***Name: Rohan Srivastwa***

***Role: Azure Data Engineer***

**Summary**

* 5.5 years of Professional experience in Data Engineering/ETL working on Azure Tech Stack with expertise in PySpark, Python and SQL.
* Developed Data pipelines from various sources like SQL Server, Oracle DB, SFTP to ingest data into ADLS Gen2 based on Medallion Architecture.

**Tools & Technologies**

* **Azure Databricks (ADB)**
* **PySpark**
* **Spark SQL (SQL)**
* **Python**
* **Excel advanced**
* **Azure Data Factory (ADF), Azure Synapse Analytics**
* **Azure Data Lake Storage (ADLS)**
* **Azure SQL Database**
* **Modern Data Warehouse**
* **Snowflake**
* **PowerBI**
* **Delta tables**
* **Azure Log Analytics**
* **Azure DevOps**
* **GitHub**

**Professional Experience**

***#01 Project***

***Role:*** Azure Data Engineer

***Duration:*** 1.5 year

**Project Description:**

This project involved refactoring legacy system tables to a modern Medallion architecture using Azure Databricks, GitHub for code development, and Snowflake for data warehousing. I created and validated procedures while employing PySpark, Spark SQL, and other tools for a seamless migration.

**Roles and responsibilities**

* Worked as data engineer for refactoring tables from legacy system to modern system based on Medallion architecture using Azure Databricks, Databricks workflows for orchestration.
* Created feature branches to develop code from GitHub repos, committed the changes and merged the PR using GitHub.
* Created silver and gold procedures using Databricks for assigned tables using PySpark, Spark SQL based on standard and user defined libraries.
* Validated completeness and equivalency check for each table migrated to Unity Catalog of databricks.
* Migrated tables to snowflake from Unity Catalog as an alternate Data Warehouse.
* **Tools used –** Azure Databricks, GitHub, Azure DevOps, VS code, PySpark, Python, Spark SQL.

***#02 Project***

***Role:*** Azure Data Engineer

***Duration:*** 2 year

**Project Description:**

I led the development of a robust Data Ingestion Framework that seamlessly collected data from diverse sources (SQL Server, Oracle, SFTP, SAP) into ADLS Gen2 using a code-driven approach on Databricks. This initiative included the creation of delta tables, a Data Quality Framework for consistency, logging, monitoring, and orchestration, ensuring data integrity and performance. Email alerts and Azure DevOps integration further enhanced the project's efficiency and reliability.

**Roles and responsibilities**

* Worked on Data Ingestion Framework to ingest data from multiple sources (SQL Server, Oracle, SFTP, SAP) to ADLS Gen2 using code driven framework built on databricks.
* Created delta table for each table ingested to ADLS Gen2
* Worked on creating Data Quality Framework to ensure data ingested has no discrepancy.
* Implemented logging & monitoring in the code for auditing and performance monitoring.
* Email alert sending system in case of any code failure and a final data quality report.
* Orchestrated the pipeline using Databricks workflows.
* Used Azure DevOps to create PR to deploy code to production.
* **Tools used** – Azure Databricks, Azure Log analytics – App insights, Azure DevOps.

***#03 Project***

***Role:***

***Duration:*** 1.5 years

**Project Description:**

I led the creation of data pipelines, ingesting data from APIs like Facebook Graph into a Medallion model in ADLS Gen2, with a focus on solution architecture. Additionally, I facilitated the Data Science team's needs, conducted API testing, and actively engaged with clients to ensure project success, utilizing Azure Databricks, Azure Synapse Analytics, and Azure Data Factory.

**Roles and responsibilities**

* Worked as a data engineer to build pipelines to ingest data from API sources – Facebook Graph API to ADLS gen2 based on Medallion model – Bronze, Silver & Gold layer.
* Responsible for creating Solution architecture for Data Engineering workflow.
* Performed API testing for data using POSTMAN.
* To facilitate Data Science team with Solution Architecture and ad-hoc requirements related to data engineering. Ex- Databricks setup, Cluster configuration etc.
* Spearheaded Client calls for requirements and updates with Weekly Presentations
* **Tools used** – Azure Databricks, Azure Synapse analytics, Azure data factory.

**Certifications:**

* DP- 203 – Azure data engineer certification
* AZ- 900 – Azure Fundamentals

**Education**

* Bachelor of Engineering, SIT, Tumkur, Karnataka.