Praveen Kumar

Lucknow, India | praveenprajapati0912@gmail.com | +91 9142889256 | linkedin/praveen-kumar

SUMMARY

Experienced Data Engineer with over 3+ years of expertise in leveraging advanced technologies for enhanced collaboration and data analysis. Proficient in SQL, Python, and Azure services such as Azure Data Factory, Azure Databricks, Azure Function App, and PySpark. Skilled in ETL development, report generation, and data security measures. Dedicated to transforming complex datasets into actionable insights that align technical solutions with organizational goals.

SKILLS

- Python
- SQL
- Azure Data Factory
- Pandas
- Azure Databricks
- ETL
- Azure Function Apps
- PySpark
- Azure Logic Apps

WORK EXPERIENCE

Data Engineer, Tranzita Systems Feb 2021 – Present

Project: OLA RCA (Online Availability Root Cause Analysis) **| Client:** P&G, Mumbai **| Role:** Data Engineer

- Engineered and maintained data pipelines to ensure consistent data accessibility. Enhanced project efficiency by architecting an ETL data pipeline using Azure Data Factory to upload input files to the database, optimizing performance and accelerating data manipulation, resulting in a 90% reduction in manual work.
- Managed a comprehensive process re-engineering project to improve and consolidate end-to-end service processes for leading e-commerce and delivery platforms, including Amazon, Nykaa, Swiggy, Zepto, Flipkart, and Blinkit. Achieved a significant 90% reduction in manual work.
- Created and deployed ETL pipelines using Python and SQL integrated into Azure Functions, resulting in an 80% reduction in data retrieval time and an 85% improvement in system performance.
- Implemented data quality checks using Azure Data Factory, SQL, and Python scripts, ensuring accurate and clean data before ingestion into the data lake, improving overall data integrity by 95%.
- Automated performance monitoring of ETL pipelines with Azure Monitor, Log Analytics, and custom Python scripts, leading to a 70% reduction in pipeline downtimes and faster issue resolution.
- Optimized SQL queries and automated data transformation scripts with Python for data retrieval from large datasets, cutting down query response time by 60% and improving overall pipeline efficiency.

Project: Transport Audit Process Automation | **Client:** P&G, Mumbai | **Role:** Data Engineer

- Automated e-commerce order processing audits using Python's pandas library, achieving 95% record iteration and ensuring thorough data validation and accuracy, which significantly reduced error rates in order tracking systems.
- Developed and implemented KNIME workflows to automate quarterly audits and analyze sample data, resulting in a 95% reduction in manual auditing time and saving approximately 120 hours per quarter, improving audit cycle efficiency.
- Collaborated with cross-functional teams to refine audit criteria and optimize workflow logic**, ensuring more accurate data sampling and faster audit turnaround times, improving overall data compliance by 30%.
- Streamlined reporting processes by integrating KNIME workflows with SQL databases, automating data extraction and report generation, which accelerated the delivery of audit results by 80%.

Project: Demand Forecasting | **Client:** P&G, Mumbai | **Role:** Data Engineer

- Utilized advanced data analysis to predict market demand for stocks with over 85% accuracy, improving profit margins and risk management strategies by identifying optimal buying and selling windows.
- Leveraged Python to efficiently manage and analyze large datasets, ensuring precise data handling and

accuracy in calculations. This approach enhanced project outcomes by 30% through more reliable, data-driven decision-making.

- Implemented predictive models using machine learning algorithms in Python, improving forecasting accuracy for stock market trends and enabling better-informed investment strategies, resulting in a 20% increase in portfolio performance.
- Optimized data transformation processes by automating data cleaning and normalization with Python, reducing processing time by 40% and enabling faster insights and improved decision-making capabilities.

Project: OSA (On Shelf Availability) | **Client:** P&G, Mumbai | **Role:** Data Engineer

- Crafted a data-driven approach for sales and inventory management by implementing a combination of weekly data and 3-day predictive analysis techniques, achieving an impressive 95% accuracy in forecasting sales and product availability. This approach led to a 20% reduction in stockouts and a 15% decrease in overstock situations.
- Developed an ETL pipeline in Azure Data Factory to load data into SQL, speeding up data integration by 50%. Created a reusable SQL stored procedure that significantly enhanced Power BI visualizations, leading to more insightful and actionable business intelligence.
- Optimized the ETL pipeline performance by incorporating data partitioning and parallel processing techniques, resulting in a 35% reduction in data processing time and improved efficiency in reporting.
- Integrated real-time data feeds into the sales and inventory management system, providing up-to-date insights and enabling more responsive decision-making. This integration enhanced forecasting accuracy and improved the replenishment process by 25%.
- Designed and implemented advanced data visualizations in Power BI, leveraging the enhanced SQL stored procedures to create interactive dashboards and reports. This development improved stakeholder engagement and facilitated more effective data-driven strategies.

EDUCATION

Dr. APJ Abdul kalam Technical University, B.Tech in ECE — Aug 2016 - Sept 2020