Hem Gandhi

(Data Scientist | ML Applications & Analytics)

Melbourne, VIC • +61 484 653 827 • **Email** • **LinkedIn** • **Website**

Work Authorizations: Full Australian work rights (Subclass 485)

PROFESSIONAL SUMMARY

Data Analyst / Data Scientist with 1.3 years of commercial analytics experience at NielsenIQ and a Master's in Data Science. Strong in Python, SQL and Excel, building repeatable reporting workflows, data checks, and dashboards for FMCG clients including L'Oréal, Budweiser and Dr. Oetker. Delivered weekly reporting with >98% on-time delivery and >95% accuracy, reducing manual fixes by ~30% and cutting analysis turnaround by ~25%. Built portfolio projects in Streamlit/FastAPI and PostGIS to demonstrate end-to-end data app delivery. Melbourne-based with full Australian work rights (Subclass 485).

SKILLS

Languages: Python, SQL (PostgreSQL), JavaScript/TypeScript

Machine Learning & Analytics: pandas, NumPy, scikit-learn, XGBoost, Feature Engineering, Model Evaluation, Variance Analysis (Driver Logic), A/B Testing Concepts

NLP & GenAI Engineering: LLM Orchestration (LiteLLM), RAG Architectures, Vector Embeddings (Sentence-Transformers), Prompt Engineering, Semantic Search (TF-IDF/Cosine Similarity)

Data Engineering & Cloud: FastAPI (Restful Endpoints), Docker, AWS (S3, EC2, Lambda, IAM), Azure, OCI, Streamlit, ETL Pipelines, Data Validation (Pydantic), Git/GitHub Actions

Geospatial & Visualisation: PostGIS (Spatial Indexing/Joins), Leaflet.js, Tableau, Excel (VBA/Macros), Matplotlib, Seaborn, Plotly

EXPERIENCE

NielsenIQ — Data Analyst (Retail Analytics) (FMCG/Retail) |Vadodara; Gujarat; India Sep 2021 — Dec 2022

Clients: L'Oréal, Budweiser, Dr. Oetker, Shell, etc
Tools: SQL, Excel (advanced), internal BI/reporting tools, basic Tableau

- Automated Reporting: Engineered weekly FMCG client reporting workflows across enterprise portfolios (L'Oréal, Budweiser), lifting on-time delivery to >98% with >95% accuracy.

- ETL Pipeline Optimization: Built SQL data extraction, cleaning, and validation workflows from internal databases, reducing manual data fixes by ~30% and significantly shortening the weekly reporting cycle.

- Variance Analysis: Developed repeatable EDA templates to diagnose sales movement drivers (price/promo/mix), cutting analysis turnaround time by ~25%.

- Dashboard Development: Built and maintained interactive Excel reporting packs (advanced formulas, pivots, VBA automation) to support monthly business reviews and client decision-making.

- Stakeholder Communication: Translated complex dataset anomalies and market trends into clear commercial storylines, directly supporting client retention strategies.

The Maharaja Sayajirao University of Baroda; Internal Quality Assurance Cell (IQAC) — Junior Analyst (Internship) | Vadodara; Gujarat; India   **Apr 2019 — Jul 2019**

- Analyzed feedback from ~36,000 students to build faculty-wise summaries supporting NAAC accreditation and annual planning.

SELECTED PROJECTS

[ATS Resume Analyzer — JD Fit Copilot Full-Stack ML Application](#)   **Nov 2025**

Tech: Python (FastAPI, Streamlit), LiteLLM, Docker, scikit-learn (TF-IDF, cosine similarity), Vector Search, uv

- End-to-End Application: Built a resume analysis web app that ingests PDF/DOCX resumes and Job Descriptions to produce fit scores, keyword coverage, and actionable rewrite guidance.

- Microservices Architecture: Designed a modular architecture where the Streamlit frontend communicates with the analysis engine via strict Pydantic schemas, ensuring type safety and robust data validation across the application.

- LLM Orchestration: Integrated LiteLLM to route prompt requests to various providers (OpenAI, Anthropic) with deterministic fallbacks, ensuring system reliability even during API outages.

- Scoring Engine: Implemented a dual-layer scoring system combining TF-IDF for keyword coverage and Sentence-Transformers for semantic vector similarity to rank candidate "evidence" against requirements.

Polimap — High-Performance Geospatial Visualization Web App   **Aug 2024 — Nov 2024**

Tech: Python (pandas, requests, FastAPI), SQL, PostgreSQL, **PostGIS**, React (TypeScript/JavaScript), **Leaflet** (markers, clustering, heatmaps), Highcharts, Power BI/DAX (prototype), Docker, MongoDB (metadata), GraphQL (select queries), Git/GitHub

- Geo-Analytics Platform: Built an interactive web application to ingest, validate, and visualize Australian electoral and census datasets for region-level decision support.

- Spatial Engineering: Designed a PostGIS backend optimized with spatial indices (GIST) to handle rapid querying of 100,000+ geospatial data points and complex spatial joins.

- API Performance: Developed read-optimized FastAPI endpoints achieving <50ms latency for viewport-based spatial queries, utilizing lazy loading and clustering for performance.

- Frontend Visualization: Created a dynamic Leaflet.js interface with heatmap rendering and cluster layers, enabling real-time demographic data exploration

CERTIFICATIONS

[BCG X — PowerCo Churn Analysis (Forage)](#)   **Sept 2025**

- Designed a churn prediction model using Random Forest/XGBoost; handled class imbalance with SMOTE and optimized thresholds for business recall targets.

Citi FP&A — Virtual Experience (Forage)   **Sept 2025**

- Built executive reporting packs for variance analysis and scenario-based forecasting under shock conditions.

EDUCATION

Performance Education — Professional Year in IT | Melbourne; VIC | Jun 2025 — Present
• Industry readiness: workplace communication; job search strategy; and Australian professional standards.

Swinburne University of Technology — Master of Data Science | Melbourne; VIC | Feb 2023 — Dec 2024
• Key subjects: Machine Learning; Statistical Modelling; Data Management; Cloud/APIs.
• Achievements: High Distinction for geo-analytics capstone (NDA).

The Maharaja Sayajirao University of Baroda — B.Sc. in Statistics | Vadodara; India | 2018 — 2021
• Core areas: Probability; Inference; Regression; Time Series; Survey Methods.