

Kshitij Nailwal

[LinkedIn](#) • [Github](#)

PROFESSIONAL SUMMARY

Performance-driven Data Analyst with 11 months of experience in automation, business intelligence, and AI evaluation at Google (via Vaco Binary Semantics). Expert in Python, SQL, and advanced data visualization. Successfully automated HITL workflows, improving evaluator efficiency by 20% and reducing manual effort. Strong foundation in machine learning, feature engineering, prompt engineering, and model building using scikit-learn. Skilled at transforming raw data into actionable insights, building dashboards, and supporting large-scale AI model evaluation for high-impact business decisions.

EDUCATION

Bachelor of Technology (Computer Science) University Of Petroleum And Energy Studies(UPES), Dehradun	2020 - 2024 Grade: 8.81/10.0
Intermediate (Class XII) Riverdale International School, U.S. Nagar	2019 - 2020 Grade: 80.0%
Matriculation (Class X) Sainik School Ghorakhal, Nainital	2017 - 2018 Grade: 80.4%

EXPERIENCE

Associate Analyst January 2025 - Present
Vaco Binary Semantics/Highspring (Onsite : Google) *Gurugram*

- Automated Human-in-the-Loop (HITL) evaluation workflows using Python + Google Sheets, increasing accuracy and saving 15+ hours weekly of manual effort.
- Built data quality and evaluator performance dashboards using PLX and Looker Studio as part of the Business Intelligence team, enabling leadership to track trends and operational KPIs.
- Designed scalable data workflows supporting AI model evaluation, improving annotation consistency and boosting labeling accuracy by 20%.
- Conducted data acquisition, cleaning, manipulation and anomaly detection using Sheets, SQL, and Python, reducing reporting errors by 25%.

INTERNSHIPS

Data Analytics Intern March 2024 - January 2025
AccioJob *Remote*

- Cleaned and transformed large datasets using Pandas, NumPy, Power Query, and Excel.
- Built dashboards using Power BI to simplify insights across customer, product, and sales dimensions.
- Utilized regression, clustering, decision trees and scikit-learn to build predictive models with 12% increase in forecasting accuracy.
- Wrote optimized SQL queries for ETL and analysis of complex datasets.
- Developed knowledge in Generative AI — LLMs, transformer models, diffusion models, prompt engineering, and evaluation.

PROJECTS

Optimizing Online Content for Enhanced Social Media Engagement ([Github](#)) ([Demo](#)) October 2024

- Analyzed 30,000+ news articles to identify key factors influencing viral reach.
- Applied NLP techniques for sentiment scoring and feature extraction.
- Built regression models to quantify impact of title, keywords, sentiment, and media elements on engagement.
- Tools: Python, Pandas, Scikit-learn, Matplotlib

Sales Analytics ([Github](#)) ([Demo](#)) August 2024

- Built a Power BI dashboard from 10+ relational tables for a wholesale business case.
- Used DAX, interactive visuals, KPIs, and data modeling to answer 30+ business questions.

- Helped identify customer segmentation opportunities, regional growth trends, product performance, and supply chain gaps.
- Tools: Power BI, Excel, SQL, DAX

CERTIFICATIONS

Cutshort Advanced Data Science Certification Cutshort ([Link](#)) November 2024

- Successfully passed a skill validation assessment in Data Science conducted by Cutshort.
- Certified at the Advanced proficiency level based on performance in an assessment test.

BCG X Virtual Experience Program – Data Science Job Simulation Forage ([Link](#)) November 2024

- Successfully completed a virtual job simulation offered by BCG X through Forage
- Hypothesis testing, EDA, feature engineering, modeling & business insights

SKILLS

Computer Languages: SQL, Python, Machine Learning, Python, SQL

Data Tools: Power BI, Power BI, NumPy

Software Packages: Excel, Pandas, Matplotlib, Excel, MySQL

Others: Generative AI, Analytics