

# SANYA SETIA

UK | P: +44 7826988866 | [LinkedIn](#) | [GitHub](#)

Results-driven Data Professional with expertise in **A/B testing, trend analysis, SQL, Python, and BI reporting**. Proven ability to **optimize data pipelines (60M+ rows/day)**, develop **real-time dashboards (Power BI, DOMO)**, and conduct **statistical modeling to drive business decisions**. Experienced in **data-driven storytelling**, passionate about leveraging **AI** to deliver faster insights in a world that is shifting at an unimaginable pace.

## WORK EXPERIENCE

---

### UNIVERSITY OF CAMBRIDGE – Career Accelerator with ML and AI

Data Scientist

Remote

Nov 2024 – Present

- Developed **NLP models**, anomaly detection systems, sentiment analysis, and financial forecasting models as part of a **7-month intensive program** focused on deep learning and data science fundamentals.
- Delivered data-driven solutions for a Global Gym Chain and a Universal Bank, providing actionable insights and strategic recommendations.
- Applied advanced machine learning techniques to real-world datasets, optimizing model performance and improving decision-making processes.

### TECHSPERT (AI-driven Healthcare Scale-Up)

Business Intelligence Developer

Cambridge, UK

Sep 2022 – Present

- Developed and maintained a centralized data visualization platform, integrating data from multiple sources (Salesforce, Redshift, Hubspot, BambooHR). Built BI dashboards for business KPIs (NRR, Client Reports, ROI), automating reports and saving 10+ hours weekly
- Optimized and developed robust SQL queries and ETL pipelines (**PostgreSQL, MySQL**), processing **60M+ rows/day** with **99.9% uptime** to ensure data accuracy
- Aggregated unstructured data from 5+ sources to build the foundation fields for a **new AI-driven Platform**; forecasted **£850,000 at >95% gross margin**
- Built an AWS Lambda UDF in PostgreSQL for millisecond-accurate business hour calculations. Led data integration across Redshift, S3, and Gainsight, improving analytics speed and consistency
- Support decision-making: Provide clean, transformed data to multiple stakeholders in a way that empowers them to answer their own questions

### UNIVERSITY OF LEEDS

Education Outreach Fellow

Leeds, UK

Dec 2021 – Dec 2022

- Curating and Delivering Business Workshops across UK schools for more than **100 participants** successfully
- Presenting interactive **2- hour Workshops** for Masters in Business Analytics Taster sessions for A-level students
- Presenting complex data and analysis concepts in a comprehensive, lucid manner

### BAIN AND COMPANY

Data Analyst

Remote - India

Mar 2021 – Aug 2021

- Led, Structured insights and proposals for a healthcare chain looking for investment opportunities, delivered clear and professional presentations
- Delivered strategic insights to senior management on **200+ healthcare investment opportunities**
- Databased **1000+ use cases** using software; Excel, Tableau and Alteryx

### THE SMART CUBE

Research Analyst

Remote - India

Nov 2020 – Mar 2021

- Consulted major defense and air services on sustainability practices; Consulted UK's police fleet of more than **40 counties**
- Analyzed market and delivery systems to assess present and future material availability and delivered a relevant report to clients

### NATIONAL UNIVERSITY OF SINGAPORE

Research Fellow

Singapore

Jan 2019 – Aug 2021

- Conducted individual lab experiments under the guidance of Director of Centre of Water Research, Professor Hu Jiangyong
- Researched the inflow and outflow of more than 50 specimens daily using Catalytic Ozonation using Excel and MATLAB for the project **“Improving Efficiency of treatment of Petrochemical Wastewater”**

## EDUCATION

---

<b>Institute of Continuing Education, University of Cambridge</b> Career Accelerator in Data Science with Machine Learning & Artificial Intelligence Focus on Deep Learning Models and Neural Networks with Hands on Project with major collaboration with a Universal Bank and 4 Deep Learning Projects Relevant Coursework: Hypothesis Testing, Regression; Clustering; Algorithms; Artificial Intelligence	Cambridge, UK Expected July 2025
<b>University of Leeds</b> MSc Business Analytics and Decision Sciences ( <b>Distinction</b> ) Core Modules that focus on digital transformation, machine learning, advanced analytics, Management decision making, and storytelling with data with like R Studio and Power BI	Leeds, UK Sep 2021-May 2022
<b>PEC University of Technology</b> Bachelor of Technology in Civil Engineering ( <b>CGPA 8.1/10</b> ) Top 15% of the Academic Year, Awarded ' <b>College Colour</b> ' for exceptional performance in academics, Administration and co-curricular activities	Chandigarh, India July 2016-May 2020

## AI&ML PROJECTS

---

<b>TOPIC MODELLING</b> ● Applied <b>BERTopic</b> to analyze customer reviews from Google and Trustpilot for an International Gym Chain ● Conducted Sentiment Analysis on <b>10,000+ reviews</b> to classify positive and negative reviews ● Identified key issues by location and extracted insights using <b>Microsoft-phi4-mini-instruct (LLM)</b> <b>Skills: NLP, Topic Modeling, BERT, LDA, Data Visualisation, Python (Pandas, NLTK, Scikit-learn)</b>	March 2025
<b>SENTIMENT ANALYSIS</b> ● Dataset from Hugging Face Library, converted text to vectors using <b>TF-IDF</b> and trained on <b>Logistic Regression Model</b> ● Accuracy Metrics – <b>Precision, Recall, f1-score</b> , with weighted average accuracy of 78% and f1-score of 60%. <b>Accuracy f1-score above 50%</b>	Mar 2024
<b>ANOMALY DETECTION</b> ● Built an outlier detection system for ship engines using <b>IQR, One-Class SVM, and Isolation Forest</b> ; resulting in Early fault detection, <b>reducing maintenance costs</b> ● Optimised model parameters to maintain an <b>anomaly detection rate of 1-5%</b> , ensuring reliability	Nov 2024

## ADDITIONAL SKILLS

---

**Programming & ML:** Python, SQL, TensorFlow, PyTorch, Hugging Face, Scikit-Learn, spaCy, BeautifulSoup  
**Data Engineering:** AWS (S3, Redshift, Lambda), ETL Pipelines, PostgreSQL, MySQL  
**BI & Analytics:** Power BI, DOMO, Salesforce  
**Version Control & DevOps:** Git, VS Code, Jira  
**Knowledgeable:** Databricks, Docker, Apache Airflow, Databases (AWS, GCP), Spark

**Certifications & Training:** AWS: Generative AI for Executives, AWS: Fundamentals of Machine Learning and AI, AWS: Responsible Artificial Intelligence Practices

## LEADERSHIP EXPERIENCE

---

<b>Guest Speak: Data Science Discussion Panel</b>   University of Leeds	Feb 2024
<b>Education Outreach Fellow (University Representative)</b>   University of Leeds, Leeds	Dec 2021 – Dec 2022
<b>Sports In-charge (Sports Secretary)</b>   Punjab Engineering College, Chandigarh	June 2019 – June 2020
<b>Branch In-charge Head</b>   Punjab Engineering College, Chandigarh	June 2019 – June 2020

## AWARDS AND ACHIEVEMENT

---

**Certificate of Excellence** | Punjab Engineering College, Chandigarh  
**Budding Water Professional Award** | National Hydrology Project, and Ministry of Water Resources India  
**Certificate of Appreciation** | Punjab Engineering College, Chandigarh