

SUBHRAJYOTI MAHANTA

+91 6002967278 subhrajyoutimahanta@gmail.com github.com/Subhrajyoti subhrajyotimahanta subhrajyoti.online

EDUCATION

Indian Institute of Engineering Science and Technology, Shibpur, Howrah

2021 - 2025

Bachelor of Technology, Civil Engineering

CGPA 7.06

PROJECTS

• GoodCabs Performance Analysis – SQL & Power BI Dashboard : [↗](#)

- Engineered an interactive Power BI dashboard to analyze 500,000+ trips, optimizing data visualization for **business insights**.
- Architected a robust **PostgreSQL database** with 7+ tables, ensuring seamless storage and retrieval of operational data.
- Executed SQL queries to track trip trends, revenue patterns, and customer retention reporting **10+ ad-hoc** requests.
- Unearthed revenue gaps, revealing repeat passengers contribute 40%+ more per trip, leading to **strategic recommendations**.

Tools Techniques - SQL, Power BI, PostgreSQL, Data Analysis, Business Intelligence, Ad-Hoc Reporting, Visualization, BI Tools, Marketing Insight Analytics, Marketing platforms, Data governance, ETL

• Data Analyst Job Market Insights: Trends, Salaries, and Optimal Skills : [↗](#)

- Analyzed **10,000+ job postings** from Luke Barousse's dataset to identify in-demand skills, salary trends, and career opportunities for data analysts.
- Developed 5+ interactive visualizations and dashboards in **Excel (Power Pivot, Power Query)** and Jupyter Notebooks to showcase salary trends and skill demand.
- Orchestrated **interactive reports and visualizations** in Jupyter Notebook, uncovering that SQL appeared in 50%+ of job postings and Python in 72% of Data Scientist roles
- Optimized career insights by identifying high-paying and high-demand skills, revealing Oracle as the top-paying skill (\$97K) and Excel as the most in-demand.

Tools Techniques - Excel, Power Pivot, Power Query, Jupyter Notebook, Python, Pandas, Seaborn, Matplotlib, Data Visualization,

• AI-Driven Blood Report Parser & Health Insight Dashboard : [↗](#)

- Engineered an end-to-end data pipeline using **Tesseract OCR** and Unstructured Library, accurately extracting text data from 200+ scanned and native PDFs.
- Implemented text preprocessing and **NLP techniques** (OpenCV, spaCy, NLTK) to enhance parsing accuracy by over 90%, structuring complex medical data via **LangChain and GPT-4**
- Deployed a scalable **RESTful API** using **FastAPI** with PostgreSQL for secure, real-time access to patient data, and visualized health metrics through **Tableau dashboards** to deliver actionable insights that improved clinical decision-making efficiency by 30%.

Tools Techniques - Gen AI, Tesseract OCR, Unstructured Library, Text Preprocessing, OpenCV, spaCy, NLTK, LangChain, GPT-4, NLP, Predictive Analytics, Python, PostgreSQL, RESTful APIs, JSON

WORK EXPERIENCE

Summer Research Internship

may 2023–July 2023

Mentor: **Dr. AK Maurya**

Indian Institute of Technology, Guwahati

- Researched and applied LIDAR technology for transportation engineering.
- Spearheaded a research project to improve the accuracy of traffic flow analysis by 20%.

Research Internship

July 2023–may 2025

Mentor: **Dr. Anuj Budhkar**

IIESTS

- Built LIDAR systems to collect and analyze traffic data.
- Investigated 100+ traffic samples to identify EV impact patterns.

Senior Trainer

August 2023–June 2024

Organization: **IHFC, IIT Delhi**

Indian Institute of Technology, Delhi

- Facilitated 50+ robotics and programming sessions in 10+ colleges in West Bengal, reaching approximately 300 students in the NURTURE program initiated by IHFC, IIT Delhi.
- Orchestrated a robotics competition for 10+ schools and colleges, engaging 100+ students and promoting STEM interest.

SKILLS

Languages — SQL, Python, R • **Tools** — Power BI, MS Excel, Tableau, Pandas, NumPy, SciPy, Matplotlib • **Database** — PostgreSQL, SQLite, MySQL • **Cloud** — Azure • **Big Data** — Google BigQuery • **Automation** — n8n, LangChain