

## SUMMARY OF EXPERIENCE

A **Software Engineer** with **7+ years of experience in the IT industry** having good knowledge in Azure Cloud, ETL Developer, and Data Analyst which includes Requirement gathering & Analysis, Development & Unit Testing.

- Experience in Azure Data Factory, Azure Cosmos DB and Azure Data bricks, Azure key vault, Azure SQL Database, ADLS, Azure Delta Lake and CICD using Azure DevOps.
- Experience in Azure Data factory Pipeline Audit logs.
- Strongly knowledge on dataflows in data factory pipelines.
- Efficient in performing Data analysis using SQL queries for different business needs.
- Excellent knowledge and experience in creating source to target mapping, validations, transformations, and business rules.
- Working in Agile Methodologies.
- Ability to quickly adapt to the technical environment as a developer.
- Good team player, committed, result oriented with a quest and zeal to learn new technologies and undertake challenging tasks
- Demonstrated track record of successful project development and implementations with strong technical, logical, analytical and problem-solving skills.

## EMPLOYMENT HISTORY

- Working as an Azure data engineer (Associate) in **Cognizant Technology Solutions**, Hyderabad India from October 2021 to till date
- Worked as a Big data developer in Versatile Softwares Private Limited, Bangalore, India from October 2016 to September 2021

## EDUCATION PROFILE:

- Post-Graduation in Master of Computer Applications (M.C.A) with 60% from Osmania University in 2012.

## CERTIFICATION COMPLETED

- Microsoft Certified: Azure Data Engineer Associate **DP 203** in June 2023
- Microsoft Certified: Microsoft Azure Fundamentals **AZ 900** in June 2022
- Microsoft Certified: Azure Data Fundamentals **DP 900** in June 2023
- Microsoft Certified: Azure AI Fundamentals **AI 900** in March 2024.

## TECHNICAL SKILLS

- Azure Cloud : Azure Data factory, Azure Databricks, Azure SQL Server, Azure Cosmos DB, Azure Data Lake Gen2, Azure blob storage
- Big Data Components : Hadoop 2.7.0, Sqoop 1.4.6, Hive- 1.1, PySpark, Spark SQL.
- Language : Python, SQL, UNIX Shell Script.
- Back End : SQL Server, MySQL
- Operating System : Windows, Linux, Ubuntu
- Packages : MS Office.
- IDE : Eclipse, IntelliJ, Visual Studio
- Source Control Tool : GitHub, Azure DevOps

## PROFESSIONAL EXPERIENCE

<b>1. Project Title</b>	<b>OSF HealthCare- KnowME - Digital Interaction</b>
Client	OSF HealthCare
Period	October 2021 – till date
Tools / Database	Azure Data Factory, Cosmos DB, Teradata, SQL Server
Team Size	4
Role	Developer
Organization	Cognizant Technology Solutions

### **PROJECT DESCRIPTION:**

The objective of this project is to create an operation report to visualize the scope and digital imprints of the OSF outreach programs using Upfront along with clinical and non-clinical patient data such as Patient contact & demographic, current health condition, social determinants of health and Chronic Epic Disease registry and to integrate the data responses, pick latest file of each data set, clean the data, join and transform and convert into meaningful information and load into SQL DB/Cosmos DB. All information in one centralized location with longitudinal data view. The stages of data load include Raw Zone, Curate Zone and Cosmos DB. The first phase is to move the data from the Teradata table or External storage or API or Azure Blob container into the Raw Zone, and from the raw zone the Required logic will be included and move it to the Curated Zone and as a Final phase the data will be move into the Cosmos DB. The responses of these outreach initiatives are brought into Azure Data Lake as CSV files. The idea of the project is to integrate the data responses and build a report visualization in Excel tool to visualize the coverage, utilization, and benefits of the program.

### **ROLES AND RESPONSIBILITIES:**

- Implement Azure Data factory pipeline, Azure Databricks Notebooks to move data from multiple source system into Azure data lake and Azure Cosmos Database
- Analysis, design and data modelling of Azure Cosmos Database and Container structure based on the front-end process which consumes data from CDB, source data & load frequency.
- Perform detailed unit testing of the ETL components built before deploying them into higher environments (like test and production) and assist the quality assurance testing group, User Acceptance Testing groups during the corresponding software development testing lifecycle phases.
- Implement CI/CD Pipeline in Azure DevOps for Production migration
- Assist in the ongoing development of technical best practices for data movement, data quality, data cleansing and other ETL related activities.
- Extensively involved in offshore team handling, assessment meetings and technical peer reviews for receiving approvals for code deployment.
- Participate in Sprint retrospective, sprint planning and requirement analysis.

<b>2. Project Title</b>	Ecommerce
Client	AAA Teleshopping
Period	Nov 2020 – September 2021
Role	Big Data Developer
Team Size	4
Organization	Versatile Softwares Private Limited

**Responsibilities:**

- Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- Involved in handling streaming Data from **Kafka** and processing it in **Spark**.
- Was responsible for **Optimizing Spark SQL** queries that helped in saving Cost to the project.
- Handled streaming data using both **Spark streaming and Spark structured streaming** APIs.
- **Managed Kafka Consumption** in an optimized way for a better performance.
- Exported necessary spark Jars to run in the cluster.
- Involved in working on the Data Analysis, Data Quality and data profiling for handling the business that helped the Business team.
- **Loaded and transformed** large sets of semi structured data likes **CSV, JSON, Avro**.
- **Code & peer review** of assigned task. Unit testing and Bug fixing.

**Technologies:** Hadoop, HDFS, Hive, Sqoop, Spark, Spark SQL, Scala, Spark streaming, Kafka

<b>3. Project Title</b>	WIS – Vehicle Insurance
Client	FCA
Period	July 2019 -Oct 2020
Role	Spark Developer
Team Size	4
Organization	Versatile Softwares Private Limited

**Responsibilities**

- Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- Involved in handling streaming Data from **PostgreSQL** .and processing it in **Spark**.
- Was responsible for **Optimizing Spark SQL** queries that helped in saving Cost to the project.
- Exported necessary spark Jars to run in the cluster.

**Technologies:** Spark, Spark SQL, Scala, PostgreSQL, AWS

<b>4. Project Title</b>	Retail Management Facilities
Client	MN Telemart
Period	Nov 2016 -June 2019
Role	Hadoop Developer
Team Size	4
Organization	Versatile Softwares Private Limited

**Responsibilities:**

- Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- Involved in **creating Hive tables, data loading** and writing **hive queries**.
- Was responsible for **Optimizing Hive** queries that helped in saving Cost to the project.
- **Managed Hive Tables** and created child tables based on partitions.
- Involved in working on the Data Analysis, Data Quality and data profiling for handling the business that helped the Business team.
- **Loaded and transformed** large sets of semi structured data.

**Technologies:** Hadoop, HDFS, Hive, Sqoop, Cloudera CDH.

**Declaration:**

I hereby declare that the above written particulars are true to the best of my knowledge and belief.

**Place: Hyderabad**

**Date: 19.06.2024**

**(Jayalaxmi Pulipyata)**