

ANKIT GUPTA

Data Analyst • SQL • Excel • Power BI • Python

+91 73037 48364 | ankit25679@gmail.com | New Delhi, India | Graduating 2027

[Portfolio](#) | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL SUMMARY

BCA (Data Science) student graduating in 2027 with hands-on analytics experience across a Power BI internship, four industry virtual simulations and a portfolio of multi-page Power BI reports built on real public data (785K job postings, NPCI's UPI statistics, DMRC ridership, Epoch AI's model database), plus three self-built products running live in production. Comfortable taking a vague business question, cleaning and modelling the data, spotting anomalies, and turning it into a dashboard or short write-up a non-technical stakeholder can act on. Based in Delhi NCR and open to relocating within India.

CORE SKILLS

- **Business Problem Framing** — turning an ambiguous stakeholder question into a measurable, scoped analysis.
- **Data Cleaning & Validation** — missing values, duplicates, outliers, inconsistent records and sanity checks before analysis.
- **Exploratory Analysis & Insight** — segment, compare and quantify to find what is actually driving a number.
- **Dashboard Design & Data Storytelling** — right chart for the question, clean layout, plain-language takeaway.
- **Stakeholder Communication** — explaining findings to non-technical reviewers; asking for direction early on unclear briefs.
- **Documentation & Ownership** — writing up methods and assumptions so work is reusable; shipping and maintaining live products end-to-end.

TECHNICAL SKILLS

- **SQL & Databases:** SELECT queries, JOINS, GROUP BY and aggregations, subqueries, CTEs, window functions; PostgreSQL / Supabase, MongoDB.
- **Spreadsheets:** Microsoft Excel — pivot tables, VLOOKUP / XLOOKUP, INDEX-MATCH, Power Query, slicers, dynamic charts, conditional formatting; Google Sheets.
- **BI & Visualisation:** Power BI — DAX (time intelligence, RANKX / TOPN rankings), Power Query, star-schema modelling, drillthrough, report-page tooltips, decomposition trees, synced slicers, custom themes, PBIP / TMDL source control; Tableau; Excel dashboards.
- **Programming:** Python — Pandas, NumPy, Matplotlib, Seaborn; Jupyter, Streamlit; JavaScript / TypeScript.
- **Statistics & Methods:** descriptive statistics, distributions, correlation, customer segmentation, outlier and anomaly detection, basic forecasting.
- **Cloud & Tools:** AWS Athena, AWS EMR, Oracle Cloud AI, Git & GitHub, Google Analytics 4, Google Search Console, Bing Webmaster Tools.

EXPERIENCE

Data Analytics Intern — Cognifyz Technologies

Remote, India • Jan 2024 – Mar 2024

- Built and delivered **4+ interactive Power BI dashboards** covering sales, operations and investor-behaviour analytics, including the internship capstone.
- Ran end-to-end data cleaning, transformation and modelling in Power Query and DAX, **cutting report generation time by roughly 30%**.
- **Global Super Store:** analysed profit and loss by segment, region and category to identify the loss-driving segments and recommend where discounting should be pulled back.
- **Investment Sector Analysis (capstone):** examined investor behaviour across demographic groups, flagged outliers and inconsistent records during cleaning, and summarised the patterns for a non-technical review.
- Awarded the official **Cognifyz Internship Completion Certificate**.

Virtual Data Analyst Simulations — TATA • Deloitte • Quantum • Lloyds Banking Group

via Forage • 2024

- **TATA — Data Visualisation:** framed executive-level questions, selected the right chart type for each, and prepared a leadership-facing summary.
- **Deloitte — Forensic Analytics:** used spreadsheet analysis to detect anomalies in operational data and classified findings for review.
- **Quantum — Retail Analytics:** performed customer segmentation and transaction analysis to recommend a category strategy.
- **Lloyds Banking Group — Data Science:** worked a banking problem statement from data preparation through to recommendation.
- **Common thread:** clarify an ambiguous brief, clean the data, quantify the business impact, communicate the takeaway in plain language.

POWER BI DASHBOARD PORTFOLIO

Multi-page Power BI reports on real public data · [see all](#)

Power BI · DAX · Power Query · PBIP

- **Data Jobs Market Intelligence (5 pages): 785,741** real 2023 job postings modelled as a star schema (fact → 3.59 M-row skills bridge → skill and date dimensions) with **54 DAX measures**; roles, skills, salaries and geography, with RANKX-over-ALLSELECTED top-N visuals that stay correct under any slicer.
- **The AI Race 1950 → 2026 (8 pages):** Epoch AI's **1,065 notable AI models** — compute frontier on a log scale, training cost, openness and the country race; **76 measures**, a drillthrough lab profile, report-page tooltips, a decomposition tree and a colour-blind-safe palette.
- **UPI Pulse (2 pages): 125 months** of official NPCI statistics (Apr 2016 – Aug 2026) pulled from NPCI's JSON APIs and gap-checked; 24 measures with filter-proof year-on-year KPIs, cumulative and fiscal-year views, app-share and P2P-vs-merchant splits.
- **Delhi Metro in Numbers (5 pages):** line-wise passenger journeys rebuilt from DMRC's monthly PDFs (74.75 crore journeys, Apr–Jul 2026), all **229 stations** on an offline map built from a scatter chart, and a clear note that a "journey" measures line load, not people.
- **Earlier work:** Global Super Store profit and loss, Spotify music trends, crypto market with basic forecasting (Power BI) · Maven Analytics workbook with pivots and slicers (Excel) · IPL Decoded, a 3D web dashboard over 18 IPL seasons (Three.js).

PRODUCT PROJECTS

Resume Banao — AI Resume Tailoring SaaS · [resumebanao.app](#)

Next.js · TypeScript · Postgres

- Designed a **5-provider LLM routing layer** with cost-aware selection, Redis caching and automatic failover — **reducing cost per request by roughly 70%** versus a single-provider setup, tracked against per-provider usage data.
- Built a hybrid ATS scoring system (rule-based + model-based) that reports keyword-coverage breakdowns, and implemented DPDP Act 2023 data access and erasure APIs on a Postgres schema with row-level security.

360 PlayZone — HTML5 Gaming Portal · [360playzone.com](#)

JavaScript · SEO · JSON-LD

- Run a live browser-games portal with **260+ games across 10 categories in 16 languages**, every title playable instantly on mobile or desktop with no download or sign-up.
- Generate the entire site from a **single game-metadata dataset**: one record produces the play page, category placement, FAQ block and embed snippet, with canonical URLs, hreflang and JSON-LD schema layered on top.
- **Treat the site as a dataset**: Google Search Console, GA4 and Bing Webmaster data drive which pages are indexed, promoted or rewritten; monetised through Google AdSense behind a certified consent flow.

CalcVerse — 140+ Tool Calculator Platform · [allsmartcalculators.com](#)

Next.js · TypeScript · SEO

- Shipped a live platform of **140+ calculators across 8 categories**, including an India payroll set (EPF, ESI, TDS on salary, in-hand ↔ CTC) whose figures are verified by a script that replicates the new-regime tax rules.
- All computations run client-side for sub-100 ms response and zero data transmission; schema markup and internal linking grow organic traffic.

College Chatbot — AI Student Helpdesk · [student1chatbot.streamlit.app](#)

Python · Streamlit · NLP

- Deployed an NLP-based helpdesk for academic, fee-structure and campus queries that **reduced manual inquiry handling by roughly 60%**.

J.A.R.V.I.S. — Voice AI Assistant

Python · Whisper · ChromaDB · LangChain

- Built a voice assistant with **50+ English and Hinglish commands**, multi-engine routing with automatic failover, and long-term semantic memory on a ChromaDB vector store.

EDUCATION

Bachelor of Computer Applications (BCA) — Data Science

Graduating 2027

Aravali College of Engineering & Management (affiliated to YMCA University) — Faridabad, Haryana, India

- **Relevant coursework:** Data Analytics, DBMS and SQL, Python Programming, Statistics, Data Visualisation, Machine Learning.

Data Analytics Certification Program — WsCube Tech

2024

- End-to-end analytics workflow: Python, statistics, data visualisation (Matplotlib, Seaborn) and SQL.

CERTIFICATIONS

- **Data Analytics Essentials** — Cisco (2026) · **Amazon Athena & Amazon EMR** — AWS Training (2026)
- **Forge Simulations:** Data Visualisation — TATA (2024) · Data Analytics — Deloitte (2024) · Data Analytics — Quantum (2024) · Data Science — Lloyds Banking Group (2024)
- **Data Science & Analytics** — HP LIFE (2024) · **Oracle Cloud AI Professional, Gen AI Infrastructure** — Oracle (2025) · **Digital Marketing & Gen AI** — Adobe / Coursera (2025)

ADDITIONAL INFORMATION

- **Languages:** English — professional working proficiency · Hindi — native.
- **Availability:** Available for full-time roles; open to in-office work and to relocating within India — Hyderabad, Bengaluru, Gurugram, Mumbai or Delhi NCR.