

# AJAY CHINTALA

[in LinkedIn](#) | [+91 9494433397](#) | [Ajaychintala98@gmail.com](#)

## Profile

Innovative Big Data Engineer with a demonstrated history of orchestrating successful data strategies and driving transformative projects. Proven expertise in building large scale data pipelines for real time data streaming and batch processing. Expertise in Azure, Snowflake, Hadoop, PySpark, Kafka and GCP. Recognized for seamlessly migrating legacy systems, enhancing operation efficiency and contributing to substantial revenue growth.

## Skills

- **Cloud Data Engineering:** Azure Data Factory | Azure Data Flows | Azure Synapse | Azure Analysis Service | PySpark | Snowflake | Pythom
- **Database Technologies:** SQL Server | MSSQL | ADW | Snowflake | Hive
- **Cloud Platforms:** Azure | Snowflake
- **Data Pipelines:** ETL Development | Data Integration | Data Warehousing
- **Programming:** Linux Basics | PowerShell Scripting | SQL | Python
- **Data Catalog:** Alation
- **DevOps:** Git | CI/CD | Airflow
- **Business Intelligence:** Power BI
- **Data Security:** Identity and Access Management (IAM) | SAMBA Application
- **Ticketing Tools:** Service Now | BSM Remedy
- **Project Management:** PMP Trained | ITIL Trained

## Experience

- |  |  |                       |                          |
|--|--|-----------------------|--------------------------|
| <b>Senior Big Data Engineer</b>  | <b><u>SANOFI</u></b>                       | <i>Hyderabad, IND</i> | <b>07/2024 - Current</b> |
| <ul style="list-style-type: none"><li>• Implemented a comprehensive data management strategy aligning with business objectives, ensuring quality, and fostering consistency.</li><li>• Developed PySpark applications for both batch and real-time processing, optimizing performance by tuning Spark jobs and leveraging Spark's in-memory computing.</li><li>• Implemented ETL processes using PySpark, transforming and loading structured and unstructured data into data warehouses.</li><li>• Applied partitioning and bucketing strategies in Hadoop and PySpark to efficiently handle large datasets and optimize query execution.</li><li>• Utilized Hadoop for distributed storage and parallel computation of large datasets, ensuring high performance in data processing tasks.</li><li>• Developed Python scripts to automate data ingestion, transformation, and loading (ETL) processes for large-scale data pipelines.</li><li>• Promoted knowledge sharing and best practices adoption, improving cross-functional communication and project outcomes.</li></ul>                           |  |                       |                          |
| <b>Big Data Engineer</b>   | <b><u>Schlumberger</u></b>                 | <i>Pune, IND</i>      | <b>06/2022 – 07/2024</b> |
| <ul style="list-style-type: none"><li>• Designed and developed real-time data ingestion and processing pipelines using Apache Kafka and PySpark Streaming.</li><li>• Wrote custom Spark SQL queries and optimized them for better performance, enabling data analysts to run queries with faster response times.</li><li>• Implemented framework for Automated Data refresh, monitoring and alerting for Analysis services / Power BI Tabular models using azure automation</li><li>• Design, implement, and optimize ETL pipelines using Apache Spark and Kafka Streams for real-time data ingestion, transformation, and loading into data warehouses.</li><li>• Experience with storage solutions like HDFS, Hive to store and manage large datasets efficiently in Hadoop-based environments.</li><li>• Spearheaded the management and analysis of diverse datasets, including real-world data, maintaining high data quality and compliance.</li><li>• Demonstrated proficiency in manipulating large datasets, assessing vendor data quality, and ensuring secure, compliant data transfers.</li></ul> |  |                       |                          |
| <b>Data Engineer</b>   | <b><u>Capgemini Technology Service</u></b> | <i>Bangalore, IND</i> | <b>04/2018 - 06/2022</b> |
| <ul style="list-style-type: none"><li>• Worked on Legacy BI Reporter Applications</li><li>• Aligned with Agile methodologies, scrum stories and sprints to develop Power BI dashboards.</li><li>• Standardized processes and configured metrics, alarms and monitors using Azure technologies.</li><li>• Conducted data analysis, providing valuable insights for informed business decisions.</li></ul>   |  |                       |                          |

## Education

- |   |                                 |                |                |
|---|---------------------------------|----------------|----------------|
| <b>Bachelor of Engineering Technology</b>   | <b><u>ANDHRA University</u></b> | <i>AP, IND</i> | <b>04/2017</b> |
| <ul style="list-style-type: none"><li>• In Electrical &amp; Electronics Engineering</li></ul> |                                 |                |                |

## Projects

- Worked on Hadoop environment, deployed all the required services like Spark, Kafka, Hive. Created Kafka brokers, consumers and producers to stream data into HDFS.
- Worked with streamsets to ingest data into HDFS from various devices.
- **Managed ETL Migration** from SSIS to Azure Data Flows.
- **AAS Migration** - OBIEE RPD Migration to Azure SSAS Models or Cubes.

## Mentorship

- Served as a Technical Interview Panel Member for Schlumberger and Capgemini Roles.
- Provided organization-wide tutoring, covering Project Technologies, High-Level Architecture, Career advice, and Professional portfolio development