SWAPNIL SHINDE
Product Data Analyst

shinde.swapnil9454@gmail.com | 9423459354 | LinkedIn

OBJECTIVE

Product Data Analyst with 5+ Years of experience leading data preparation, QA automation, and delivery of enterprise-scale datasets. Adept at translating product needs into reliable data outputs and using SQL., Python, and AWS to improve product performance and customer value.

EXPERIENCE

DEMANDBASE INDIA PVT. LTD., *DATA ANALYST*   07/2023 – Present
Core Responsibilities & Skills
- Led product analytics initiatives across multiple data ecosystems, uncovering behavioral insights that directly influenced roadmap priorities, feature optimization, and customer experience improvements.
- Built and automated scalable data pipelines and analytical workflows in AWS, Python, and SQL, improving data reliability, reducing manual effort, and enabling faster experimentation.
- Designed end-to-end KPI frameworks, dashboards, and product performance models to monitor user journeys, funnel efficiency, and product health at scale.
- Partnered with Product, Engineering, and GTM/ABM teams to convert ambiguous business problems into structured analyses, actionable hypotheses, and data-driven recommendations.
- Improved data quality and analytical rigor by integrating new sources, enhancing data models, and establishing robust validation and QA processes.
- Presented insights through clear, executive-ready storytelling, influencing strategic decisions and accelerating adoption of data-driven product practices.

LOGRAN TECHNOLOGIES PVT. LTD., *DATA ANALYST*   07/2020 – 02/2023
Core Responsibilities & Skills
- Data Visualization & Reporting: Developed and deployed 40+ interactive dashboards and reports using Power BI, connecting to diverse data sources (SQL Server, AWS S3, Redshift, Snowflake). Expertise in creating a variety of visualizations (e.g., Line, Stacked, Pie, Matrix, Tree Map, Maps) with advanced features like Drill Down/Through and dynamic parameters.
- Data Transformation (ETL/ELT): Performed end-to-end Data Cleaning and Transformation using M Language (Power Query Editor) and DAX functions to ensure data quality and integrity (e.g., rectifying duplicates and obsolete records).
- Database Querying & Analysis: Wrote complex SQL queries (DDL, DML) utilizing Window Functions, Aggregate Functions, CASE Statements, and Joints in MS-SQL Server and Oracle to extract and prepare data for in-depth analysis.
- Requirements & Deployment: Gathered and translated end-user/client requirements into impactful reporting solutions. Managed report scheduling and subscriptions in Report Manager and published final reports/dashboards to Power BI Service.

EDUCATION

MBA, *SYMBIOSIS SCHOOL FOR ONLINE AND DIGITAL LEARNING*   Pursuing
B-TECH, *MAHARASHTRA INSTITUTE OF TECHNOLOGY, AURANGABAD*   05/2020 – 06/2020

SKILLS & ABILITIES

SQL: Wrote complex SQL queries using joins, CTEs, and window functions to extract actionable insights from large datasets.

PySpark: Built and optimized PySpark pipelines to process, transform, and analyze large-scale datasets efficiently.

PYTHON: Developed Python scripts to automate repetitive data processing and validation tasks. Implemented Python-based data quality checks to identify duplicates, nulls, and inconsistencies.

ADVANCE EXCEL: Applied advanced Excel functions (VLOOKUP, INDEX-MATCH, IF, SUMIFS, COUNTIFS) for complex data analysis. Analyzed large datasets using Pivot Tables to uncover trends and insights.

TABLEAU: Built visually compelling charts and dashboards to communicate insights effectively. Partnered with business teams to deliver Tableau insights that support decision-making.

POWER BI: *Built efficient Power BI data models using star schema and optimized relationships. Developed advanced DAX measures for time intelligence, rolling metrics, and KPI tracking. Solved complex business logic using CALCULATE, FILTER, and iterator functions in DAX.*

AWS S3: *Designed efficient S3 folder structures to support scalable analytics and reporting workflows. Integrated AWS S3 with analytics tools for seamless data ingestion and processing.*

AWS ATHENA: Queried large-scale datasets directly on AWS S3 using Amazon Athena for ad-hoc and exploratory analysis.

AWS GLUE: Integrated AWS Glue with S3, Athena, and Redshift for end-to-end analytics workflows.

SNOWFLAKE: Performed high-performance analytics on Snowflake using optimized SQL queries.

SELLER DATA: *Validated and maintained seller data accuracy for reliable analytics. Performed data cleansing and reconciliation on seller datasets.*

AWS LAMBDA: *Developed AWS Lambda functions to automate data processing and trigger analytics workflows. Used AWS Lambda for event-driven execution of data ingestion and transformation tasks. Integrated AWS Lambda with S3, Glue, and Redshift to enable serverless data pipelines.*

AWS REDSHIFT: *Executed complex analytical queries on AWS Redshift to support business reporting and insights. Designed and optimized Redshift tables using appropriate distribution and sort keys for performance. Handled large-scale datasets in Redshift for fast, reliable analytics workloads.*

BigQuery: *Performed large-scale analytical querying in BigQuery using optimized SQL. Analyzed high-volume datasets in BigQuery to generate actionable business insights.*

JIRA (Agile Methodologies): *Used JIRA to manage Agile workflows, including sprint planning, backlog grooming, and task tracking. Collaborated with cross-functional teams in an Agile environment using JIRA to deliver on sprint goals.*

MCP (Model Context Protocol): Owned end-to-end testing and deployment of new tools on the MCP server, collaborating with stakeholders to enable reliable production releases.

PROJECTS

ADVERTISING AND AD CAMPAIGNS ANALYSIS
06/2022 – 02/2023
- Analyzed affiliate-based and ad campaign performance, calculating key metrics such as Cost Per Click (CPC), Clicks, Views, Cost, and Revenue across multiple regions (USA, UK, Australia, Denmark).
- Produced performance reports to support data-driven decision-making for Marketing, Sales, Legal, and Compliance teams.

CUSTOMER INSIGHTS VIA CLICKSTREAM ANALYSIS
11/2021 – 05/2022
- Retrieved clickstream data from AWS S3 via AWS Lambda and analyzed customer behavior metrics (e.g., rewards redeemed, daily reach, demographics) to generate customer insights reports.

REAL STATE ASSET MANAGEMENT ANALYSIS
01/2021 – 10/2021
- Utilized Snowflake to retrieve and analyze real estate asset data from AWS S3, calculating key financial estimates (equipment total estimates, long-term cost estimates, capital/maintenance budgets).

ANALYSIS & ENRICHMENT OF CLIENT CRM DATA
- Performed comprehensive analysis of client CRM datasets to identify data gaps, inconsistencies, and enrichment opportunities.
- Cleaned, standardized, and validated CRM records to improve data accuracy, completeness, and usability for downstream analytics.
- Enriched customer and company profiles by integrating external datasets, firmographic attributes, and missing contact details.
- Built automated data quality checks and QA workflows to detect duplicates, invalid entries, and broken linkages across tables.
- Developed enrichment pipelines to enhance CRM attributes, including job titles, company identifiers, segmentation labels, and persona tags.
- Generated data-driven insights highlighting Data Points based on Family Tree, Type-Status-Structure, Geographical Location, Title-Job Level-Job Function, market segments, and opportunities for targeted outreach.
- Created final consolidated datasets and dashboards to support decision-makers in campaign planning, lead scoring, and customer profiling.
- Collaborated with cross-functional teams (sales, marketing, operations) to align enrichment rules, attribute definitions, and delivery standards.
- Improved CRM reliability and actionability, contributing to better segmentation, increased campaign precision, and enhanced customer understanding.

CONTACT & COMPANY LIST BUILD
- Built high-quality Contact and Company lists by aggregating, cleaning, and validating data from multiple internal sources.
- Applied data enrichment techniques to append missing firmographic and demographic attributes such as job titles, job levels, industries, revenue, employee size, and company hierarchy.
- Designed and executed QA workflows to identify and remove duplicates, invalid titles, defunct companies, and low- quality records.
- Implemented rule-based and persona-based matching to ensure accurate alignment of contacts to target audience criteria.
- Performed segmentation and categorization to create tailored lists for specific market tiers, verticals, and campaign requirements.
- Automated parts of the list-build process to reduce manual effort and improve turnaround time for large datasets.
- Collaborated with sales, marketing, and customer teams to refine targeting criteria and ensure list alignment with campaign objectives.
- Delivered final, ready-to-use contact and company datasets with standardized fields, high accuracy, and strong coverage across defined criteria.
- Improved downstream campaign performance by ensuring high-quality, precisely filtered, and enriched data outputs.

PREPARING COMPANY & CONTACT FLAT FILES ON DEFINED CRITERIA
- Led end-to-end preparation of high-volume company (100M+) and contact (25M+) datasets by executing optimized SQL extractions on AWS EC2 and orchestrating automated batch workflows via shell scripts.
- Drove data drift analysis by comparing current and historical raw files, identifying schema changes, quality gaps, and anomalies impacting downstream product and analytics reliability.
- Built and maintained scalable Python-based QA and transformation pipelines to validate data integrity, standardize fields, and remove invalid or non-compliant records.
- Ensured consistent monthly dataset releases by implementing automated validation checks and historical comparisons aligned to business rules and product requirements.
- Delivered final production-ready datasets in .txt.gz and Parquet formats with strict adherence to schema definitions, delivery criteria, and client expectations.
- Collaborated cross-functionally with engineering, analytics, and product teams to refine data definitions, improve pipeline efficiency, and enhance customer-facing data quality.
- Championed data quality and operational excellence by establishing repeatable workflows, improving pipeline reliability, and reducing manual QA overhead.

DATA ANALYST AT LOGRAN TECHNOLOGIES PVT.LTD.

07/2020 – 02/2023
- Technical Focus: AWS Services (S3, Glue, Redshift, Athena, Lambda), Power-BI Visualization, DAX, Power Query, SQL (Joints, Functions, PL-SQL), and Snowflake Architecture/Data Loading.

TECHNOLOGIES

DATA ANALYST AT DEMANDBASE INDIA PVT. LTD.
Advance Excel, PySpark, SQL, Python, Tableau & AWS

CONTACT & COMPANY LIST BUILD
Advance Excel, PySpark, SQL, Python, Tableau & AWS

PREPARING COMPANY & CONTACT FLAT FILES ON DEFINED CRITERIA
SQL, PySpark, Python, Shell Script & AWS EC2

ADVERTISING AND AD CAMPAIGNS ANALYSIS
- POWER BI, AWS S3, SQL SERVER, AWS REDSHIFT, AWS GLUE, JIRA.

CUSTOMER INSIGHTS VIA CLICKSTREAM ANALYSIS
- SQL SERVER, AWS S3, REDSHIFT, GLUE, POWER-BI.

REAL STATE ASSET MANAGEMENT ANALYSIS
- SQL, POWER BI, S3, Snowflake, AWS.