

PAVAN KALYAN BERA

Fullerton, CA | 714-936-5448 | pavankalyanbera2401@gmail.com | [LinkedIn](#)

PROFESSIONAL SUMMARY

I am a passionate and results-driven Data Analyst with strong expertise in data analytics, software development, and cloud computing. Skilled in programming languages like Python, Java, and SQL, I have hands-on experience with big data technologies, cloud platforms, and ETL pipelines. Adept at solving complex problems with clean, efficient solutions while following industry best practices. My focus on continuous learning helps me stay updated with the latest advancements, ensuring innovative and efficient solutions.

Data | Cloud | Big Data | ETL | Analytics Engineer at core, proficient in SQL, Python, and cloud technologies. Experience includes:

- Leading and collaborating with Agile, hybrid distributed, and cross-functional teams.
- Designing, architecting, and deploying multi-tier, distributed computing frameworks.
- Experience in project delivery, report development, and ETL platforms.
- Managing multiple projects and presenting to management and stakeholders.
- Proficiency in usability testing, performance testing, automated testing, and test plan development.
- Core knowledge of data pipelines, API development, and cloud-based data solutions.
- Consuming RESTful web services and processing data for business validations.

PROFESSIONAL HIGHLIGHTS

- Developed and maintained high-quality, scalable data analytics dashboards using PowerBI and Tableau.
- Worked closely with stakeholders to implement data-driven solutions and improve business decision-making processes.
- Applied strong computer science fundamentals (data structures, algorithms) to deliver optimized and efficient ETL pipelines.
- Participated in code reviews, providing constructive feedback to peers and ensuring best practices.
- Contributed to the development of internal data frameworks, improving team efficiency and product delivery.

EDUCATION

California State University, Fullerton, CA

Aug 2022 – May 2024

Master of Science in Computer Science

GPA: 3.65/4

Sathyabama Institute of Technology and Science, Chennai, India

Aug 2016 - July 2020

Bachelor of Technology in Computer Science

GPA: 7.18/10

TECHNICAL SKILLS

- **Programming Languages:** Java, C++, C#, Python, PySpark, C, HTML, CSS, SQL, Scala, Unix, Powershell
 - **Big Data & Cloud Technologies:** MongoDB, MySQL, SQL server, HDFS, Hive, Spark, Snowflake, Apache Airflow, Kafka, Databricks, Oracle, Teradata
 - **Data Visualization & Tools:** AWS, Azure Cloud, Spark Streaming, API Gateway, Flask, Jira, Git, Tableau, Informatica, PLSQL, PowerBI, Azure Analysis Services
-

PROFESSIONAL EXPERIENCE

California State University, Fullerton — *Fullerton, CA*
Data Analyst

May 2024 - Present

- Developed and managed parking and transportation data analytics using Excel and SQL, leading to a 75% improvement in ride-matching efficiency and a 50% increase in event support effectiveness
- Assisted in collecting, organizing, and evaluating large data sets, leveraging Excel VBA and Python to streamline and automate processes.
- Developed interactive dashboards and reports in PowerBI for the parking and transportation department, effectively communicating findings and driving a 35% improvement in decision-making processes
- Processed and analyzed monthly Commuter Tracking Forms for AQMD compliance, using data insights to streamline submissions

Accenture — *Hyderabad, India*
Software Analyst

July 2020 - July 2022

- Developed and optimized ETL pipelines on AWS, enhancing data ingestion, cleansing, and processing efficiency by 40%, and facilitated seamless integration with Amazon Redshift for advanced analytics.
 - Implemented AWS SNS to establish a reliable notification and email alert system, resulting in a 50% increase in timely communication and engagement with clients and users
 - Employed Athena for data crawling and executed vital sanity checks on S3, enhancing data accuracy and integrity, which led to a 35% reduction in data-related issues
 - Integrated Amazon Redshift with Tableau, improving data accessibility and enabling the generation of actionable client reports.
-

PROJECTS

Real-Time Stock Market Data Analysis Using Apache Kafka

Oct 2022 - Dec 2022

- Built and deployed a Kafka-based messaging system on AWS EC2 instances, enabling scalable and distributed data processing
- Utilized Glue Crawler to automate ETL of data, facilitating seamless updates to the Glue Data Catalog for query optimization

Twitter Data Extraction and Analysis Pipeline

Oct 2023 - Dec 2023

- Designed and implemented a scalable ETL pipeline to extract data from Twitter via its API, process this data with Apache Airflow, and store the processed results in AWS S3, deploying the entire solution on an AWS EC2 Ubuntu instance for robust and automated data handling.
- Configured Apache Airflow for sophisticated workflow management, orchestrating data extraction, transformation, and loading tasks to ensure efficient, automated operations.

CERTIFICATIONS

- AWS Certified Solutions Architect - Associate