

ANJALI SALUJA

anjalisaluja305@gmail.com | +91 7470848106 | [LinkedIn](#) | [GitHub](#)

Data Analyst with 2+ years of experience in building scalable data pipelines and utilizing AI-driven insights to enhance business decision-making. Skilled in transforming complex data into actionable strategies to drive efficiency and growth.

SKILLS

Programming & Databases: Python (Pandas, Numpy, Matplotlib, Seaborn, Transformers, PyTorch,), SQL, R, MS SQL, PostgreSQL, RDBMS

Gen AI & LLM Tools: Databricks DBRX, BGE-Large-EN, OpenAI, Vector Databases (FAISS), Prompt Engineering

Data Engineering & Analytics: Databricks, Apache Spark, Power BI, Tableau, Apache Superset, Looker, Metabase, Excel

Automation & Scripting: Python (BeautifulSoup, Selenium, Scrapy)

Statistical & Predictive Analysis: SPSS, Scikit-learn, Time Series Forecasting

Marketing & Business Analytics: Marketo, Google Analytics, A/B Testing, Customer Segmentation

WORK EXPERIENCE

Data Analyst | LatentView Analytics

February 2024 – Present

Data Lineage Dashboard

Deployed a centralized data lineage dashboard using Databricks, PySpark, Power BI, and SQL, improving data traceability by 30% across Adobe applications.

- Extracted and structured table- and application-level lineage in Databricks, capturing complete upstream and downstream flows.
- Integrated Databricks with Power BI to visualize lineage across 100% of applications and tables.
- Conducted end-to-end validation with 15+ application owners, ensuring accuracy, adoption, and usability of the dashboard.

AI PenPal – Intelligent Email Management (Internal Project)

Built and deployed an AI-driven email outreach system using Databricks, Python, PySpark, LLMs, boosting conversion rates by 10–15% and sales closures by 2x.

- Utilized Python to scrape and process structured/unstructured data from 10,000+ records across Salesforce, Salesloft, and LinkedIn, enriching 95% of prospect profiles.
- Performed API integrations and batch processing with Python to enable real-time data synchronization and scalable workflow automation.
- Developed end-to-end lead data pipelines and integrated Databricks DBRX-Instruct and BGE-Large-EN, enabling generation of 1000+ hyper-personalized emails/week.
- Conducted vector search and model-serving endpoints, reducing manual input for content generation by 70%.

Data Analyst Intern | KocharTech

June 2022 – May 2023

- Automated internal site data extraction and Excel reporting using Python & Selenium, reducing manual effort by 10+ hours/week.
- Designed interactive dashboards in Apache Superset, Metabase and Plotly Dash, accelerating stakeholder insights by 30%.
- Integrated Cube.js with databases to boost query performance and cut data retrieval time by 40%.
- Created an executive-level, sales-focused Power BI dashboard for Samsung, enhancing sales visibility and enabling strategic decisions across 5+ business units, resulting in a 15% improvement in reporting efficiency.

CERTIFICATION

Databricks Certified Generative AI Associate

Databricks Certified Data Engineer Associate

AWARDS & ACHIEVEMENTS

Runner-Up | Quid Hackathon – LatentView Analytics

October 2024

- Developed "DarkIntel", an innovative solution, as part of a team-based competition.
- Secured Runner-Up position among 42 teams in a company-wide hackathon.

Runner-Up | TNT (Data Science Hackathon) – LatentView Analytics

February 2023

- Selected for an interview opportunity at LatentView Analytics through a competitive hackathon with 5,000+ participants.

EDUCATION

Lovely Professional University

Bachelor of Technology (BTech), Computer Science

July 2019 – June 2023