| A picture containing text, sign, blue  Description automatically generated | Nivethika J  **469-532-0003**  **nivethikajmani@gmail.com** |
| --- | --- |

**Professional Summary:**

* Around 7 years of IT experience in implementing the end-to-end Data Warehouse and Data engineering solutions.
* Having 4+ years of professional experience as Azure Cloud Data Engineer in with Azure Data Factory(ADF), Azure Data Lake Storage(ADLS), Azure Synapse Analytics (SQL Data warehouse), Azure SQL Database, Azure Analytical services, Azure Cosmos, Azure Key vaults, Azure HDInsight and Azure Data bricks.
* Experience in developing pipelines in spark using Scala and PySpark.
* Experience in building ETL (Azure Data Bricks) data pipelines leveraging PySpark, Spark SQL.
* Proficient in Azure Data Factory to perform Incremental Loads from Azure SQL DB to Azure Synapse.
* Experience in building the Orchestration on Azure Data Factory for scheduling purposes.
* Experience in Azure cloud worked on App services, Azure SQL Database, Azure Blob storage, Azure Functions, Virtual machines, Azure AD, Azure Data Factory, event hub and event Queue.
* Experience working with systems processing event based logs, user logs, etc.
* Experience working with Azure Logic APP Integration tool.
* Experience with the Azure logic apps with different triggers
* Experience in Orchestrated data integration pipelines in ADF using various Activities like Get Metadata, Lookup, For Each, Wait, Execute Pipeline, Set Variable, Filter, until, etc.
* Experienced in working with structured data using HiveQL, and optimizing Hive queries.
* Running and scheduling workflows using OOZIE and Zookeeper, identifying failures and integrating, coordinating, and scheduling jobs.
* Experience in Database Architecture for OLAP and OLTP Applications, Database designing, Data Migration, Data Warehousing Concepts with emphasis on ETL.
* Experienced in Data Modeling & Data Analysis experience using Dimensional Data Modeling and Relational Data Modeling.
* Experience in using Star Schema/Snowflake Modeling, FACT & Dimensions tables, and Physical & Logical Data Modeling.
* Defining user stories and driving the agile board in JIRA during project execution, participate in sprint demo and retrospective.
* Experience in GIT source code repository and GitHub Enterprise.

**Technical Skills:**

* **Azure Cloud:**Azure SQL Database, Azure Data Lake, Azure Data Factory (ADF), Azure SQL Data Warehouse, Azure Analysis Service (AAS), Data Migration Service (DMS), Azure SQL Data, Elastic Pools, Geo Replication, Geo-restore, Azure SQL Analytics.
* **ETL Tools:**Azure Data Factory (ADF), Azure Database migration Service (DMS), ETL SQL Server Integration Services (SSIS), SQL Server Reporting Services (SSRS), BCP.
* **Storage:**Azure Storage, Azure Blob Storage, Azure Backup, Azure Storage Disks- Premium, Azure Files.
* **Performance Monitoring Tools used:**Azure SQL Analytics, Spotlight, IDERA, Red Gate, SQL SentryOne, SQL Server Profiler, System Monitor.
* **CI CD Tools:** Terraform, Jenkins.
* **Programming:** Scala, Python, PySpark, SQL,Shell script, Perl script.
* **Operating System:**Azure Virtual Machine (VM), Windows Server 2019, Windows Server 2016, Windows Server 2012 R2, Windows 10, Windows 7, Windows XP, Unix, LINUX. Shell Script.
* **Scripting:**Windows PowerShell, Shell Script, Azure CLI, Transact SQL, UNIX
* **Databases:**Azure SQL Database, Azure data warehouse, Azure Data azure data factory, Azure Data Synch, Elastic Pools, SQL Server 2019, SQL Server 2017, SQL Server 2016, Microsoft Azure VM, Business Intelligence (BI), Oracle 19g/11g, MYSQL, MS Access, OLAP, OLTP.

**PROFESSIONALEXPERIENCE:**  
**Pepsico, TX October 2022 – Till Date**

**Azure Data Engineer**

* Designed and executed pipelines, data flows, and intricate data transformations and manipulations utilizing Azure Data Factory and PySpark on Databricks.
* Implemented data ingestion into one or more Azure Services, such as Azure Data Lake, Azure Storage, Azure SQL, and Azure DW, with data migration processes in Azure Databricks.
* Utilized Data Flow Debug to construct efficient Azure Data Factory data flow pipelines, and improved performance by implementing optimization options, including effective partition utilization during transformations.
* Constructed intricate ETL pipelines in Azure Data Factory using mapping data flows with multiple input and output transformations.
* Experienced in working with Azure BLOB and Data Lake storage, as well as loading data into Azure SQL Synapse analytics.
* Familiar with the Azure SQL Database Import and Export Service.
* Incorporated Azure Key vault as the central storage for secrets, referencing them in both Azure Data Factory and Databricks notebooks.
* Developed a common SFTP download/upload framework using Azure Data Factory and Databricks.
* Designed ETL pipelines in Databricks using notebooks, Spark DataFrames, Spark SQL, and Python scripting.
* Created Databricks job workflows that extract data from SQL Server and upload files to SFTP using PySpark and Python.
* Built the logical and physical data models for Snowflake, based on required changes.
* Determined virtual warehouse sizing for Snowflake for various types of workloads.
* Implemented Spark applications with PySpark and Spark SQL for data extraction, transformation, and aggregation from multiple file formats.
* Accountable for estimating cluster size, monitoring, and troubleshooting Spark Databricks clusters.
* Wrote Spark code in Scala and Spark SQL/Streaming for faster data processing.
* Implemented Spark Streaming to break down streaming data into batches for processing in Spark Engine.
* Worked on ETL pipelines outside of the data warehouse using a combination of Python and Snowflake Snow SQL, writing SQL queries against Snowflake
* Participated in branching, tagging, and release activities on GitHub version control.

**Charles Schwab, TX October 2021 – September 2022**

**Azure Data Engineer**

* Analyze, design and build modern data solutions using Azure PaaS service to support visualization of data. Understand current production state of application and determine the impact of new implementation on existing business processes.
* Extract Transform and Load data from source systems to Azure Data Storage services using Azure Data Factory. Data Ingestion to one or more Azure Services- (Azure Data Lakes, Azure Storage, Azure SQL) and Snowflake.
* Created Pipelines in ADF using Linked Services/Datasets/Pipeline/to Extract, Transform and load data from different source like Azure SQL, Blob storage, Azure SQL Data warehouse.
* Developed JSON Scripts for deploying the Pipeline in Azure Data Factory (ADF) that process the data using the SQL Activity.
* Have good experience working with Azure BLOB and Data Lake Storage and loading data into Snowflake.
* Hands-on experience on developing SQL Scripts for automation purpose.
* Worked on Azure Cloud services and Snowflake development.
* Writing SQL Scripts to extract the data from Database and for Testing Purposes.
* Extracted the data from relational databases using the data factory and stored the data in data lake.
* Stored the data in the data lake and access the data from the data lake.
* Worked on the data factory to copy data from on prem relational data bases into data lake.
* Developed script to decode the data and transforms the data secured and a script for receiving unstructured data from flume and structured data from RDBMS.
* Actively participate in scrum calls, story points, estimates, and own the development piece.
* Analyze the user stories and understand the requirements and develop the code as per the design.
* Developed test cases, perform unit testing and integrating testing.

**Shell, India July 2018 – June 2021**

**Azure Data Engineer**

* Design and implement end-to-end data solutions (storage, integration, processing, visualization) in Azure.
* Implement ETL and data movement solutions using Azure Data Factory, SSIS
* Migrate data from traditional database systems to Azure SQL databases.
* Implement ad-hoc analysis solutions using Azure Data Lake Analytics/Store, HDInsight
* Design and implement streaming solutions using Kafka or Azure Stream Analytics
* Experience managing Azure Data Lakes (ADLS) and Data Lake Analytics and an understanding of how to integrate with other Azure Services. Knowledge of U-SQL and how it can be used for data transformation as part of a cloud data integration strategy
* Work with similar Microsoft on-prem data platforms, specifically SQL Server and related technologies such as SSIS, SSRS, and SSAS
* Extract Transform and Load data from source systems to Azure Data Storage services using a combination of Azure Data Factory, T-SQL, Spark SQL and U-SQL Azure Data Lake Analytics. Data ingestion to one or more Azure services (Azure Data Lake, Azure Storage, Azure SQL, Azure DW) and processing the data in Azure Databricks.
* Recreating existing application logic and functionality in the Azure Data Lake, Data Factory, SQL Database and SQL Data Warehouse environment. experience in DWH/BI project implementation using Azure DF
* Involved in designing Logical and Physical Data Model for Staging, DWH and Data Mart layer.
* Created POWER BI Visualizations and Dashboards as per the requirements.
* Used various sources to pull data into Power BI such as SQL Server, Excel, Oracle, SQL Azure etc.
* Engage with business users to gather requirements, design visualizations, and provide training to use self-service BI tools.
* Participating in Technical Architecture Documents, Project Design, and Implementation Discussions
* Azure Automation through Runbooks Creation, Migration of existing .PS1 scripts, Authorizing, Configuring, Scheduling.

**KPIT, India March 2016 – June 2018**

**ETL Developer**

* Design, Develop ETL process and create UNIX shell scripts to execute Teradata SQL, BTEQ, jobs.
* Performed bulk data load from multiple data source (ORACLE 8i, legacy systems) to TERADATA RDBMS using BTEQ, TPT, Fastload, Multiload and TPump.
* Assisted DBA to Create tables and views in Teradata, Oracle databases.
* Dealt with Incremental data as well Migration data to load into the Teradata.
* Enhanced some queries in the other application to run faster and more efficiently.
* Used various Teradata Index techniques to improve the query performance.
* Developed the shell scripts to automate the Call Detail Records – Informatica processes and subsequent concatenation of load ready files.
* Extensively used Informatica Client tools - Source Analyzer, Warehouse Designer, Mapping Designer, Mapplet Designer, and Informatica Workflow Manager.
* Developed Source to Target Mappings using Informatica PowerCenter Designer from Oracle, Flat files sources to Teradata database, implementing the business rules.
* Used BTEQ and SQL Assistant (Query man) front-end tools to issue SQL commands matching the business requirements to Teradata RDBMS.
* Modified BTEQ scripts to load data from Teradata Staging area to Teradata data mart.
* Updated numerous BTEQ/SQL scripts, making appropriate DDL changes and completed unit and system test.
* Created series of Macros for various applications in TERADATA SQL Assistant.
* Responsible for loading data into warehouse from different sources using Multiload and Fastload to load millions of records.
* Performed tuning and optimization of complex SQL queries using Teradata Explain.