

SHUNYA KIMURA

Status: PhD Researcher at University College London
Field: Data Mining, Machine Learning, Spatial Statistics
Techs: R, Python, SQL, Apache Hive, Google BigQuery, Git

London, UK
+447399980267
kimushun618@gmail.com

EDUCATION

MPhil/PhD Social Data Science - University College London

Sep. 2021 - Exp Apr. 2025

- **Project:** *Typologies of Online Gambling Behaviour in Great Britain: Insights from Geographic Smart Data Analysis*; supervised by Professor Paul Longley and Dr Justin Van Dijk.
- **Funding:** Economic and Social Research Council (ESRC) Studentship; co-funded by GambleAware.
- **Data Mining:** In collaboration with our industry partner, I work with consumer datasets distributed across their data warehouse system. External data linkages are conducted while adhering to appropriate procedures for information governance. [Technologies: SQL, Apache Hive, Google Bigquery, R, GitLab, etc.]
- **Feature Engineering:** Designed novel behavioural markers to facilitate cross-customer comparisons, providing enhanced analytical insights into how users engage with the platform. [Libraries: arrow, datatable, tidyr, lubridate, sf, ggplot2, etc.]
- **Unsupervised Machine Learning:** Segmented customers into 12 distinct groups based on their revealed behaviour by applying dimensionality reduction techniques combined with clustering algorithms. [Libraries: arrow, datatable, tidyr, stats, factoextra, motifcluster, sf, ggplot2, etc.]

BA Geography with Quantitative Methods (1st Class) - University College London

Sep. 2018 - Aug. 2021

- **Relevant Modules:** Introduction to Quantitative Research Methods (73%), Geocomputation (68.5%), Data Analysis (74.5%), Mining Social and Geographic Datasets with Python (80.9%)
- **Independent Work:** Dissertation (74%) utilised mobility data to conduct an empirical investigation into the usage of the JUMP (now LIME) e-bike share system in London. Several spatial and statistical analyses were conducted, including DBSCAN, KDE clustering and multilevel negative binomial regression. This work was nominated for the 2021 Transport Geography Research Group Undergraduate Dissertation Prize.
- **Advanced Data Analysis:** Garnered experience with analysing randomised experiments, observational causal inference methods, quasi-experimental methods and with panel data methods.

CONFERENCES AND TALKS

Bristol Hub for Gambling Harms Research Joint Hybrid Seminar - Oral presentation

Mar. 2024

- *Unveiling Online Gambling Behaviour: A Descriptive Analysis of Consumer Data*. The recording and slides are available on [the github link](#).

GIS Research UK (GISRUK) Glasgow - Oral presentation

Apr. 2023

- The slides are available on [the github link](#).
- Kimura, S., Van Dijk, J.T., and Longley, P.A. *The Use of Consumer Data to Explore Geographic and Social Variations in Online Gambling*. Proceedings of the 31st Conference of GIS Research UK (GISRUK). Glasgow: University of Glasgow. doi: 10.5281/zenodo.7828456.

GIS Research UK (GISRUK) Glasgow - Co-author

Apr. 2023

- Winner of the Best Paper Award
- Todd, J., Kimura S., O'Brien, O., and Cheshire, J. *Dominant Trip Purposes within a Dockless Bicycle Sharing System in London*. Proceedings of the 31st Conference of GIS Research UK (GISRUK), Glasgow: University of Glasgow. doi: 10.5281/zenodo.7828473.

CDRC Data Partner Forum 2022 Oxford Saïd Business School - Oral presentation

Nov. 2022

- *A Classification of Online Gambling Behaviour in Great Britain*. The slides are available on [the github link](#).

- Winner of the CASA Prize for Best Paper in Spatial Analysis. The slides are available on [the github link](#).
- Kimura, S., Longley, P. A., and Van Dijk, J.T. *Exploring the spatial disparities in gambling risk and vulnerability*. Proceedings of the 30th Conference of GIS Research UK (GISRUK). Liverpool: University of Liverpool. doi: 10.5281/zenodo.6407146.

GRANTS

UBEL DTP International Institutional Visit Scheme - GBP 1,200

Dec. 2023 - Jan. 2024

- A travel grant for a 35-day visiting research period at the Institute of Disaster Mitigation for Urban Cultural Heritage, Ritsumeikan University at Kyoto, Japan. Invited by Professor Keiji Yano.

RECENT EXPERIENCES

Research Assistant - University College London

Jan. 2025 - Exp Mar. 2025

- **Safeguarding:** Leading the design and development of a new data product titled: *An Output Area Level Taxonomy of Online Gambling Behaviour in Great Britain* for the Geographic Data Service (GeoDS), with a focus on implementing rigorous disclosure control measures to protect commercially sensitive data.
- **Collaboration:** Partnering with an industry leader and a UCL researcher to deliver impactful insights.
- **Innovation:** Utilising Bayesian modelling to enhance the prediction, accuracy and applicability of the data product.

Senior Postgraduate Teaching Assistant - University College London

Jan. 2022 - Apr. 2023

- **Teaching:** Led seminars and held academic support and feedback hours for the following modules: (1) Geography in the Field I and (2) II; (3) Geocomputation; (4) Introduction to Quantitative Research Methods; and (5) Data Analysis

Postgraduate Teaching Assistant - Queen Mary University London

Nov. 2022 - Jan. 2023

- **Teaching:** Led seminars for a module: Introduction to Research Methods

Data Analyst Intern - BBC

Aug. 2021 - Sep. 2021

- Generated actionable business insights by utilising R and Python to conduct analysis on unstructured data sourced from web scraping and social media pages.
- Rapidly acquired and implemented new skills within a tight deadline.
- Presented findings to various teams, providing key benchmarks for assessing the success of BBCs social media platforms.

PROFESSIONAL DEVELOPMENT

Introduction to Bayesian Inference and Modelling - UCL Social Data Institute

Jul. 2023

- Participated in a four-day intensive course on Bayesian inference. The course covered the fundamentals of Bayesian statistics using RStudio and Stan. Taught by Dr Anwar Musah.
- Acquired skills in performing spatial risk prediction and quantifying uncertainty levels using exceedance probabilities for aerial data, which are currently employed in my PhD project.

Advanced Topics in Social & Geographic Data Sciences - UCL Geography

Jan. 2022 - Mar. 2022

- Audited a module in the Social & Geographic Data Science MSc course. Led by Dr Stephen Law.
- Revised topics on statistical learning, including Deep Learning, Convolutional Neural Networks, and GeoAI.

LANGUAGE PROFICIENCY

- **Japanese:** Native proficiency; **English:** Full professional proficiency; and **Arabic:** Beginner proficiency.