|  |  |  |
| --- | --- | --- |
| Ravi Sankar Karri*Senior Data Engineer*  *Certified Azure Data Engineer / Snowflake* | Microsoft Fabric certification thoughts - Kevin ChantSnowflake SnowPro Certification | [rskarri.17@gmail.com](mailto:rskarri.17@gmail.com)(480)-758-0896 [linkedin.com/in/rskarri](http://www.linkedin.com/in/rskarri) |

# Career Objective:

# A seasoned Data and Analytics professional with over 20 years of experience, I am seeking a challenging role where I can leverage my extensive expertise in cloud technologies, data engineering, and process optimization to drive business success and innovation.

# Summary:

Experienced Data and analytics professional with over **20+ years** of proven expertise in designing, planning, and delivering complex technical solutions on-premises and on Cloud. Skilled in process improvement, automation, and leveraging various technologies to achieve business goals. Demonstrated leadership in introducing and evangelizing software design patterns, coding standards for various products across technical teams for better quality, scalability, and maintenance. As a **Senior Data Engineer** for the last 8+ years, I specialize in Azure Data Factory, Azure Databricks, Snowflake, Azure Synapse DW, Microsoft Fabric. Outstanding analytical thinking, logical reasoning, communication, interpersonal skills, and technical skills.

# Core Competencies:

Technical Architecture | Data Engineering | Software Development | Team Development & Leadership | Distributed Teams | Agile

|  |  |
| --- | --- |
| **Cloud Services**  Azure Cloud, Azure Blob/Data Lake Storage, Azure Data Factory, Azure Databricks**,** Delta Lake**,** Azure Synapse, Microsoft Fabric, Dremio, Logic Apps**,** Power Platform, Azure SQL Database, PowerBI, Snowflake, Azure DevOps | **Database**  Postgres SQL, SAP HANA, MS SQL, MS SQL MI, Oracle, Mongo DB, Azure Cosmos DB |
| **Programming**  Java, Python, PySpark, Pandas, JavaScript,  Servlets, JSP, jQuery, Angular, SQL | **Web Technologies**  HTML5, CSS, GitHub, Adobe XD |
| **SAP Stack**  HANA DB, HANA Modeling, SAP BODS, SAP SLT, Enterprise Portal, OData | **Operating Systems**  Microsoft Windows, macOS, Linux, Shell  Scripting |

# Professional Experience:

# QBE DataInsights (QBE Insurance)– Senior Azure Data Engineer USA, New York, NY *Sep 2024 – Present*

# QBE DataInsights analytics solution for QBE Insurance Group Limited, a multinational insurance and reinsurance company offering commercial, personal, specialty, and risk management products.

# Solution- This solution ingested and transformed data from multiple critical transactional and master data sources, creating reconciled Data Marts to support key business dashboards and reports. Focused on delivering insights across commercial, personal, specialty, and risk management insurance products by leveraging Azure Data Factory, Databricks, PySpark, and Snowflake for optimized data processing, integration, and business intelligence.

# • Connecting to On-Prem Sources: Established a pull mechanism where source connections were enabled via Azure Data Factory through Integrated Runtime for both on-premise/QBE-owned and third-party sources (e.g., Majesco). Sources pushing data to Network File Servers were picked up by Azure Data Factory through Integrated Runtime.

# • Developed main/parent data pipelines scheduled based on defined frequency and timelines using Azure Data Factory.

# • Utilized Databricks activities to execute custom PySpark codes leveraging Accenture Intelligent Data Foundation assets, facilitating data movement from raw to clean zones through spark-submit jobs and driver functions tailored by configuration files.

# • Managed data processing and movement from the Atomic to Provisioning Zone by loading data into Snowflake Data Warehouse/SQL Database, creating summary tables for optimized storage.

# • Configured Power BI dashboards on the provisioning layer to report key performance indicators (KPIs) and metrics.

# • Ensured secure connections to on-prem sources through private peering via ExpressRoute.

**Valcourt data migration - Microsoft Fabric data engineer  
USA, Houston, TX***April 2024 – Sep 2024*

# At Valcourt Building Services, we implemented Microsoft Fabric to modernize data management, enhance analytics capabilities, and streamline operational efficiency. The solution aimed to consolidate data from multiple sources, improve reporting accuracy, and enable real-time insights for better decision-making. By leveraging Microsoft Fabric’s Lakehouse architecture, we integrated data from various operational systems, automated ETL workflows.

# Solution: As part of Data needs to be migrated from each branch into Salesforce and NetSuite, extract and transform data as per business standards before load into above target systems by using Microsoft Fabric platform.

* Designed and implemented Microsoft Fabric architecture, including OneLake, Data Factory, and Synapse Analytics for a unified data platform.
* Developed and optimized ETL pipelines using Microsoft Fabric Data Factory to automate data ingestion and transformation from SAP, CRM, and operational databases.
* Built a scalable Lakehouse architecture with bronze, silver, and gold layers, ensuring efficient data processing and management.
* Optimized data storage and query performance using Delta tables, partitioning to enhance processing speed and efficiency.
* Collaborated with business stakeholders to define key data requirements, ensuring alignment with organizational goals and decision-making needs.
* Led the deployment of Microsoft Fabric components in production environments using CI/CD pipelines and Azure DevOps for seamless integration.
* Provided ongoing support and troubleshooting, resolving data quality issues, improving performance, and implementing enhancements for continuous improvement.

**Shell Trading – Senior Azure Data Engineer  
USA, Houston, TX***Nov 2022 – Mar 2024*

# As part of the Shell Energy Brazil Gas business, designed and implemented an Azure data platform solution for the Gas Trading desk that will process daily operations data, including monthly volumes , daily production volumes, wellhead pressure, and temperature onto the data platform and make this data available to the business in a standardized way to provide insight on volumetric movements and enable the business to perform optimization analytics via Power BI. This solution supports downstream operations by enhancing trading efficiency and market-facing analytics.

**Solution**- Analyze, design, and build modern data solutions using Azure PaaS services to handle variety of data formats into a data platform and support visualization of data via Power BI. Understand current Production state of application and determine the impact of new implementation on existing business processes.

* Extensive ADF development building pipelines for scanning across multiple emails and attachments (excel, pdf), SharePoint documents and created rules using Graph API and ingest into ADLS Gen2.
* Implemented Power Automated workflow to automate the process of downloading data from an online source processing the data, and sending an email notification with the processed data attached.
* Developed notebooks using PySpark, Pandas and Spark-SQL for data extraction, transformation, and aggregation from multiple file formats for analyzing & transforming the data to uncover insights into the customer usage patterns.
* Image processing- Processed images using Python Imaging library and converted to Spark data frames for further processing. The images were chemical composition reports sent by field inspectors.
* Developed Power BI report for presenting daily Ops data that includes composition analysis as well as estimated, nominated and actualized plant volumes received daily.
* Installed the Snowflake connector in Databricks and send curated data from Databricks to the Snowflake data warehouse.
* Responsible for estimating the cluster size, monitoring, and troubleshooting of the Spark databricks cluster.
* Monitoring, Schedule ADF jobs and troubleshoot the issues using Azure Monitor.
* Worked with DevOps engineers and promoted code from development to production following client guidelines.
* Conducting code reviews for team members to ensure proper test coverage and consistent code standards.
* Worked in Agile development environment with offshore and onshore team members.

**Bridgestone – Senior Azure Data Engineer**

**USA, Phoenix, AZ**

*Feb 2022 – Oct 2022*

Bridgestone is one of the largest tire manufacturing company and having retread manufacturing plants across the globe and having more than 500 dealers.

**Solution -**By connecting these various types of data from Microsoft Dynamics and other sources, Bridgestone can gain comprehensive insights into its operations, supply chain, customer relationships, and financial performance. This data integration supports data-driven decision-making, operational efficiency, and continuous improvement across the organization.

* Set up linked services in ADF for both ADLS Gen2 and Snowflake.
* Develop ADF pipelines to read data from ADLS Gen2 folders and load it into Snowflake.
* Use ADF Copy Data activities to ingest data from ADLS Gen2 to Snowflake, ensuring efficient data transfer.
* Create transformation scripts in Snowflake to handle Slowly Changing Dimension Type 1 (SCD1) updates by overwriting existing records.
* Develop transformation scripts in Snowflake for SCD2 to maintain historical data by creating new records with versioning.
* Apply data masking techniques in Snowflake to protect sensitive information after data ingestion.
* Create and manage Snowflake compute resources (virtual warehouses) effectively to balance performance and cost.
* Utilize Snowflake's storage capabilities to efficiently manage data, including compression and partitioning.
* Implement an archival strategy for old data using Snowflake’s Time Travel and data retention features to manage storage costs.
* Set up monitoring for ADF pipelines and Snowflake operations, ensuring timely detection and resolution of issues, and perform regular maintenance activities for optimal performance.

**Fulton Bank – Senior Azure Data Engineer**

**IND, Hyderabad**

*Feb 2021 – Dec 2021*

The Fulton Data Marketplace (FDM) involve migrating retail customers, accounts, balance statements, loans and credit card transactions from on-premise to Azure cloud.

**Solution-** Fulton Data Marketplace (FDM) project represents a strategic investment in modernizing Fulton Bank's data infrastructure and unlocking the full potential of its data assets. By migrating to the Azure cloud and leveraging Power BI for reporting and analytics, Fulton Bank is poised to drive innovation, enhance customer experiences, and achieve operational excellence in the ever-evolving banking landscape.

* As part of data migration, worked on Hydration and Delta Hydration activities.
* Work with Azure Meta Data framework & configuration, modified metadata files.
* Handled various sources- Salesforce, Oracle, SQL Server, file systems.
* Developed Synapse pipelines to ingest data into the data lake.
* Used various file formats avro, parquet and delta lake in the data lake.
* Created Databricks/Spark Notebooks
* Developed Control scripts for data validations.
* Collaboratively work with teams using Teams & Azure Dev

**Stanley Black & Decker – Senior Azure Data Engineer**

**IND, Hyderabad**

*Feb 2020 – Jan 2021*

Stanley Black & Decker has been setting the standard for innovation and design of power tools, outdoor yard care equipment and home products. The inventor of the first portable electric drill with pistol grip and trigger switch, Stanley Black & Decker has evolved from a small machine shop in Baltimore, Maryland to a global manufacturer with a broad line of products.

**Solution**- extract data from SAP, MongoDB, and SQL databases to the Azure Cloud Platform represents a transformative step towards unlocking the full potential of data-driven innovation and agility at Stanley Black & Decker. By harnessing the power of advanced analytics, cloud computing, and integrated data management, Stanley Black & Decker is poised to drive operational excellence, enhance customer experiences, and maintain its position as a leader in innovation and design within the global marketplace.

* Created Synapse Workspace/SQL Pools.
* Connect to External Data Source outside of your Synapse Workspace
* Create reusable and scalable Synapse pipelines.
* Extract data from SAP tables, MongoDB tables and CSV files using Synapse pipelines.
* Designed Synapse pipelines and Data Flows for Initial Load and Delta Load.
* Created and Scheduled Azure Synapse Pipeline Daily Triggers
* Stored Data in Azure Data Lake/Azure Synapse analytics and Power BI reports are consuming data from ADL/ Synapse.
* Responsible for providing Sizing & Cost Estimates for required Azure services.
* Configure DevOps CI/CD pipelines migrate ARM templates from different environments.
* Monitoring and troubleshoot Azure Synapse pipelines.
* Installing Self-hosted Integration runtime/configuration
* Created Azure Key vaults/securely connect source systems.
* Created logic app to send e-mail alerts to users if any Azure Synapse pipeline job fails.

**Unilever – Azure Data Engineer**

**IND, Hyderabad**

*Sep 2017 – Dec 2019*

Unilever is a multinational consumer goods company renowned for its diverse portfolio of brands spanning food, beverages, cleaning agents, and personal care products and operates in over 190 countries, emphasizing sustainability and innovation.

**Solution** - enhance factory performance monitoring and optimization across Unilever's global manufacturing network. Leveraging SAP data, the solution provides real-time visibility into production volumes, efficiency, and capacity utilization, enabling proactive decision-making and continuous improvement initiatives.

* Load data from Universal Datalake into Azure Databricks.
* Developed Spark applications using Pyspark and Spark-SQL for data extraction, transformation, and aggregation from multiple file formats for analyzing & transforming the data to uncover insights into the customer usage patterns.
* Estimated the cluster size, monitoring and troubleshooting ­of the Spark databricks cluster.
* Responsible for reviewing Design Documents and code, Test coverage and its alignment with technical objectives.
* Responsible for leading the specification of complex customer requirements in one's own area of expertise.
* Responsible for leading/ participating in technical reviews to ensure quality parameters are being met.
* Create reusable and scalable azure data factory pipelines.
* Created and Scheduled Azure Data Factory Pipeline Daily Triggers
* Monitoring and troubleshoot Azure data factory pipelines

**Education**:

* M.Sc. Computer Science

*Andhra University, India*

**Professional Certifications:**

* Certified Azure Data Engineer – DP-203

*Microsoft*

* Certified – SnowPro Core

*Snowflake*

* Certified – JAVA Developer

*Sun*