Swapnil Deore

315-883-7097 | deoreswapnil@outlook.com | linkedin.com/in/swapnildeore | Github

SUMMARY
Data Engineering and Analytics professional with 3+ years of experience in building scalable data pipelines, ETL processes, data modeling, tracking KPIs and designing data-driven solutions. Proficient in Python, SQL, Azure Databricks, PySpark, Tableau, AWS, Snowflake, Power BI and Dbt with strong focus on optimizing data workflows and enabling business intelligence solutions.

EXPERIENCE

Data Analyst, Syracuse University   September 2024 - Present
- Engineered ETL pipelines in Azure Databricks using PySpark to process 10 million alumni records from Azure Data Lake into Snowflake, reducing data processing speed by 60% and enabling real-time insights for career networking opportunities.
- Created Dbt models to transform raw alumni data into scalable OLAP star schema, making it easier to build materialized views/Tables for Tableau dashboards that support alumni engagement and networking initiatives for university students.
- Developed Tableau KPI dashboards integrated with campus store's data warehouse to analyze sales trends and restocking cycles, driving data-driven inventory strategies that resulted in $50K+ annual savings by minimizing excess stock.
- Automated ad-hoc reporting with optimized SQL scripts, saving 32 hours/week, streamlining reporting processes, enabling more efficient data insights, enhancing decision-making and achieving significant reduction in manual workload across teams.
- Conducted A/B testing on product pricing, analyzing sales data to identify the most effective approaches, resulting in 10% increase in sales and more efficient inventory management at Syracuse University campus store.

Data Science Intern, Prompt Marketing Agency   June 2024 - August 2024
- Led conversational analysis of follower interactions on Lidl CEO's LinkedIn profile using SQL and Langchain (ML) models; uncovered top-performing content themes and trending topics, leading to 80,000+ increase in profile engagement.
- Conducted data analysis efforts using Salesforce CRM systems to explore market trends, customer behavior and audience segmentation, contributing insights that helped my client Cetaphil increase annual consulting bookings by approximately 4%.
- Built KPIs in Power BI using DAX and calculated columns to centralize metrics across email, social and paid campaigns reducing manual reporting time by 70% and enabling managers to reallocate budget mid quarter driving 12% increase in ROI.
- Analyzed over 70K+ marketing interaction records using Python libraries (NumPy, Pandas, Matplotlib), performing exploratory data analysis (EDA) surfaced high-performing segments and enabled 15% increase in ad engagement rates.
- Collaborated with cross-functional teams utilizing Jira and Confluence to gather requirements and create Business and Functional Requirement Documents (BRDs/FRDs), driving delivery of data-driven solutions in agile sprints for marketing.

System Engineer, Tata Consultancy Services   September 2020 - July 2022
- Performed User Acceptance Testing (UAT) for Financial Information Exchange (FIX) API data pipelines, configuring message sessions, troubleshooting latency and network issues for clearing corporations and broker firms at the National Stock Exchange.
- Managed 100 million daily trading transactions, ensuring data integrity and security across FIX applications on servers.
- Created and maintained comprehensive data lineage and metadata documentation to support regulatory audits, ensuring full transparency in data transformations and meeting compliance requirements for accurate financial reporting and audits.
- Devised robust checksum-based data validation mechanism like MD5/SHA-256 in ETL pipelines for broker and clearing firm trades, ensuring end to-end data integrity and 100% accuracy, leading to reduced manual reconciliation efforts by 70%.
- Developed data governance policies at NSE to ensure alignment with data protection regulations, while utilizing VBA, Lookup functions and Pivot Tables to automate reporting processes and enhance accuracy and reliability of regulatory data outputs.

EDUCATION

Master of Science in Information Systems	CGPA: 3.8/4.00	**May 2024**
Syracuse University, Syracuse, New York, USA		

Bachelor of Engineering in Computer Science	CGPA: 3.6/4.00	**June 2020**
University of Mumbai, Mumbai, Maharashtra, India		

PROJECTS

Airbnb ELT Pipeline | Dbt, Dagster, SQL, AWS S3   March 2024 - April 2024
- Transformed 250K+ Airbnb booking records using Dbt models, enabling insights on booking trends and host profitability.

Analyzing Engineering Subreddits | Airflow, AWS S3, Redshift, Tableau   June 2023 - July 2023
- Built real-time data pipeline to process engineering subreddit conversations, revealing 42% growth in AI/ML related topics.

SKILLS
- **Programming and Analytics:** Python, SQL, R, C++, HTML/CSS, NumPy, Pandas, Matplotlib, Seaborn, PySpark, NLP, PyTorch, TensorFlow, Jupyter, PyCharm, Visual Studio, Data Wrangling, Data Mining, Visualization, Statistical Modeling, EDA, A/B Testing
- **Data Engineering/Big Data:** Airflow, Kafka, Hadoop, Spark, Informatica, Dbt, Kubernetes, EMR, AWS Glue, Azure Data Factory
- **Business Intelligence Tools:** Tableau, Power BI, Looker, Qlik, Alteryx, GIS, Gitlab, Docker, SDLC, Agile/Scrum, Waterfall
- **Database and Cloud Platform:** Databricks, Snowflake, AWS Redshift, PostgreSQL, MySQL, Google BigQuery, CRM, SAP, AWS (S3, IAM, Glue), GCP, SPSS, Azure, ERP systems, Salesforce CRM, MongoDB, SAS, SSRS, SSIS, Fivetran, Matillion