | Data Engineer | Business Analyst

**Mob No.:** +91-8309162265 | **Email:** [dshanumanthu1@gmail.com](mailto:dshanumanthu1@gmail.com)

---

### **Summary**

Dynamic and results-driven Azure Data Architect & Data Engineer with 14.8+ years of experience in data engineering, cloud architecture, and business intelligence. Proven expertise in Microsoft Fabric, Azure Data Services, SQL, Python, and Power BI. Strong background in building scalable ETL pipelines, data governance, real-time analytics, and data warehousing solutions for enterprise domains such as HR, Finance, and Supply Chain.

---

### **Certifications**

- Microsoft Azure Data Engineer Associate (DP-203) (In Progress)
  - PGP in Data Science – INSOFE, Hyderabad
  - Capstone Business Simulation
- 

### **Skills & Technologies**

- **Data Engineering & Cloud Technologies:** Microsoft Fabric, Azure Data Factory, Azure Databricks, ADLS, Synapse, Snowflake, AWS Redshift, Google BigQuery
  - **Programming & Scripting:** Python, PySpark, SQL, Scala, C#, DAX, MDX
  - **ETL & Data Pipeline Orchestration:** Airflow, Apache Kafka, DBT, Presto, Spark, Airbyte
  - **Data Modeling & Warehousing:** Dimensional Modeling, Data Vault, Star & Snowflake Schema, MPP Systems
  - **BI & Analytics:** Power BI, Tableau, Zoho Analytics, Business Intelligence Strategy
  - **Data Governance & Security:** Microsoft Purview, Role-Based Access Control (RBAC), Compliance Standards
  - **Agile & DevOps:** GitHub Actions, CI/CD, DataOps, JIRA, Confluence
- 

### **Professional Experience**

**TAFE India Pvt Ltd, Chennai**

**Azure Data Architect & Data Engineer | Apr'24 – Present**

## Project: Real-time Data Ingestion and Analytics using Microsoft Fabric

**Technologies Used:** Microsoft Fabric, Azure Event Hub, Fabric Event Stream, Fabric Eventhouse, Table Update Policy, Materialized Views, Power BI, Python, PySpark, KQL, Azure Data Factory

### Project Description:

This project involved designing and implementing a **real-time data ingestion and analytics platform using Microsoft Fabric**. The system processes high-volume data from multiple external API sources (both **push and pull APIs**) through **Azure Event Hub and Fabric Event Stream**, ensuring **real-time data processing, transformation, and visualization**.

### Key Responsibilities:

- **Real-time Data Ingestion:** Configured **Azure Event Hub** to receive real-time data from multiple external APIs and seamlessly stream the data into **Microsoft Fabric Event Stream**.
- **Data Processing & Storage:** Stored raw and processed data in **Fabric Eventhouse**, ensuring optimized query performance and scalability.
- **Data Transformation & Enrichment:**
  - Implemented **Table Update Policies** to dynamically transform and process ingested data.
  - Created **Materialized Views** to pre-aggregate and enhance analytical query performance.
- **Event-driven Processing:** Designed an event-driven **streaming pipeline** to ensure **low-latency** data processing for near real-time analytics.
- **Data Governance & Security:** Applied **Role-Based Access Control (RBAC)** and **Microsoft Purview** for **data governance, compliance, and security measures**.
- **Business Intelligence & Reporting:** Developed **interactive Power BI dashboards** to provide **real-time insights** on business metrics, enabling stakeholders to monitor trends dynamically.
- **Error Handling & Monitoring:**
  - Implemented **automated alerting** for event failures using **Azure Monitor**.
  - Set up **Fabric Eventhouse logging** to track data anomalies and ensure data integrity.

### Achievements:

- **Reduced data latency by 40%** through real-time streaming via **Fabric Event Stream and Eventhouse**.
- **Enabled near real-time analytics** with **Materialized Views**, improving query performance by **60%**.
- **Enhanced business decision-making** with **real-time Power BI reports**, providing actionable insights within seconds.

Data Engineer | Mar'23 – Apr'24

## Project: Real-time Data Integration and Analytics

**Technologies Used:** Azure Data Factory, Azure Databricks, Python, PySpark, Azure SQL Database, RESTful APIs, Microsoft Fabric

### **Project Description:**

In this project, I worked on integrating **data from various API sources** into a **unified data warehouse**. The APIs used are **push and pull** types, where data is both retrieved and sent to external sources. The project aimed to create a scalable pipeline for handling large volumes of real-time data while maintaining high efficiency and reliability.

### **Key Responsibilities:**

- **Data Extraction:** Designed and implemented data extraction pipelines from multiple external APIs using **Azure Data Factory** and **Azure Logic Apps**, ensuring smooth data flow.
- **Data Transformation:** Used **Azure Databricks** and **PySpark** to process and transform incoming data for better analysis and reporting, implementing **data wrangling** techniques.
- **Storage:** Stored the processed data in **Azure SQL Database** and **Azure Data Lake**, optimizing for speed and access for further analytics.
- **Real-time Data Handling:** Utilized event-driven architectures to ensure real-time data processing with minimal latency, integrating with APIs that push data at regular intervals.
- **API Integration:** Worked with both **GET** and **POST** API methods for pulling and pushing data between various systems, ensuring secure and efficient API management.
- **Reporting & Analytics:** Built **real-time dashboards** and reports in **Power BI** and **Microsoft Fabric** to provide insights into business metrics, ensuring stakeholders had up-to-date information.
- **Error Handling:** Implemented **robust error handling and logging** mechanisms to ensure continuous data flow and troubleshoot potential issues without manual intervention.

### **Achievements:**

- Reduced **data latency** by **30%** with optimized data extraction and transformation processes.
- Enabled **real-time reporting**, allowing stakeholders to make **data-driven decisions** promptly.

**AAF India Pvt Ltd, Bangalore**

**Data Engineer | Apr'21 – Jan'23**

### **Project: Supply Chain Optimization for Production**

**Technologies Used:** Azure Data Factory, Power BI, SQL, Python, PySpark, Azure Synapse Analytics, Azure SQL Database, Tableau, Materialized Views, Data Warehouse

### Project Description:

This project focused on optimizing the **supply chain process** for a **manufacturing and production environment**. The objective was to integrate and analyze **real-time production data** to improve forecasting, inventory management, and order fulfillment. The solution involved automating the flow of data from various sources, including production systems, ERP systems, and external suppliers, to generate actionable insights and improve the efficiency of the supply chain.

### Key Responsibilities:

- **Data Integration & ETL:** Designed and implemented **ETL pipelines** using **Azure Data Factory** to extract data from multiple systems, including **SAP, Oracle ERP, and production systems**. The data was transformed and loaded into an **Azure SQL Data Warehouse** for centralized storage.
- **Supply Chain Analytics:**
  - Created **data models** for tracking **inventory levels, demand forecasting, and production timelines**.
  - Used **Materialized Views** and **Azure Synapse Analytics** to pre-aggregate data, enabling faster and more efficient querying for real-time reporting.
- **Inventory Management & Order Fulfillment:** Implemented solutions to automatically track stock levels and generate real-time alerts when inventory fell below predefined thresholds. This helped reduce stockouts and overstock situations by **20%**.
- **Business Intelligence & Reporting:**
  - Developed **interactive Power BI dashboards** for **inventory management, production monitoring, and order status** to provide key stakeholders with real-time insights into supply chain performance.
  - Integrated **Tableau** for advanced reporting and KPI tracking, allowing supply chain managers to make data-driven decisions.
- **Optimization & Automation:** Designed **automated workflows** using **Azure Logic Apps** to streamline data flow between systems and trigger notifications for key actions like stock reordering and production delays.
- **Collaboration & Communication:** Worked closely with **operations teams, suppliers, and production managers** to understand business requirements and align the solution to meet business needs.

### Achievements:

- Successfully reduced **order fulfillment time** by **30%** through automated data flow and optimized inventory management.
- Improved **forecast accuracy** by **25%** through advanced **demand planning** and data analytics.
- Increased **production throughput** by **15%** by ensuring better alignment between inventory levels and production capacity

### Project: Sales and Billing Data Integration for Real-time Analytics

**Role:** Data Engineer / Business Analyst / Data Architect

**Technologies Used:** Azure Data Factory, SQL Server, SAP, Power BI, Python, PySpark, Azure Synapse Analytics, REST APIs, Joho CRM, ETL

## Project Description:

This project involved creating an **end-to-end ETL pipeline** to integrate **Sales data from Joho CRM** with **billing data from SAP** to provide comprehensive and actionable insights to business stakeholders. The goal was to build a system that enables real-time decision-making for the **sales and dealer network**. By connecting the customer, sales, and billing data, the system provides an integrated view of business activities, improving the efficiency of decision-making, forecasting, and business operations.

## Key Responsibilities:

- **Data Integration:**
  - Designed and implemented **ETL pipelines** to integrate data from two primary sources: **Joho CRM** for customer and sales data and **SAP** for billing and payment information.
  - Used **REST APIs** to extract data from **Joho CRM** and **SAP** (via OData or direct connections) and transformed it using **Azure Data Factory**.
- **Data Transformation:**
  - Applied **data wrangling** techniques using **PySpark** to clean and transform raw data from both systems into structured, analytical formats.
  - Combined customer data with billing details, creating a unified customer sales profile.
  - Ensured **data quality**, checking for duplicates, missing values, and discrepancies between CRM and billing data.
- **Data Storage & Management:**
  - Stored transformed data in **Azure Data Lake** and **Azure Synapse Analytics**, implementing **data partitioning** for scalability and performance optimization.
  - Created a **data warehouse** to centralize customer, sales, and billing data, enabling efficient querying and analysis.
- **Real-time Analytics:**
  - Developed **real-time data pipelines** for continuous data synchronization between Joho CRM, SAP, and the central data warehouse.
  - Used **Azure Stream Analytics** to provide real-time insights into sales activities and billing status, enabling stakeholders to monitor performance continuously.
- **Business Intelligence & Reporting:**
  - Created **Power BI dashboards** for stakeholders to track sales performance, billing status, customer trends, and dealer activities.
  - Implemented **KPIs** to monitor the health of the business, such as **sales conversion rate**, **average deal size**, **billing cycle time**, and **dealer performance**.
- **Dealer Access & Real-time Decisions:**
  - Developed a **dealer portal** where dealers can log in to view their performance, sales figures, and billing status in **real-time**.
  - Enabled **real-time decision-making** for dealers by providing them access to critical data points for business actions.
- **Automation & Scheduling:**
  - Automated data synchronization tasks and reporting with **Azure Data Factory** scheduling to ensure timely updates of business data.
  - Used **version control** and **CI/CD pipelines** in **Git** for deployment management.

## Achievements:

- Improved **business decision-making** by providing real-time analytics and visualizations to stakeholders and dealers.
- Increased **operational efficiency** by automating data integration and reporting processes, reducing manual interventions.
- Enhanced **sales tracking accuracy** by consolidating customer and billing data into a unified dashboard, enabling better customer segmentation and insights.
- Reduced **billing discrepancies** by implementing an integrated data flow between **CRM** and **SAP** systems, ensuring data consistency.

## TATA AIA Life Insurance, Hyderabad

Data Engineer | Jul'20 – Apr'21

### Project: Insurance Data Integration and Analytics for Real-time Decision Making

**Role:** Data Engineer / Business Analyst / Data Architect

**Technologies Used:** Azure Data Factory, Azure Synapse Analytics, Power BI, SQL, Python, PySpark, Azure Blob Storage, Azure Data Lake, REST APIs, Azure Logic Apps

### Project Description:

In this project, I worked on integrating multiple data sources such as **customer data**, **policy information**, and **claims data** within the **insurance domain** to enable **real-time reporting** and **data-driven decision making**. The goal was to create a unified platform that could process large volumes of data from legacy systems, optimize the customer experience, and reduce manual effort in policy management, claims processing, and customer analytics.

### Key Responsibilities:

- **Data Integration & ETL Pipelines:**
  - Designed and implemented **ETL pipelines** using **Azure Data Factory** to extract data from multiple sources including **CRM systems**, **policy management systems**, and **claims databases**.
  - Utilized **REST APIs** to integrate with third-party services for retrieving customer data, policy information, and claim records.
  - Built **data pipelines** to transform raw data into usable formats, ensuring high data quality and reducing discrepancies across data systems.
- **Data Storage & Management:**
  - Designed a **data warehouse** in **Azure Synapse Analytics** for storing consolidated policy, claims, and customer data, optimized for both reporting and real-time access.

- Used **Azure Data Lake** to store raw and semi-structured data (e.g., claims reports, customer details in JSON format), improving accessibility and enabling future data mining processes.
- Implemented **Azure Blob Storage** for storing large volumes of unstructured data related to customer claims and insurance policies.
- **Real-time Analytics:**
  - Implemented **event-driven pipelines** using **Azure Event Hubs** and **Azure Stream Analytics** for real-time claims data processing.
  - Built real-time dashboards to monitor claims status, policy renewals, and customer activities, enabling insurers to take immediate action where required.
- **Data Transformation & Reporting:**
  - Used **PySpark** for **data transformations** and **aggregations** on large datasets to derive actionable insights like **claim frequency analysis** and **customer churn prediction**.
  - Developed **Power BI dashboards** to visualize key metrics such as **claim processing time**, **policy sales performance**, and **customer segmentation**.
  - Implemented **Power BI reports** for senior management to track KPIs, such as **total claims processed**, **claim denial rates**, and **policy renewal rates**.
- **Automation & Data Quality:**
  - Automated **data synchronization** between different systems using **Azure Logic Apps** to trigger workflows when certain conditions are met (e.g., claim approval or policy renewal).
  - Implemented **data quality checks** and validation rules to ensure that the claims and policy data were consistent and accurate across the system.
- **Compliance & Data Security:**
  - Ensured **data governance** by implementing role-based access control (RBAC) using **Azure Active Directory** to restrict access to sensitive customer and claims information.
  - Worked with the compliance team to ensure that all data handling followed **GDPR** and **HIPAA** regulations.

#### **Achievements:**

- Successfully automated the **claims processing workflow**, reducing the processing time by **30%** and enabling real-time insights into claims status.
- Enabled **predictive analytics** for customer behavior, helping to identify customers at risk of **policy lapses** and **claims fraud**.

- Increased **operational efficiency** by automating **data integration** from CRM, policy, and claims systems, reducing manual data entry by **40%**.
  - Created **interactive dashboards** in **Power BI** that enabled stakeholders to make real-time decisions regarding policy pricing, claims management, and customer retention.
- 

#### **Key Benefits to the Business:**

- **Improved Claims Management:** Real-time tracking and monitoring of claims, allowing for quicker decision-making and improved customer satisfaction.
- **Enhanced Policy Management:** Optimized policy renewal processes, reducing churn by identifying customers with higher renewal likelihoods.
- **Customer Segmentation:** Used analytics to classify customers into different groups based on risk, claim history, and policy preferences.
- **Data-Driven Decisions:** Empowered business stakeholders with real-time insights into the business, leading to quicker and more informed decisions.

#### **IKEA India Pvt Ltd, Hyderabad**

##### **Data Engineer | May'18 – Jul'20**

- Implemented Azure-based data solutions to enhance customer insights and store operations.
- Developed Power BI, Zoho A dashboards to track checkout efficiency and reduce wait times.

#### **Sleek Modular Kitchens, Hyderabad**

##### **Management Trainee (Data visualization) | Apr'17 – Apr'18**

- Led data migration projects integrating ERP systems with Azure & SAP.
- Designed reporting dashboards to track supply chain KPIs.

#### **Mphasis, Pune**

##### **Software Engineer | Jun'2009 – Jun'15**

- Managed big data processing using Spark and Azure Data Lake.
  - Optimized SQL queries to improve database performance by 45%.
- 

#### **Education**

- **Pursuing M.Tech (AI)** – SRM University (2024-2026)



- **MBA (Marketing & Business Analytics)** – Amrita School of Business, Bengaluru
  - **B.Tech (ECE)** – JNTU Anantapur
- 

### **Achievements**

- Successfully led data modernization initiatives, improving data accessibility & efficiency.
  - Optimized ETL pipelines, reducing data processing time by 50%.
  - Designed scalable data architectures using Microsoft Fabric & Azure Data Services.
- 

### **Areas of Expertise**

- Data Engineering & Cloud Solutions
  - Business Intelligence & Data Analytics
  - AI & Machine Learning for Business Insights
  - Data Governance & Security Compliance
- 

### **Tools & Platforms**

- **BI & Reporting:** Power BI, Tableau, Zoho Analytics
- **Data Engineering:** Azure Data Factory, Databricks, SQL, Snowflake, Fabric
- **Programming:** Python, PySpark, SQL, APIs
- **Cloud & Platforms:** Azure, AWS, Data Warehousing
- **ETL & Workflow Orchestration:** Airflow, DBT, Presto, Kafka
- **Agile & DevOps:** GitHub Actions, CI/CD, JIRA, Confluence