**Vishal Kumar**

**Data Engineer**

**Email:** [vishalvodala6@gmail.com](mailto:vishalvodala6@gmail.com)

**Phone:** 913 346-1701

**Summary**

* Driven data practitioner with 7 years of diverse experience from analyzing, data modeling, develop data pipelines, wrangling big data sets (SQL), create BI reports. Flexible and eager to work on robust, complex challenges and game changing insights.
* Hands-on experience in Azure Cloud Services (PaaS & IaaS), **Azure** **Synapse** **Analytics**, SQL Azure, **Data** **Factory**, Azure Analysis services, Key Vault, Azure Data Lake.
* Have Data warehousing experience in Business Intelligence Technologies and Database with Extensive Knowledge in Data analysis, TSQL queries, ETL & ELT Process, Reporting Services (using SSRS, Power BI) and Analysis Services using SQL Server 2012/2014 /2016 **SSIS**, **SSRS** and **SSAS**, SQL Server Agent.
* Hands-on experience in using **Informatica** Power Center for ETL including Hadoop (MapReduce & Hive), Spark (SQL, Streaming), SQL Datawarehouse, **AWS** Redshift, Athena, Lambda, Step Function etc.

**Skills**

* Data Modelling Erwin, MS Visio, Toad Data Modeler
* Query & Database PL/SQL & SQL Server, Oracle, Azure Synapse, Azure PolyBase
* ETL & BI SSIS, Informatica, Data Factory, Talend, SSRS, Tableau, PowerPivot, Synapse Analytics
* Programming SQL, PowerShell, C# or Vb, Pyspark(Python), Scala
* Big Data HDFS, Sqoop**,** Hive, Apache Spark, Databricks, Azure DataLake, S3
* Document/Version Confluence, JIRA, SourceTree - Git, Azure DevOps, BitBucket, Bamboo, Kubernetes

**Work Experience**

**Sr Data Engineer – Microsoft**

**Redmond, WA**  **Oct 2021 to Present**

* Designed and Configured Azure Cloud relational servers and databases analyzing current and future business requirements.
* Worked on migration of data from On-prem SQL server to Cloud databases (**Azure Synapse Analytics** & Azure SQL DB).
* Have extensive experience in creating pipeline jobs, scheduling triggers, Mapping data flows using **Azure Data Factory**(V2) and using **Key Vaults** to store credentials.
* Worked on creating tabular models on Azure analysis services for meeting business reporting requirements.
* Have good experience working with Azure **Blob** and Data lake storage and loading data into Azure SQL Synapse analytics (DW).
* Designed **SSIS** Packages using Business Intelligence Development Studio (BIDS) to extract data from various data sources and load into SQL Server database for further Data Analysis and Reporting by using multiple transformations.
* Worked on creating correlated and non-correlated sub-queries to resolve complex business queries involving multiple tables from different databases.
* Developed business intelligence solutions using **SQL** server data tools and load data to SQL & **Azure** **Cloud** databases.
* Perform analyses on data quality and apply business rules in all layers of data extraction transformation and loading process.
* Perform validation and verify software at all testing phases which includes Functional Testing, System Integration Testing, End to End Testing, Regression Testing, Sanity Testing, User Acceptance Testing, Smoke Testing, Disaster Recovery Testing, Production Acceptance Testing and Pre-prod Testing phases.
* Have good experience in logging defects in **Azure Devops** tools.
* Migrated code (Pyspark, Scala, Spark SQL) from existing **databricks** notebooks into **Synapse** notebooks for Microsoft sales data.
* Migrated SQL code from existing Stored Procs into **Synapse** notebooks.
* Created numerous pipelines using Azure Data Factory to get the data from disparate source systems by using different Azure Activities like Move &Transform, Copy, filter, for each, Databricks etc.
* Created Linked service to land the data from SFTP locations, Blob to Azure Data Lake.

**Data Engineer – Advantasure**

**SouthField, MI**  **July 2017 to Sep 2021**

* Working closely with business users and management to gather requirements and develop work plans.
* Translate complex requirements to data models and qualitative queries, prepare POC’s and design documents.
* Built scalable ETL SSIS packages, **ADF** Data pipelines and evaluate the workflow process to increase the efficiency.
* Developed C# code for custom data validations, transformations, and dynamic SQL in SSIS script task for reusability.
* Fine tuning the existing queries and making sure of proper table designs with indexes and partitions for quicker results.
* Built scalable databases capable of handling **ETL** process using SQL. Developed Databricks spark code to create UDF for reusability. C# script to define custom functions, generate **BCP** script, **REST API** functionalities in SSIS.
* Developing single pseudo code **Pyspark/Scala** to reuse and scale it for similar datasets with varied requirements.
* 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.
* Analyzed existing systems and propose improvements in processes and systems for usage of modern scheduling tools like Airflow and migrating the legacy systems into an Enterprise data lake built on Azure Cloud.
* Implement software enhancements to port legacy software systems to **Spark** and **Hadoop** ecosystems on Azure Cloud.
* Writing Hive SQL queries for data transformation or aggregations, create delta tables and data warehouse in Databricks.
* Involved in writing Databricks read and write streams to extract real time data and load to synapse or redshift.
* Worked on MPP databases like Synapse, Redshift and file formats like ORC, Parquet, and merge operations on delta lake.
* Build data pipelines in Data Factory using SQL procedures, **databricks** notebooks.
* Created ad-hoc reports to users in **Tableau** by connecting various data sources.
* Generated Tableau Dashboards with Quick filters, Parameters and sets to handle views more efficiently.
* Published Tableau Workbooks by creating user filters so that only appropriate teams can view it.
* Imported data from SQL Server DB, Redshift and Azure DB to generate reports in **PowerBI**.
* Installed and configured Power BI Gateways to keep dashboards and reports up to date.
* Creating rowlevel security, integration with PowerBI service and schedule automatic refresh in PowerBI Service
* Used Power BI, Power Pivot to develop data analysis prototype, and used Power View and Power Map to visualize reports.

**Data Engineer – BCBSM**

**Detroit, MI**  **Apr 2016 to Jun 2017**

* Created database objects like tables, views, procedures, triggers, and functions using **T-SQL** to provide definition, structure and to maintain data efficiently.
* Effectively used Informatica parameter files for defining mapping variables, workflow variables, FTP connections and relational connections. Used **Informatica** file watch events to pole the FTP sites for the external mainframe files.
* Used various transformations like Filter, Expression, Sequence Generator, Update Strategy, Joiner, Router and Aggregator to create robust mappings in the Informatica Power Center Designer.
* Created on various SQL window functions, indexes, partitions, dynamic SQL, stored procs, views, and performance tuning.
* Perform in-depth analysis of research information for the purpose to identify opportunities and developed proposals.
* Created a common support platform for both **ETL** & **BI** tasks which reduce his overall annual maintenance budget.
* Created shell scripts to fine tune the ETL flow of the **Informatica** workflows.
* Used debugger in identifying bugs in existing mappings by analyzing data flow, evaluating transformations.
* Building, publishing customized interactive reports and dashboards, report scheduling using **Tableau** server.
* Hands-on experience with Amazon EC2, Amazon S3, Amazon RDS, VPC, IAM, Amazon Elastic Load Balancing, Auto Scaling, CloudWatch, SNS, SES, SQS, Lambda, EMR and other services of the AWS family.
* Created action filters, parameters, and calculated sets for preparing dashboards and worksheets in Tableau. Restricted data for users using Row level security and User filters.
* Developed Tableau visualizations and dashboards using **Tableau Desktop**. Created action filters, parameters, and calculated sets for preparing dashboards and worksheets in Tableau.
* Developed Tableau workbooks from multiple data sources using Data Blending.
* Performed end- to-end Architecture & implementation assessment of various AWS services like Amazon EMR, Redshift, S3. Extensively used AWS Step functions services using various flows and conditions.
* Migrated data from various external applications to AWS services using S3, Redshift, Glue and rest AWS family services. Extensively used AWS Step Functions with various functions and scheduled them through AWS Event Bridge Services.
* Analyzing of Large volumes of structured data using **SparkSQL.**
* Used and Tuned relational databases (e.g. Microsoft SQL Server, Oracle, MySQL) and columnar databases (e.g. Amazon Redshift, Microsoft SQL Data Warehouse)
* Supporting Continuous storage in **AWS** using **S3**, Glacier. Created Volumes and configured Snapshots for **EC2** instances
* Acquired good domain knowledge on medicare, member enrollment, provider, claims, risk, stars and **Hedis**. Created knowledge documents and shared through confluence. Good knowledge in health insurance **Stars** and Hedis program. Knowledge in provider data, member **enrollment**, **medical** claims in MA of health insurance.

**SQL/ETL/BI Developer – Sriven Info Consultants**

**Hyderabad, INDIA** **May 2013 to Dec 2013**

* Interacted with Business Users, Architects and Designed ER diagrams and normalization of database.
* Involved in various Data Analysis process right from Data Collection, Processing & Cleansing.
* Generated **SQL** Scripts to create tables, Views, columns, and constraints as per the requirement.
* Worked on agile project, part of scrum meeting and status of tasks are monitored using **JIRA**.
* Created configuration tables using merge scripts and used in packages to fetch connections, config data based on environment. Performed validations to ensure the tasks are configured correctly.
* Implemented configurable Error Logging and Exception handling to the **SSIS** packages
* Created stored procedure and function to support efficient data storage and manipulation.
* Experience in using for each loop, sequential container, and object type variables to load data in iterative manner using foreach Item Enumerator and foreach **ADO** enumerator.
* Created **VB**.**Net** Script for Data Flow and Error Handling using Script component in MSBI (SSIS).
* Extensively used Extract Transform Loading (ETL) tool of SQL Server to populate data from various data sources to SQL Server. Experience in using Execute Package Task, FTP, Script, and email tasks.
* Created various MSBI(SSIS) packages using data transformations like Execute SQL task, Merge, Aggregate, Sort, and transformation of **XML** based on **XSD** File Experience with designing the ETL process, including data read/write, data transformations, error handling and logging.
* Experienced in enhancing and deploying the SSIS Packages from dev server to production server.
* Created and deployed SSIS packages using various transformations such as Slowly Changing Dimension, Multicast, Merge Join, Lookup, Fuzzy Lookup, Conditional Split, Aggregate, Derived Column, and Data Conversion Transformations. Developed a **SSAS** Cube which is used as the Data Source of **SSRS** Report
* Responsible for creating and maintaining Analysis services objects such as cubes, dimensions, measures. Created report model on **SSAS** cubes as well as changing default configuration on existing cubes.
* Designed tabular, matrix reports, drilldown, drill through, Parameterized and linked reports in SQL Reporting Services (SSRS). Built, Published and Scheduled SSRS Reports for both Dev and Test Environment.

**Education**

**BTech** in Computer Science, Jawaharlal Nehru Technological University, India **2013**

**MS** in Computer Science, New York Institute of Technology, New York, USA **2015**