ANURAG MESHRAM

Data Analyst

+91 9996323959 | anuragmesh@gmail.com | Nagpur, Maharashtra | LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

Data Analyst with 1+ year of experience (CenterPoint Group, USA – Remote) transforming complex datasets into actionable business insights using SQL, Python, Power BI, Tableau, and Qlik Sense. Proven track record of delivering $650K+ in identified cost savings, building 12+ executive dashboards, and pioneering GenAI-powered analytics (Google Gemini) for automated item matching and spend intelligence. Expert in data cleaning, KPI design, ETL automation, web scraping, and supplier performance analysis. MBA in Business Analytics with strong stakeholder management and business acumen.

TECHNICAL SKILLS

Analytics & BI:	Power BI (DAX, Power Query), Tableau, Qlik Sense, Excel (Advanced: PivotTables, VBA, Power Query)
Languages:	SQL (Advanced – MySQL, PostgreSQL, Azure SQL), Python (Pandas, NumPy, Matplotlib, Seaborn)
Databases & Tools:	MongoDB, MySQL, PostgreSQL, Azure SQL, Git, Azure DevOps, Jupyter Notebook, VS Code
Cloud & Data Eng.:	Azure (Data Lake, SQL, Data Factory), ETL Pipelines, Web Scraping (Selenium, BeautifulSoup)
AI & Analytics:	Google Gemini API, RAG Systems, Semantic Search, EDA, Statistical Modeling, A/B Testing, Forecasting

PROFESSIONAL EXPERIENCE

Data Analyst | CenterPoint Group – USA (Remote)   *Aug 2024 – Jan 2026 | Nagpur,India*

- Analyzed 1.5M+ supplier records using SQL and Python, identifying **$450K** in annual cost savings through competitive pricing analysis across Office Depot, Staples, WB Mason, and Grainger.
- Designed 12+ Power BI dashboards analyzing spend, supplier performance, and market trends for C-suite executives, maintaining 98% data accuracy across 80K+ SKUs.
- Developed an automated multi-site SKU-based web scraping system (Selenium, dynamic DOM parsing, intelligent match logic) processing 100K+ products into MongoDB and SQL for AI-driven substitution analysis.
- Pioneered GenAI analytics using Google Gemini API for automated item matching — improving matching accuracy by 35%, reducing analysis time by 60%, and uncovering an additional **$200K+** in savings.
- Executed large-scale spend analysis (10K–30K SKUs) using advanced Excel modeling, directly influencing $5M+ in procurement decisions.
- Built automated reporting solutions reducing report generation time by 70% and enabling self-service analytics for 50+ users across Procurement, Finance, and Operations.

Marketing Intelligence Analyst (Internship) | Escorts Kubota Limited   *Apr 2023 – Jun 2023 | Faridabad, India*

- Built interactive Qlik Sense dashboards analyzing 200K+ sales records with advanced SQL for a ₹2,000Cr+ revenue organization, enabling data-driven decisions worth ₹16Cr+ quarterly across the tractor division.
- Designed geospatial datasets for sub-district visualizations using SQL spatial analysis, identifying 3 high-potential growth markets for targeted sales expansion.
- Developed an industry-wide market intelligence dashboard consolidating nationwide tractor sales data across 10+ suppliers for competitive benchmarking and territory planning.
- Used QGIS to transform and process KML location files into structured geospatial datasets, enabling sub-district level territory mapping and visualization across dealer networks for targeted sales planning.

KEY PROJECTS

AI Product Substitution Engine (RAG + Chatbot) | Python · Selenium · MongoDB · MySQL · ChromaDB · Gemini Flash 2.5 · RAG

- Scraped 100K+ product SKUs, specs, pricing, descriptions, and images from Office Depot and Staples using Selenium and BeautifulSoup with adaptive rate limiting; stored raw data in MongoDB as the staging layer.
- Transformed and normalized MongoDB data into a structured MySQL schema (3NF) with product hierarchies, categories, and pricing tables — creating a clean, queryable foundation for downstream AI processing.
- Embedded product descriptions and specifications into ChromaDB vector store using Gemini text embeddings, enabling semantic similarity search across the full product catalog.
- Built a RAG pipeline using Gemini Flash 2.5 that retrieves the top 5 functionally equivalent substitution candidates per SKU from ChromaDB — ranked by semantic similarity, unit-normalized price (gallons→oz, feet→inches), and functional equivalence score — with AI-generated reasons for each recommendation.

Power BI Sales & Demographic Dashboard | Power BI · DAX · Data Modeling · Power Query

- Built a 4-page interactive Power BI dashboard — Executive Summary, Sales Performance Analysis, Product Analysis, and Demographic & Category Analysis — with 45+ visuals including KPI cards, line charts, area charts, donut charts, map, pivot tables, ribbon chart, and slicers.
- Authored 15+ DAX measures: Total Revenue, Total Profit, Profit Margin %, Total Orders, Total Customers, Avg Order Value, Avg Customer Value, Total Units Sold, 30-Day Moving Average, Monthly Avg Revenue, Best/Worst Month Display, Best Product Display; delivered age group segmentation, gender distribution, and country-level revenue share analysis.

E-Commerce & Retail Sales EDA | Python · Pandas · Matplotlib · Seaborn · Statistical Analysis

- Conducted end-to-end EDA on two datasets: an e-commerce dataset and a 62K+ row retail sales dataset spanning 8 product categories, 58 stores, and 9 countries — covering revenue trends, brand analysis, outlier detection, and geographic insights.
- Delivered interactive charts, statistical summaries, data quality assessments, and cohort-level breakdowns; findings published as live portfolio pages.

Normalized Database Design & ETL Pipelines (E-Commerce + Retail) | Python · MySQL · ETL · 3NF Schema Design

- Designed two fully normalized 3NF database schemas — an 8-table e-commerce schema and a 7-table retail sales schema — with foreign keys, indexes, and an interactive browser-based SQL playground for live query execution.
- Built automated ETL pipelines in Python transforming denormalized Excel files (62K+ retail rows, 10,500+ e-commerce records) into MySQL with product hierarchies, multi-store orders, and line items; achieved 3–5x faster ingestion via multi-threaded processing.

Multi-Platform Web Scraper (ODP + Staples) | Python · BeautifulSoup · Selenium · JSON · MongoDB · Threading

- Built a dual-platform web scraping system extracting product SKUs, specs, pricing, descriptions, and images from Office Depot (HTML + Selenium) and Staples (internal API reverse-engineering) with intelligent duplicate detection and adaptive rate limiting.
- Processed 100K+ products into structured JSON and MongoDB; used as the data foundation for downstream AI-driven product substitution and spend analysis pipelines.

EDUCATION

Master of Business Administration (MBA) – Business Analytics | CGPA: 7.06/10.0
Jun 2022 – Mar 2024
Symbiosis Institute of Business Management (SIBM), Bangalore

Bachelor of Engineering (B.E.) – Electronics Engineering | CGPA: 7.23/10.0
Apr 2017 – Mar 2021
Yeshwantrao Chavan College of Engineering (YCCE), Nagpur University

CERTIFICATIONS & ADDITIONAL

Certifications: Google Data Analytics Professional Certificate (Coursera) | SQL Intermediate – HackerRank | Power BI Data Analyst Associate (PL-300) – In Progress
Languages: English (Professional), Hindi (Native), Marathi (Native)
Open to: Data Analyst, Business Analyst, Reporting Analyst roles – Nagpur, Pune, Mumbai, Hyderabad, or Remote