

# Ashutosh Mohanty

✉ ashutosh70161@gmail.com ☎ 6370270613 💬 ashutosh-mohanty

## Profile Summary

Versatile Data Professional with 3 years of experience spanning Data Engineering and Data Science. Proven track record in building robust data pipelines, deploying scalable machine learning models, and performing advanced analytics to drive business impact. Skilled in Python, SQL, Apache Spark, and cloud platforms (AWS, Azure). Adept at bridging the gap between raw data and actionable insights.

## Experience

### Data Engineer

FactSet

Hyderabad, IN

August 2023 – Present

- Designed and implemented data pipelines using **PySpark** and **Databricks** to process and standardize large-scale financial datasets from multiple sources, improving data processing efficiency by 35%.
- Designed and deployed a scalable **ETL** pipeline that ingests daily stock market data from vendors, transforms it using PySpark, and loads it into AWS **Redshift** for analytics.
- Built and deployed ML models (churn prediction, LTV scoring) using Scikit-learn, integrated into data pipelines.
- Built an end-to-end ETL pipeline using **AWS Glue**, S3, and Redshift to process 10M+ daily records. Automated orchestration with Airflow and containerized the workflow using Docker.
- Automated data validation and alerts using **Airflow**, improving pipeline reliability by 40%.
- Built automated data quality checks for 50+ financial data sources, reducing manual QA effort by 60%.
- Collaborated with product and data teams to define data contracts and improve overall data quality.
- Developed **Python** based analytical workflows for **A/B testing**, user segmentation, and churn driver analysis enabling product teams to make data-backed roadmap decisions.
- Built interactive **Power BI** dashboards to visualize key KPIs such as feature adoption, user cohorts, and campaign ROI; reduced manual reporting time by 40%.

### Data Analyst

FactSet

Hyderabad, IN

October 2022 – Aug 2023

- Automated reporting workflows with **Python** and **SQL**, reducing turnaround time by 60%.
- Analyzed large datasets to identify fraud patterns using clustering techniques (**KMeans**, DBSCAN).
- Built a time-series forecasting model for product demand using Prophet; improved forecast accuracy by 20%.
- Engineered data pipelines using Azure Data Factory and Snowflake for daily reporting.
- Applied **statistical modeling** techniques (hypothesis testing, regression, correlation) to identify key behavioral drivers influencing user retention.

## Certifications

**CKA:** Certified Kubernetes Administrator

**Microsoft:** Certified Azure Fundamentals AZ-900

**AWS:** Certified Cloud Practitioner

## Technologies

**Programming:** Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn)

**Databases & Data Warehousing:** MySQL, PostgreSQL, SQL Server, MongoDB, Databricks, Snowflake, Amazon Redshift

**Machine Learning:** Linear/Logistic Regression, Gradient Descent, Clustering, Decision Trees, Random Forest

**Data Engineering:** ETL pipelines, Apache Spark, AWS (S3, Redshift, Glue), Databricks, Azure Data Factory

**Visualization & BI:** Power BI, Tableau, Matplotlib, Seaborn

**DevOps:** Docker, Kubernetes, Ansible, Terraform, GitHub Actions, GitLab CI, Jenkins, CloudWatch, Airflow

## Education

---

**Odisha University of Technology and Research**  
*B.Tech in Electronics and Instrumentation Engineering*

2018 – 2022