

## Shekhar Gund

Azure Data Engineer & Power BI Developer

☎: - +91-9859490333

E-mail: - [shekhargund0@gmail.com](mailto:shekhargund0@gmail.com)

LinkedIn - <https://www.linkedin.com/in/shekhargund/>

### Professional Summary :

- 4+ years of professional experience in information technology as a Data Engineer with an expert hand in **Database Development**, **ETL Development**, **Data modeling**, **Report Development**, and **Big Data Technologies**.
- Experience in Software Industry as **Azure Data Engineering Services** such as (**Azure Data Factory**, **Azure Data Bricks**, **Azure SQL**, **Azure Synapse**, **Azure BLOB**, **Azure Data Lake gen-2**), **Azure Cosmos DB**, **Azure Streaming Jobs**, **MS Power BI**, **Python**, **PySpark**, **Spark SQL** and **MS SQL**, **Oracle Database Development & Support**.
- Creating Pipelines in **ADF** using **Linked Services/Datasets** to Extract, Transform, and load data from different sources like **Azure SQL**, **Blob storage**, **ADLS**, **Oracle**, and **My SQL**.
- Good understanding of various Activities (**Copy**, **Metadata**, **Delete**, **Lookup**, **Stored procedure**, **Until**, **For Each**, **IF**, and **Execute Pipeline**).
- Hands-on experience in creating **Mapping data flows** as per the requirement.
- Configured and implemented the Azure Data Factory **Triggers and scheduled** the ADF Pipelines.
- Experience developing and executing **Databricks notebooks** using languages like **PySpark** and **Spark SQL** to perform data processing, transformation, and analysis tasks on large datasets.
- Implemented **reusable Data Bricks notebooks** and incorporated them across all notebook projects for consistency and efficiency.
- Worked with multiple file formats, including **CSV**, **JSON**, and **Parquet** for data processing and storage.
- Worked in Extracting, transforming, and loading the data from different formats like **JSON** and exposing it for interactive queries using **Spark SQL**.
- Monitor the performance of **pipelines**, **troubleshoot incidents**, and provide support to end-users.
- Created, monitored, and restored **Azure SQL database** and performed migration of **On-Prem to Azure SQL database**.
- Experience in integrating **Azure Databricks** with **Azure Key Vault** using secret scopes, enabling secure access and management of sensitive information.
- Develop Batch processing solutions by using **Data Factory** and **Azure Data Bricks**.
- Experienced in **IoT hub** for **Azure Streaming Job** real-time data processing.
- Experienced in building a **Data Warehouse** in **Azure platform**, using **Azure data bricks** and **data factory**.
- Implement **Azure Data Bricks** clusters, notebooks, Jobs, and Auto Scaling.
- Experience in **Databricks** to provide a reliable and scalable platform for data pipelines, data lakes, and **data platforms**. Manage the full data journey, so you can ingest, process, store, and expose data throughout your organization
- Implemented **Spark Scripts**, and **Spark SQL** to access tables into Spark for faster processing of data.
- Designed **Azure Logic apps** application to send pipeline success/failure alert email notifications etc.
- Implemented **ETL pipelines** using **Azure Data Factory**, extracting data from various sources, transforming it in **Databricks**, and loading it into **Azure Data Warehouse**.
- Managed large datasets in **Delta Tables** using the **OPTIMIZE** command, enabling **ACID** transactions, schema evolution, and enhancing query performance through **Z-Ordering**.
- Managed **Unity Catalog** to streamline data governance, control access, and unify data assets across multiple environments.
- Worked with **Managed** and **External Delta Tables**, leveraging features like **tombstoning**, **data versioning**, **time travel**, and **vacuum** for efficient data management.
- Assisted individual teams in setting up and managing their repositories in **Azure DevOps** and **GitHub**, maintaining code quality, and configuring jobs to leverage **CI/CD pipelines** across **DEV**, **UAT**, and **PROD** environments.
- Expertise in **Data warehousing** concepts, **Dimensional Modeling**, **Data Modeling**, **OLAP** and

OLTP systems.

- Good Experience on Azure SQL Database, creating Tables and **Views**.
- Good experience in **Power BI Desktop, Power BI Service** and writing **DAX** functions.
- Proficient in writing complex and advanced **DAX calculations** using MS Power BI Desktop like **aggregation, cumulative total, and calculations with time intelligent function**.
- Experience in providing row-level securities (**RLS**) for reports/dashboards by assigning roles.
- Hands-on experience in creating Power BI Reports and Dashboards connecting various data sources in all modes like **Import, Direct Query, and Live** Connections.
- Worked with dimensional modeling and created **star schema** for the given system and good knowledge of the development of Snowflake Schema.
- Experience in Creating Filters, Hierarchies, Interactions, Joins, Parameters, grouping and Formatting.
- Knowledge in **performance tuning** and knowledge to work according to the Work Request and Change Request.
- Good Experience in Databases PL/SQL and MS SQL Server.
- Good experience in writing complex **MS SQL Queries** and concepts like DML, DDL operations, views, Index, stored procedures, and joins.
- Good work experience in **performance tuning on SQL** queries.
- Experience in Agile project management methodologies and proficiency in using **JIRA**, a leading project management and issue tracking tool.
- Having worked in various domains like **Retail, Sports, Pharma, and Healthcare Insurance**.

#### Education :

- Education: Master of Business Administration (Business Analytics) from Savitribai Phule Pune University (**First Class with Distinction**)

#### Work Experience:

- Worked at **Explarity Solution Pvt Ltd.** (Sept\_2022 - Current)
- Working at **Mindstix Software lab.** (Sept\_2020 – Aug\_2022)

#### Technical Skills:

- Cloud Services : Azure Data Factory, Azure Data Bricks, Azure Synapse Analytics, Azure Data Lake storage, Azure SQL Database, Azure Blob Storage, Azure Cosmos DB, Azure Streaming Analytics.
- Databases : MS-SQL Server, Oracle 11g, Azure SQL
- Programming : Python, SQL, PY-Spark
- Reporting Tools : Power BI, Tableau, Looker
- Working Tools : JIRA, GitHub
- Other ETL : Informatica Cloud (IICS), Alteryx, Talend open Studio.
- Domains : Retail, Pharma, Sports, Healthcare Insurance

#### Project - 1

Title	:	Data Migration and Analytics Solution using Azure Services and Power BI
Role	:	Data Engineer
Duration	:	Aug 2023 to current
Technologies	:	Azure data factory, Azure Data bricks, Azure Data Lake, Azure SQL database SQL Server, MySQL DB, Azure Synapse Analytics, Power Bi

#### Description:

The Landmark Group is a leading retailer with a diverse portfolio of brands like Centrepont, Max Fashion, Home Centre, Baby Shop, Lifestyle, and more, spanning fashion, home, groceries, and electronics. Migrated data from on-premises MSSQL Server to the cloud, leveraging Azure Data

Factory for seamless data flow and Databricks notebooks for data validation, quality checks, and deduplication. Built a centralized data warehouse system to store and analyze sales, profits, margins, and customer information, enabling insights into product trends, regional sales, and top-performing products to support strategic decision-making and profit growth.

### Roles & Responsibilities:

- Analyze, design, and build Modern data solutions using Azure service to support visualization of data. Understand the current Production state of application and determine the impact of new implementation on existing business processes.
- Data Ingestion to one or more **Azure Services** - (**Azure Data Lake, Azure Storage, Azure SQL, Azure DW**) and processing the data in Azure Databricks.
- Extract Transform and Load data from source systems to **Azure Data** Storage services using a combination of **Azure Data Factory, Spark SQL, and Azure Data Lake Analytics**.
- Enhanced and optimized **Spark/python** jobs to aggregate, group, and run data mining tasks using the Spark framework.
- Worked on migration of data from On-prem SQL server to Cloud databases (**Azure Synapse Analytics (DW) & Azure SQL DB**).
- Designed Data Stage ETL jobs for extracting data from heterogeneous source systems, transforming and finally loading into the **Data Warehouse**
- Experience in developing Spark applications using **Spark-SQL** in **Databricks** for data extraction, transformation, and aggregation from multiple file formats for Analyzing& transforming the data to uncover insights into customer usage patterns.
- Primarily involved in Data Migration using **SQL, SQL Azure, Azure Storage, and Azure Data Factory**.
- Responsible for estimating the cluster size, monitoring, and troubleshooting of the Spark **Data Bricks** cluster
- Created relationships between different tables to create of **data model**.
- Created New Columns, New Measures, and New Tables over data models using **DAX functions**.
- Used “**LINE CHART**” for sales trend, “**BAR CHART**” for regional sales, “**COLUMN CHRT**” for sales profit, “**CARD VISUAL**” for total sales, “**COMBO CHART**” for comparing last year's sales, this year's sales and total sales variance, “**PIE CHART**” to compare Country wise profit.
- Used **bookmarks, buttons**, and actions to create Manager View and Business View.
- Designed different types of reports like **Drill Down, Drill Through, and Sync Slicers**. Develop **Visual Level Filters, Report Level Filters, and Page Level Filters** in Power BI.
- Created Microsoft Power BI Dashboards in **Power BI Service** using Reports.
- Installed and Configured **Data Gateways** to Refresh and Schedule the Refresh of the Dataset.
- Managed access to Reports, Dashboards, and Data for individual users using roles.
- Implemented **Row Level Security** as part of Security as an aspect of ensuring data confidentiality and integrity.

### Project - 2

Title	: Healthcare Data Integration and Analytics: Empowering HMOs with Azure Technologies
Client	: HMOs (Health Maintenance Organizations)
Environment	: Azure data factory, Azure Data bricks, Azure Data Lake, SQL server, PowerBi
Duration	: Sept 2022 To July 2023
Role	: Data Engineer

### Description:

Healthcare insurance project for an organization providing health insurance to employer groups, individuals, and families. The project involved integrating Health Maintenance Organizations (HMOs) with a network of doctors and hospitals to offer medical coverage. Members receive ID cards upon enrollment, which are used to verify the validity of their policies. Members can visit a service provider (hospital, doctor, pathology lab), where their eligibility for services is checked. Once the service is

provided, the provider or member files a claim, which is validated by the insurer. Upon validation, the insurer reimburses the provider or member for the medical expenses incurred.

#### **Roles & Responsibilities:**

- Understanding the user **requirements** and existing system.
- Design and implement data pipelines using **Azure Data Factory** for the extraction, transformation, and loading (ETL) of healthcare data from various sources into **Azure Data Lake**.
- creation of Pipelines with different activities like **Copy Data, Lookup, Switch, Filter, Metadata, For Each, Pipeline, Validation, set variable, Wait, Until, etc. and Data Flow activity and its transformations like Select, Join, Derived column, Lookup, alter row, exists, conditional split** etc.
- Connect to **data lake** with **mount point** and prepare **delta tables** with **PySpark**.
- Utilize PySpark for data processing and transformation tasks within the **Azure Databricks** environment.
- Design and build data bricks notebooks with **PySpark** and **spark SQL** to move & transform data for data ingestions from source to Target files
- Implement data quality checks and ensure data integrity throughout the entire data lifecycle.
- Involved in deploying Notebooks into the development environment using GIT.
- Develop interactive and insightful **Power BI dashboards** and reports for visualizing key performance **indicators, trends, and metrics** related to health insurance.
- Collaborate with business analysts to understand reporting requirements and design data visualizations that meet the needs of various stakeholders.
- Created different visualizations like **Lines, Bars, Histograms, Scatter plots, waterfalls, Pie, Area, Donut Charts and Bullet graphs, Heat Maps Tree Maps, etc.**
- Implement data-driven insights and enable self-service analytics for end-users.
- Ensure the security and access control of **Power BI reports**, considering sensitive healthcare data.

#### **PERSONAL DETAILS:**

- Name : Shekhar Banshi Gund
- Marital status : Married
- Language : English, Hindi, Marathi
- Address : New Sangavi, PCMC, Pune India.

#### **Declaration:**

I hereby certify that all the information provided above is true to the best of my Knowledge

Place: Pune

**Shekhar Gund**