

SOUMAJYOTI DHUT

Phone : +91 - 7029658149 | Email : soumajyotidhut@gmail.com

LinkedIn : [linkedin.com/in/soumajyoti-dhut](https://www.linkedin.com/in/soumajyoti-dhut)

GitHub : github.com/SoumajyotiDhut

SUMMARY

Versatile tech enthusiast with a passion for data science and data analyst, problem-solving, and building meaningful digital solutions. Experienced in handling real-world datasets through personal projects involving data cleaning, visualization, and predictive modeling. Self motivated, curious, and committed to continuous learning and turning raw data into actionable strategies for business decision-making.

EDUCATION

TECHNO MAIN SALT LAKE ,KOLKATA

Bachelor of Technology in Information Technology; CGPA - 7.12 (till 6th Semester) June 2022 - June 2026

BARADONGAL RAMANATH INSTITUTION (H.S)

Higher Secondary - 81% April 2020 - March 2022

SKILLS

- **Languages:** Python , Java, SQL
- **Frameworks:** NumPy ,Pandas, Matplotlib, Seaborn
- **Tools:** Power BI, Excel, MySQL
- **Platforms:** Jupyter Notebook, Visual Studio Code, IntelliJ IDEA
- **Concepts:** DBMS , OOPS , Computer Networks, SDLC

PROJECTS

E-commerce Data Analysis | [Link](#) May 2025 - June 2025

- Executed 15+ advanced SQL queries (JOINS, subqueries, CTEs, window functions) on 8 relational tables with 2,900+ order records to analyze customer behavior, payments, and product sales.
- Analyzed data from 122 customers, 273 payments, and 110 products using Pandas & SQL, and built visualizations with Matplotlib & Seaborn to uncover order trends, top-performing categories, and regional preferences.

Fake News Detection Web App | [Live Link](#) June 2025 - July 2025

- Built a Fake News Detection web application using Flask, scikit-learn, and TF-IDF vectorization, achieving over 95% accuracy with Random Forest Classifier.
- Integrated confidence score display and highlighted top influential keywords contributing to model predictions.
- Implemented NLP preprocessing, text cleaning, and TF-IDF feature extraction to classify news articles as FAKE or REAL .

Regional Sales Analysis | [Link](#) July 2025 - July 2025

- Performed EDA on 64K+ sales records (21 features, 5 years) to uncover regional and product-level insights.
- Preprocessed and merged multi-sheet Excel datasets (Sales, Products, Regions, Budgets) into a unified dataset for analysis & visualization.
- Built an interactive Power BI dashboard with drill-downs by region, product, channel, and time, highlighting trends and top products (Products 25 & 26 = 25%+ revenue).

ACHIEVEMENTS

- **1st Runner-up** among 71 teams in a college-level project hackathon organized by the IT department.
- Achieved 97.34 percentile in Naukri Campus Young Turks Assessment. | [CERTIFICATE](#)
- HackerRank Python (Basic) Certification – Passed official skill certification test . | [CERTIFICATE](#)
- HackerRank SQL (Advanced) Certification – Passed official skill certification test . | [CERTIFICATE](#)

CERTIFICATES

Introduction to Data Analytics – Meta | [CERTIFICATE](#) Aug 2025

- Applied OSEM framework to marketing data, assessing gaps and dataset quality.
- Analyzed and compared structured, semi-structured, and unstructured data formats for real-world applications

Introduction to Data Science – Cisco Networking Academy | [CERTIFICATE](#) July 2025

- Gained foundational knowledge in analytics, data handling, AI/ML concepts, and career insights in data science.
- Covered experience analytics, data handling, and machine learning, with a final exam certification.