

SMITH PAREKH

Business Intelligence | Analytics Engineering

Mumbai, India | +91 86899 60348 | smithparekh107@gmail.com | [linkedin.com/in/smith-parekh](https://www.linkedin.com/in/smith-parekh) | github.com/smithparekh

Professional Summary

Data Analyst with 4+ years turning raw, multi-source data into revenue and growth decisions for performance marketing and fintech businesses. Works across the full stack: SQL/Python analysis and executive BI dashboards (Power BI, Metabase); data modeling and warehousing (Snowflake, dbt, ClickHouse); and applied AI systems (RAG pipelines, autonomous agents, fraud-detection ML).

Skills & Technologies

- **Languages & Query:** Python (Pandas, NumPy), SQL, DAX, Power Query (M)
- **Data Warehousing & Modeling:** Snowflake, ClickHouse, PostgreSQL, MySQL, SQL Server, dbt, Dimensional Modeling (Star Schema), ETL Pipelines, Incremental Processing
- **BI & Visualization:** Power BI, Metabase, Tableau, Excel (Pivot Tables, VLOOKUP), Dashboard Development, Data Storytelling
- **Analytics Methods:** Cohort Analysis, Funnel & Attribution Analysis, A/B Testing, Retention & Churn Analysis, RFM Segmentation, Trend & Statistical Analysis
- **AI, ML & Automation:** LangChain, LangGraph, OpenAI API, Groq API, RAG, Qdrant, FAISS, AWS Fraud Detector, Autonomous Agents
- **Cloud & Tools:** AWS (Lambda, S3), MCP (Model Context Protocol), REST APIs, Git, Jira

Work Experience

Data Analyst — ParityCube

Dec 2024 – Present

Performance Marketing / Affiliate Platform

- Own 3 executive BI dashboards (Metabase/Power BI on PostgreSQL & ClickHouse) across Cuelinks, Zingoy, and DeSiDime, tracking 6.8M+ affiliate transactions; cohort, funnel, and trend analysis on CVR, EPC, and commission margins drove a 10–14% platform profit-margin improvement over 3 fiscal years at 83–93% monthly data-validation accuracy.
- Built a cross-platform campaign comparison tool in ClickHouse for flexible period-over-period analysis of GMV, commission, transactions, and conversion rate — using templated dual date-window filters and conditional aggregation to compute both periods in a single query pass.
- Designed a multi-product executive dashboard unifying gift cards, prepaid cards, cashback, and affiliate revenue across 8+ source tables on a fiscal-year basis, tracking YoY commission trends, gift-card lifecycle profitability, and prepaid-card activation rate; built companion self-serve tools (dual-date-window CTEs, safe percentage-change logic) surfacing top gaining/declining campaigns and publishers at daily-to-yearly granularity, plus fiscal-year-aware Top-20 merchant/publisher rankings.
- Deployed an AWS Fraud Detector model for gift-card fraud with custom HLR mobile-number enrichment features, cutting fraudulent transactions by 80% and manual review by 90%, eliminating police escalations.
- Built a self-learning RAG deal-category classifier (Qdrant vector DB, OpenAI embeddings, cosine similarity) processing ~1,000 deals/day, with an AWS Lambda LLM fallback that writes back to the vector store — fully eliminating manual categorization; also built an MCP server letting non-technical teams self-serve campaign reports via plain-English prompts.
- Automated data collection end-to-end: a Python + Cloudflare Worker pipeline auto-extracts listings for 600–700 of ~1,000 daily deals, and a competitor-monitoring scraper across 6 platforms surfaces 10–15 exclusive stores/coupons per cycle.

Data Analyst — Fin Rise Softech

Jun 2022 – Dec 2024

Fintech

- Built and maintained a Power BI risk-monitoring dashboard tracking market data across 200+ client portfolios, improving data accuracy by 25% and reducing risk exposure by 30%.
- Contributed to database schema and data-model design for a Risk Management System, improving system performance by 15% and strengthening data integrity across all tracking workflows.

- Wrote and optimized SQL queries and stored procedures (SQL Server), cutting query execution time by 10% and eliminating production deadlocks; automated Excel reporting pipelines, increasing analysis throughput by 30%.
- Applied regression and pattern-recognition analysis on trade data to surface trends, enabling strategic adjustments to client trading models.

Independent Projects

- **Affiliate Revenue Intelligence Platform** — Analytics engineering platform on Snowflake + dbt modeling affiliate click, conversion, campaign, and merchant data into a tested dimensional layer (18 models, 150 dbt tests, 0 failures) with incremental processing and full lineage documentation; marts answer affiliate, merchant, campaign, and product performance questions while explicitly preventing aggregation fanout. ([View on GitHub](#))
- **AI Job Application Copilot** — Streamlit application using the Groq LLaMA-3.3-70B API to score a resume against a job description (match %, missing skills, keyword gaps) and generate tailored resume/cover-letter drafts; includes a background job scanner covering 9 job boards (Greenhouse, Lever, Ashby, SmartRecruiters, and others) with smart location/keyword matching, automatic 6-hourly scans, and a local application tracker with conversion analytics. ([View on GitHub](#))

Certifications

- Oracle Cloud Infrastructure 2024 — Generative AI Certified Professional, *Oracle* (2024)
- Google Data Analytics Professional Certificate, *Coursera* (2022)
- Advanced SQL, SQL (Intermediate), SQL (Basic), *HackerRank*
- SQL for Data Science, *Coursera* (2021)
- Deep Learning, *Udemy* (2021)
- IBM Data Analysis with Python, *Cognitive Class* (2020)

Education

B.E. in Information Technology, Universal College of Engineering

2016 – 2022