Shekhar Gund

Azure Data Engineer & Power BI Developer **☎: - +91-9011876027**

**LinkedIn -** [**https://www.linkedin.com/in/shekhargund/**](https://www.linkedin.com/in/shekhargund/) **E-mail: -gundshekhar99@gmail.com**

**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 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 **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 reliable and scalable platform for data pipelines, data lakes, and **data platforms**. Manage full data journey, so you can ingest, process, store and expose data throughout your organization
* Implemented **Spark Scripts**, **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.
* 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 on 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 on Databases PL/SQL and MS SQL Server.
* Good experience of 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 **Mindstix Software lab**. (2020 - 2022)
* Working at **Explarity Solution**. (2022 - Current)

## 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 client group is involved in the retailing of apparel, footwear, consumer electronics, cosmetics & beauty products, home improvement, and baby products. The group also has interests in hospitality & leisure, healthcare, and mall management. The group has several in-house brands and also works with other brands, acting as a retailer. This Project is involved in migrating the data from the on-premises MSSQL server to the Cloud and Applying the Azure activities, scheduling the pipelines by scheduling triggers. Azure Data Factory facilitates seamless data flow, while Databricks notebooks play a key role in executing essential tasks. These tasks encompass checking for duplicates, validating date formats, ensuring null checks, and performing row-level quality checks. It has data that maintains information like sales, profits, margin, and contacts of customers. Using this existing data, we analyze their business moment like different product sales trends, sales profits, regional sales, and top N products sales. This predictive capability proves invaluable for decision-makers, contributing to the growth in profits and strategic business planning.

# 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).**
* For real-time/batch analysis, reporting, optimization, data collecting, and associated tasks, design, manage, and maintain **Azure data pipelines.**
* 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**, SSIS, PowerShell
* Performance analysis of batch jobs by using **Spark** Tuning parameters.
* Responsible for estimating the cluster size, monitoring, and troubleshooting of the Spark **Data Bricks** cluster
* Created views in **Synapse** and connected Synapse to Power BI for visualization.
* Created relationships between different tables to create of **data model**.
* Created New Columns, New Measures, and New Tables over data models used **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 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, PySpark, Azure Synapse, Power BI

Duration : Sept 2022 To July 2023

Role : Data Engineer

**Description:**

Their primary function is to provide Health Insurance to Employer Groups Individuals and family members. The basic idea behind this is that the HMOs tie up with a group of Doctors and Hospitals (Providers of Healthcare) and allow the members to visit one of the Providers within their list. Healthcare Insurance is a medical insurance policy that offers financial coverage for medical expenses when the policyholder/Customer/Member is hospitalized. Once enrolled, an insurer provides ID cards to its members. These will be used to show the validity of the policy taken. Members can go to a service provider (hospital, doctor, Pathology lab) to avail of the service. Provider will check the Eligibility of member for that service. The provider/member then will file a claim to the insurance company. Once, the claim gets validated the insurer will reimburse the amount to the provider/member.

## 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 poin**t 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 Plot, Water Fall, Pie, Area, Donut Charts and Bullet Graph, 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.

**Project - 3** Title : ITC, INDIA: Wells Fargo, EDW migration to Azure

Environment : Azure data factory, Azure Data bricks, Azure Data Lake, SQL server, Azure Synapse, Spark SQL, Power BI

Duration : Dec 2020 To July 2022

Role : Data Engineer

**Description:**

Wells Fargo is investment banking which deals with many investment products. Wells Fargo & Company is an American multinational financial services company with corporate headquarters in San Francisco, California, operational headquarters in Manhattan, and managerial offices throughout the United States and internationally. We create centralized database.

**Roles & Responsibilities:**

* Primarily involved in Data Migration using SQL, SQL Azure, Azure Data lake and Azure Data Factory.
* Build analytic tools that utilize the data pipeline to provide actionable insights into customer acquisition, operational efficiency, and other key business performance metrics.
* Working with Source team to extract the data and it will be loaded in the ADLS.
* Creating the linked service for source and target connectivity Based on the requirement.
* Creating the pipelines and datasets which are deployed in ADF non-restricted
* Once it’s created pipelines and datasets will be triggered based on LOAD (HISTORY/DELTA) operations.
* Based on source (big or small) data loaded files will be processed in AZURE DATABRICKS by applying operations in Spark SQL. Which will be deployed through AZURE DATA FACTORY pipelines.
* Involved in deploying the solutions to QA, DEV, and PROD.
* Involved in setting up the environments for QA, DEV and PROD using VSTS.
* Professional in creating a data warehouse, design-related extraction, loading data functions, testing designs, data modelling, and ensure the smooth running of applications.
* Responsible for extracting the data from OLTP and OLAP using Azure Data factory and Databricks to Data Lake.
* Developed pipelines that can extract data from various sources and merge into single source datasets in Data Lake using Databricks.
* Created connections from different sources from On-Prem and cloud to only source for Power BI reports.
* Coordinate with external teams to reduce data flow issues and unblock team members
* Always actively participate in four ceremonies: Sprint planning meeting, Daily Scrum, Sprint review meeting, and Sprint retrospective meeting.

## PERSONAL DETAILS:

* Name : Shekhar Bansi 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**