

Lucas Kent

Data Analyst

07740383917

LSK1@hotmail.co.uk

<https://www.linkedin.com/in/lucas-kent-212a1b16a/>

https://isabella13022012.github.io/lucas_kent/

Flat 40, Vida House

50 Trundleys Road

Bermondsey

London

SE8 5EJ

Core Skills

AI / Machine Learning

- Supervised & Unsupervised Learning
- Regression
- Classification
- Clustering
- PCA
- Time-Series Analysis

Natural Language Processing

- Sentiment Analysis
- Topic Modelling
- Document Classification
- Text Preprocessing
- TF-IDF
- NER Foundations

Deployment & Architecture Awareness

- Flask APIs
- Model Packaging
- Virtual Environment Management
- CI/CD Pipelines for model deployment in GitHub Actions and Jenkins

Web & Data Ingestion

- Selenium
- BeautifulSoup
- Requests

Tools & Platforms

- Git/GitHub
- SQL
- Excel
- Power BI
- Tableau
- PowerShell for automation and data management

Data Pipelines & Engineering

- ETL/ELT pipelines for large datasets
- Feature Engineering
- Model Evaluation
- Cross-Validation
- Hyperparameter Tuning

Libraries & Programming Frameworks

- Python
- Pandas
- Numpy
- Scikit-learn
- SpaCy
- NLTK
- VaderSentiment

Key Projects

Aspect-based sentiment analysis of London hospital reviews: 2020-2025

February 2025 to July 2025

- Developed large-scale data pipeline analysing 207,000+ hospital reviews
- Designed rigorous multi-stage data preprocessing workflows (tokenisation, lemmatisation, TF-IDF, encoding) to prevent data misclassification and duplication
- Combined manual annotation and machine learning to improve data integrity and classification reliability
- Applied sentiment analysis models to generate 3-dimensional sentiment labels
- Used PCA for dimensionality reduction and improved model interpretability
- Achieved 96% classification accuracy, improving performance by 13% through GridSearchCV optimisation
- Delivered advanced visual analytics including geographic mapping and time-series trend analysis

Customer Survey ML Analysis – Airline Industry

https://isabella13022012.github.io/lucas_kent/projects/proj-1

June 2024 to July 2024

- Developed supervised and unsupervised ML models from a 129,000-row mixed-type dataset
- Applied PCA, SVM, Random Forest, and XGBoost for classification and feature importance analysis
- Conducted scenario analysis and synthetic result generation to explore business outcomes
- Delivered findings through 8 clear, stakeholder-friendly visualisations

Employment history

Freelance data scientist– Self-employed

August 2024 to present

- Performed rigorous data quality checks including, cleaning, de-duplication, and alignment of customer and operational data, guaranteeing 100% data integrity
- Wrote and optimised complex SQL queries for 500,000-row+ datasets to generate actionable business insights
- Performed exploratory data analysis and statistical testing to identify trends and validate hypotheses
- Optimised and built classification, NLP, and regression models using cross-validation, hyperparameter tuning, and performance benchmarking, achieving an average of 42% reduction in execution time
- Delivered 100+ clear visualisations and executive summaries for both technical and non-technical stakeholders
- Developed reproducible ML pipelines for structured and unstructured data, ensuring performance, stability, and transparency
- Produced proof-of-concepts to test new modelling approaches and technologies
- Employed Google BigQuery to run large-scale queries and manage large datasets, optimizing query performance and storage costs by 24%
- Delivered actionable insights through over 50 data visualizations and reports for both technical and non-technical stakeholders using tools such as Power BI and Tableau

Freelance model evaluator – Outlier

October 2025 to present

- Evaluated and refined data-analysis-focused AI 5+-step workflows, identifying logical gaps and inefficiencies
- Collaborated with groups of up to 20 people in Agile environments, delivering work to tight schedules with consistent quality
- Assessed analytical reasoning, code structure, and reproducibility of model outputs

Freelance model evaluator – Dataannotation

May 2025 to present

- Designed structured evaluation frameworks for analysing AI-generated analytical solutions
- Reviewed and scored model outputs for correctness, efficiency, and clarity
- Developed evaluation criteria, creating atomic, self-referencing measures to assess model responses systematically
- Conducted response evaluation, assessing and rating AI model outputs against the established criteria to ensure rigorous performance analysis

Compliance Caseworker – HM Revenue & Customs

March 2025 to present

- Conducted structured data analysis in Excel and Power BI, including time-series and 95% confidence interval outlier analysis
- Produced executive summaries and dashboards for 20+ stakeholders
- Ensured strict adherence to compliance, data protection, and internal policy standards
- Managed stakeholder communications, coordinating schedules, correspondence, and meetings to ensure timely project delivery, resulting in 95% project punctuality

Customer Services Consultant – HM Revenue & Customs

September 2021 to March 2025

- Delivered 30+ analytical reports and visualisations to senior leadership
- Developed strong stakeholder communication and requirements-gathering skills, handling more than 13,600 interactions with professionalism and efficiency
- Trained and developed over 50 colleagues in a floorwalking role, enhancing team capability and performance

Education

BSc Data Science – Birkbeck UoL, (part-time)

September 2025 to present

Professional Certificate in Data Analytics – Imperial College Business School

November 2023 to June 2024

Microsoft Power BI Data Analyst Professional Certificate

August 2024 to present

Bachelor's degree in French & TESOL – York St John University

September 2016 to April 2020

Languages

French, (CEFR C2)

English, (native language)