

PRIYANSHU SINGH

Data Analyst • SQL • Python • Data Validation • Business Intelligence

Gurugram, India | maanapriyanshu@gmail.com | +91-9102462129 | [linkedin.com/in/maanapriyanshurajput](https://www.linkedin.com/in/maanapriyanshurajput) | [GitHub](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

Junior Data Analyst with experience cleaning, analyzing, and reporting on data across pharmaceutical, retail, and insurance datasets. Work independently and present findings to managers and project stakeholders. Projects include a retail store performance analysis across 1,115 German stores and a Random Forest forecasting model reaching R^2 of 0.848 on 213K records. MSc in Data Science at Chandigarh University, expected 2026. German B1, looking for analyst roles in Germany.

PROFESSIONAL EXPERIENCE

Data Analyst - Contract | Ikvans Healthcare Pvt. Ltd. Aug 2025 - Feb 2026

- Cleaned and analyzed 30,000+ pharmaceutical sales records in SQL and Excel to track demand patterns across distribution cycles. Flagged slow-moving inventory early enough for the team to act before it hit stock levels.
- Built monthly sales and distributor performance reports and presented findings to management, covering territory performance and procurement trends.
- Segmented products by sales velocity into fast- and slow-moving categories, which informed pricing and inventory planning discussions.
- Used Excel for data validation and pivot-based reporting across the full record set before migrating the workflow to SQL.

Data Analyst Intern | AAM Infotech Pvt. Ltd., Gurugram Jan 2025 - Jul 2025

- Cleaned and validated datasets in Python (Pandas, Regex) and SQL across multiple sources, fixing duplicate records, format mismatches, and referential integrity issues.
- Built Power BI dashboards for KPI tracking that cut out 4 hours of manual reporting per day, automating recurring reports that previously ran on Excel.
- Presented dashboard findings to the team and updated visuals based on feedback from managers.

PROJECTS

Rossmann Performance Insights • PostgreSQL • Python • Power BI • DAX – [GitHub](#)

- Analyzed 844,338 sales records across 1,115 German retail stores to answer 8 business questions on store performance, promo ROI, competition proximity, and seasonal demand. Promotions lifted daily sales by 39% on average, but the effect varied: Type a stores responded at 43% vs 18% for Type b, a gap large enough to suggest promo budget allocation could be reviewed by store type.
- Found that stores within 500m of a competitor averaged 14% higher daily sales than isolated stores. The data pointed the opposite direction from the usual assumption about nearby competition.
- Caught two issues that would have produced wrong conclusions before presenting any findings: only 33 of 1,115 stores trade on Sundays, which skewed day-of-week averages, and 2015 data runs only through July, making raw YoY totals about 36% lower than a fair comparison would show. Corrected both.
- Used PostgreSQL to identify 630 stores below portfolio average, flag stores with statistically abnormal sales days using Z-score analysis, and build a weighted scoring model to prioritize stores for expansion investment. Queries used LAG, LEAD, NTILE, DENSE_RANK, PERCENT_RANK, ROW_NUMBER.
- Built a 4-page Power BI dashboard with 14 DAX measures covering executive KPIs, store performance ranking, promo impact by segment, and growth opportunity scoring.

Predictive Analytics for Sales Forecasting • Python • Scikit-learn • SQL • Power BI – [GitHub](#)

- Built a Random Forest regression model on 213,328 insurance agency records (147,760 after cleaning) to forecast written premium, reaching $R^2 = 0.848$ compared to 0.309 for a linear baseline, nearly a 3x gain in explained variance.

- Found incurred losses and new business premium to be the strongest predictors through feature importance analysis.
- Designed a 3-page Power BI dashboard for non-technical stakeholders, surfacing \$4.01B in total premium, 88.75% customer retention, and a regional concentration risk. Ohio alone made up 58% of the portfolio.

TECHNICAL SKILLS

Languages: SQL (CTEs, Window Functions, Joins), Python (Pandas, NumPy, Regex, Scikit-learn, XGBoost)

Data Processing: ETL, Data Cleaning, Data Validation, Feature Engineering, Reporting Automation

BI & Visualization: Power BI (DAX, KPI Cards, Slicers, Drill-through), MS Excel (Advanced), Matplotlib, Seaborn

Databases: PostgreSQL, MySQL, Google BigQuery

Tools: Git, GitHub, VS Code, Jupyter Notebook

Machine Learning: Predictive Modeling, Regression Analysis, Random Forest, Model Evaluation, Statistical Analysis

EDUCATION

MSc in Data Science | [Chandigarh University](#) Aug 2024 - Jul 2026 (Expected)

Bachelor of Computer Applications (BCA) | [L.N. Mishra College of Business Management](#) 2020 - 2023

CERTIFICATIONS & LANGUAGES

Data Analytics Training | [Ducat IT Training School](#) Jul 2024 - Jul 2025

IBM Data Science - [Coursera](#) **Career Essentials in Data Analysis** - [Microsoft & LinkedIn](#)

Journey into Data Science with Python - [CodeSignal](#)

English (Professional) **Hindi** (Native) **German** (B1 - actively learning)