# Summary

Pranay

pranayjunnuthula67@gmail.com

Highly skilled Data Engineer with over 6 years of extensive experience in designing, building, and optimizing data pipelines and architectures across diverse industries. Proficient in utilizing advanced big data technologies such as Hadoop, Apache Spark, and Kafka, alongside expertise in cloud platforms like AWS and Azure, to develop scalable and efficient data solutions. Adept in the full data lifecycle, including ETL processes, data modeling, and database management using tools like Apache NiFi, Informatica, ErStudio, and ErWin. Proven ability to implement complex data transformations and automations using Python, SQL, and PL/SQL, and integrate data from multiple sources, including structured and unstructured databases, to drive actionable insights. Demonstrates proficiency in visualization tools such as Tableau and Power BI to deliver impactful, data-driven dashboards and reports, supporting business intelligence and strategic decision-making. Collaborates effectively with cross-functional teams to implement agile methodologies and deliver high-performance data solutions tailored to business needs.

# Skills

Programming Languages: Python, Java, SQL

Database Management Systems: Oracle, MS SQL Server, PostgreSQL, DB2 zOS, Cassandra, Couch, Netezza (PDA/DB2AA)

Big Data Technologies: Hadoop, Apache Spark, Apache Kafka ETL Tools: Apache NiFi, Informatica, SSIS

Cloud Platforms: AWS (Amazon S3, EC2, EMR, RDS, Glue), Azure (Azure Synapse, Azure Databricks) Data Modeling Tools: ErStudio, ErWin

Visualization Tools: Tableau, Power BI

Development Practices: Agile Methodologies, Metadata Management, Database Management, REST/SOAP API Integration

# Education

University of Alabama at Birmingham

Master's degree, Data Science. Aug 2021 - Apr 2023

GITAM Deemed to be University Hyderabad, India

Bachelor of Technology in Electronics & Communication Engineering Jul 2015 - Apr 2019

# Work Experience

State Farm Atlanta, GA

Data Engineer Apr 2023 – Present

* Analyzed complex data requirements and sources to support enterprise data architecture, ensuring comprehensive and accurate insights for decision-making.
* Developed and maintained data models using ErStudio and ErWin, integrating them seamlessly into existing data architectures to support scalable business solutions.
* Enhanced data structures and database designs across various platforms, including DB2 zOS, Cassandra, Oracle, and PostgreSQL, optimizing performance and scalability.
* Utilized big data technologies such as Hadoop and Apache Spark for processing and analyzing large datasets, driving actionable insights and data-driven strategies.
* Implemented ETL processes with Apache NiFi and SSIS, streamlining data extraction, transformation, and loading to improve data workflows and efficiency.
* Designed and managed APIs for data integration and real-time streaming using Apache Kafka, ensuring robust and efficient data exchange between applications.
* Performed database tuning and optimization, leveraging tools like ErStudio and various monitoring tools to achieve high performance and reliability.
* Provided production support for database systems, maintaining high availability and performance while addressing any issues promptly.
* Developed interactive dashboards and reports using Power BI, creating dynamic and visually engaging data visualizations to support business intelligence and strategic decision-making.
* Designed and implemented data models in Power BI, utilizing DAX for advanced calculations and measures to enhance report functionality and user experience.
* Integrated diverse data sources into Power BI reports, including SQL Server, Oracle, and Hadoop HDFS, ensuring accurate and comprehensive data representation.
* Optimized Power BI reports for performance by implementing data compression, aggregations, and efficient query execution, handling large datasets effectively.
* Created custom visualizations and interactive elements such as slicers and drill-throughs in Power BI, improving data exploration and user engagement.
* Developed Power BI datasets and dataflows to streamline data preparation and transformation processes, facilitating efficient data analysis and reporting.
* Collaborated with stakeholders to translate business requirements into Power BI solutions, delivering dashboards and reports that meet organizational needs.
* Implemented row-level security in Power BI to ensure data confidentiality by restricting access based on user roles and compliance requirements.
* Provided training and support on Power BI functionalities, helping users create custom reports, troubleshoot issues, and maximize the platform's capabilities.
* Designed and implemented complex SQL queries and stored procedures for data retrieval and manipulation to support business operations and reporting needs.
* Developed and optimized database schemas and table structures in SQL Server and Oracle, enhancing data integrity, performance, and scalability.
* Created and managed views, indexes, and triggers to improve database performance and automate routine data management tasks.
* Utilized PL/SQL for writing efficient scripts, stored procedures, and functions to automate and streamline data processing and business logic.
* Conducted data profiling and data cleansing using SQL techniques to ensure data quality and accuracy in reports and analytics.
* Implemented ETL processes for data extraction, transformation, and loading using SQL Server Integration Services (SSIS), handling large volumes of data efficiently.
* Performed query optimization and performance tuning to enhance the speed and efficiency of SQL queries and database operations.
* Developed complex joins, subqueries, and window functions to perform advanced data analysis and reporting.
* Managed data migration and integration projects, ensuring smooth transitions between different database systems and platforms.
* Wrote SQL scripts for generating ad-hoc reports and data analysis to meet the immediate needs of business users and stakeholders.
* Provided database support and troubleshooting for production systems, addressing performance issues and ensuring system reliability.
* Collaborated with data architects and business analysts to understand data requirements and develop appropriate SQL solutions to meet business objectives.

University of Alabama at Birmingham Birmingham, Al

Graduate Teaching Assistant - Cloud Security Jan 2022 - Dec 2022

* Assisted in teaching Cloud Security from August 2022 to December 2022 under Dr. Ragib Hassan at the University of Alabama at Birmingham, enhancing students' understanding of cloud security principles.
* Taught AWS courses to undergraduate and graduate students, improving their proficiency in cloud computing and AWS services.
* Graded student assignments and exams and provided feedback to help students improve their skills.
* Worked with students one-on-one to help them troubleshoot problems and understand concepts.
* Utilized AWS services such as Amazon S3, Amazon EC2, AWS Lambda, Amazon RDS, and Amazon CloudWatch to support course activities and enhance students' practical skills in cloud computing and monitoring.
* Explored AWS IAM for managing access to AWS services and resources, AWS CloudFormation for automating resource provisioning, and AWS Kinesis for real-time data streaming exercises.
* Introduced students to AWS VPC (Virtual Private Cloud) for secure networking within the cloud and demonstrated the use of AWS CloudTrail for governance, compliance, and operational auditing.
* Leveraged AWS Security Hub to monitor and manage security and compliance across AWS accounts, providing hands-on experience with cloud security best practices.

JH SOFTECH Hyderabad, India

Data Analyst May 2018 - Dec 2021

* Developed and deployed SSIS packages for ETL processes, including unit testing and deployment across Development and QA environments. Translated business requirements into technical specifications for ETL solutions.
* Optimized advanced T-QL queries for Azure cloud and other applications, improving performance and data processing efficiency.
* Collaborated with software engineers to deliver high-quality deployments, meeting departmental and customer quality and scheduling standards.
* Managed SQL Server jobs and alerts, including backup, recovery, and disaster recovery planning for production databases. Utilized SQL Mail for alert notifications.
* Ensured adherence to ETL standards and architecture, enhancing deployment procedures and organizational compliance.
* Designed and implemented ETL packages using SSIS, integrating data from SQL Server and flat files into target data stores with complex transformations.
* Worked with business teams to optimize onboarding processes for bulk claims data, ensuring accurate and efficient integration.
* Evaluated and leveraged legacy solutions to meet new application requirements.
* Developed SSAS Tabular Models, including shared dimensions, measures, and cubes, for enhanced data analysis.
* Implemented Copy Activity and custom pipelines in Azure Data Factory for data migration, using SQL, SQL Azure, Azure Storage, and PowerShell.
* Designed and reviewed SSAS tabular cubes, utilizing Star Schema, MDX, and DAX queries for advanced data modeling.
* Created and managed Power BI reports, integrating data from SharePoint and other sources, and embedded reports in SharePoint with role-based access control.
* Developed data extracts and reports for SSRS and Power BI to support comprehensive data analysis and visualization.

Bharat Sanchar Nigam Limited Hyderabad, India

Data Analyst April 2016 - March 2017

* Conducted data cleaning and preprocessing on large datasets to ensure data accuracy and consistency.
* Utilized SQL for data extraction and querying, supporting various business intelligence initiatives.
* Performed data manipulation and calculated key metrics using MySQL queries, including window functions and subqueries, alongside MS Excel for comprehensive reporting.
* Collected and aggregated data from internal sources using MySQL, ensuring accurate and clean datasets for reporting and analysis.
* Designed and maintained logical data models using ErWin, ensuring flexibility and scalability in data architecture.
* Created XML/JSON schemas for data exchange and storage, facilitating seamless integration with various applications.
* Collaborated with senior data engineers to develop platform-level data models and database designs.
* Monitored and tuned databases for high performance, using tools to analyze data access patterns and optimize configurations.
* Documented business data requirements and maintained accurate metadata, supporting ongoing data governance efforts.
* Assisted in the development of data-driven insights to optimize telecom services and improve customer satisfaction.
* Collaborated with cross-functional teams to identify key performance indicators and track metrics for continuous improvement.
* Visualized and monitored email performance by creating Tableau dashboards, integrating data from MySQL and MS Excel.
* Utilized the RMySQL package in R to connect to MySQL databases, retrieving data for detailed statistical analysis and data visualization.