

Harsh Koli

+91 9518951072 harshkoli568@gmail.com [linkedin.com/in/harsh-koli21](https://www.linkedin.com/in/harsh-koli21) github.com/harshkoli21
Mumbai, India

Professional Summary

Data Analyst with hands-on experience in data collection, cleaning, statistical analysis, and stakeholder communication. Currently pursuing MSc in Data Science (University of Essex) with strong proficiency in Python, SQL, and Excel for data manipulation, preprocessing, and reporting. Proven ability to translate complex data findings into clear, actionable insights for both technical and non-technical audiences. Experienced working cross-functionally in fast-paced, remote team environments with a strong ownership mindset.

Technical Skills

Languages & Tools	Python (pandas, numpy, scikit-learn), SQL (PostgreSQL, PL/SQL), Excel (pivot tables, VLOOKUP, data validation), R
Data Analysis	Data Cleaning & Preprocessing, Statistical Analysis, Exploratory Data Analysis, Feature Engineering, A/B Testing, Pattern & Trend Identification
Visualisation & BI	Power BI, Matplotlib, Python-based Dashboards, Executive Reporting
Data Engineering	Scalable Data Pipelines, OCR Pipelines, Structured & Unstructured Data Processing, API Data Collection
Soft Skills	Stakeholder Communication, Cross-functional Collaboration, Client Support, Attention to Detail, Problem-Solving
Governance	Data Integrity & Accuracy, PII-Aware Data Handling, Responsible AI, Data Governance

Work Experience

AI & Data Engineering Extern

December 2025 – Present

Extern, Inc. (Outamation Project), Remote

- Collected and processed data from multiple unstructured sources, building Python-based pipelines for structured data extraction and downstream analytics.
- Applied rigorous data quality validation and transformation techniques across large datasets, ensuring accuracy, completeness, and integrity for reporting and model consumption.
- Collaborated cross-functionally with engineering and analytics teams to deliver scalable automation solutions, communicating progress and findings to stakeholders regularly.
- Ensured responsible data handling including PII-aware processing and governance compliance in a production-adjacent environment.

Data Analyst Intern

September 2023 – December 2023

Trinity, Remote

- Collected, organised, and analysed large structured datasets using Python, SQL, and Excel to identify trends, patterns, and correlations across customer and business datasets.
- Performed systematic data cleaning, validation, and preprocessing to improve data quality and reliability for downstream reporting and business use cases.
- Translated complex analytical findings into clear, actionable business strategies and communicated insights to non-technical stakeholders through visual summaries and reports.
- Independently managed end-to-end analysis tasks on real-world business challenges, demonstrating strong ownership and quality-focused delivery.

Education

University of Essex

Postgraduate Diploma in Data Science

Relevant Modules: Machine Learning, Statistical Modelling, Predictive Analytics, Data Engineering, Responsible AI

University of Mumbai

BE in Data Engineering, 2:1

Colchester, UK

2025 – 2026

Mumbai, India

June 2021 – May 2024

Projects

Credit Risk & Fraud Detection: Predictive Modelling

2024

Python, scikit-learn, SQL, XGBoost, pandas

- Collected and cleaned transactional financial datasets; engineered domain-relevant features and applied resampling strategies (SMOTE) to handle class imbalance.
- Performed statistical analysis across risk segments and threshold settings, monitoring model performance and identifying patterns and correlations in complex datasets.
- Communicated model outputs and risk findings in a clear, business-readable format for non-technical stakeholders.

Customer Purchase Prediction & Recommendation System

2023

Python, Supervised Learning, pandas, Excel

- Analysed large transactional customer datasets to identify purchasing behaviour patterns and segment customers using statistical and classification techniques.
- Cleaned and validated behavioural data, applying feature transformation to prepare datasets for modelling and reporting.
- Translated model outputs into clear pricing and targeting recommendations for non-technical business stakeholders, supported by visual summaries.

Additional Experience

Tata Group: Data Visualisation Job Simulation

Forage, 2024

- Produced executive-level dashboards and data stories to communicate analytical findings and support strategic decision-making for senior stakeholders.

JPMorgan Chase: Investment Banking Job Simulation

Forage, 2024

- Analysed financial data and communicated risk findings in a concise, client-ready format, reinforcing skills in stakeholder reporting and data-driven decision support.

Professional Interests

Data Analysis & Business Intelligence — Financial & Customer Data — Responsible AI — Continuous Learning & Technical Certifications