NEHA DAHIYA

DATA ANALYST

nehaadahiya | nehaadahiya07@gmail.com | nehaadahiya | +91 8955083757

EDUCATION

Bachelor of Technology | Computer Science Engineering
Gautam Buddha University
2021-2025 CGPA: 9.18

Relevant Coursework: Data Structures & Algorithms, Machine Learning, Artificial Intelligence, Web Development

Certification:
IBM Data Analyst Professional Certificate | Coursera, June 2025

Volunteering:
IEEE IC3ECSBHI-2025, UNESCO India Africa Hackathon

Hobbies: Creative writing, data exploration, graphic design

SKILLS

- **Languages & Tools:** Python (Pandas, NumPy, Matplotlib, Seaborn), SQL (MySQL, PostgreSQL), Excel
- **Data Visualization:** Power BI, Tableau, Matplotlib
- **Databases & Querying:** Joins, Aggregations, CTEs, Window Functions, Indexing, Query Optimization
- **Machine Learning:** XGBoost, Isolation Forest, SMOTE, Logistic Regression
- **Others:** GitHub, Jupyter Notebook, VS Code

LANGUAGE

English
Hindi

PROFILE

Analytical and detail-oriented Computer Science undergrad with hands-on experience in data cleaning, visualization, and business insight generation using Python, SQL, Power BI, and Tableau. Strong background in end-to-end data projects including fraud detection, customer churn analysis, and dashboard development. Eager to drive decisions through data storytelling and technical excellence.

EXPERIENCE

Data Analyst Intern – Unified Mentor, Gurugram (Remote)
Jan 2025 – Mar 2025

- Conducted EDA on Netflix dataset (8,000+ rows) using Python (Pandas, Seaborn, Matplotlib) to uncover genre and region-based content trends.
- Built a Tableau dashboard enabling pattern recognition by genre, release year, and geography.
- Enhanced storytelling with data and improved content visibility through visual analysis.

PROJECTS

Customer Churn Analysis – SQL & Power BI
Tools: MySQL, Power BI, SQL Window Functions, CTEs

- Analyzed telecom dataset (~100K records) to calculate monthly churn rate and identify high-risk customer segments using SQL.
- Used CTEs, CASE WHEN, and window functions to generate churn cohorts and flag high-risk profiles.
- Built an interactive Power BI dashboard with filters (contract type, tenure, charges) to help visualize churn trends and propose retention strategies.

Health Insurance Fraud Detection System – Python, Streamlit, ML
Tools: Isolation Forest, XGBoost, SMOTE, Streamlit, Matplotlib, SQLite

- Analyzed 583K+ insurance claims, detecting ₹3.2M+ in suspected fraud using ensemble models (F1: 82.5%, AUC: 90.3%).
- Highlighted key fraud indicators such as geographic anomalies, physician claim trends, and excessive reimbursements.
- Built a Streamlit dashboard with bar charts and histograms to present fraud probabilities and feature insights for audit teams.

Stock Performance Dashboard – Python, Plotly
Tools: Python

- Scraped 5 years of stock data for Tesla, Amazon, and AMD using yfinance and web scraping.
- Built an interactive dashboard using Plotly to visualize volatility patterns and long-term trends.
- Provided actionable insights into stock movement and investment timing.