BAICHEN DU

+852-56964800 | baichen@connect.hku.hk | https://baichen-du.github.io/ | Google Scholar

EXECUTIVE SUMMARY

Experienced Data Analyst with over 4 years of expertise in standardizing, analyzing, and interpreting complex data to drive informed decisions. Proficient in Python and R, with solid skills in database management in SQL. Skilled in using data visualization tools and performing advanced statistical analysis to extract actionable insights from large datasets. Strong communicator with the ability to translate complex data into clear, strategic insights for both technical and non-technical stakeholders.

EDUCATION

University of California, Los Angeles, *PhD in Linguistics*　　　　　**Sept 2025 - Present**
Expertise: Computational Linguistics and Language Processing

The University of Hong Kong, *Bachelor of Social Sciences (First Class Honor)*　　**Sept 2019 - Jul 2024**
Majors: Politics and Public Administration & Linguistics; CGPA: 3.94/4.30
Awards: Dean's Honor List (2019-22), Soong Ching-ling Foundation Scholarship, Globalink
　　　　Research Scholarship, K.K.Ho Prize, Ms Cecilia Wong Prize, Transdisciplinary
　　　　Research Fund, Leung Kau Kui Research and Teaching Endowment

University of California, Berkeley, *Exchange Program*　　　　**Jan 2023 - Dec 2023**
Awards: Acoustical Society of America Student Fund, UC Berkeley Departmental Fund

PROFESSIONAL AND PROJECT EXPERIENCE

Teaching Assistant, *Part-time Placement*　　　　**The University of Hong Kong**
Department of Politics and Public Administration　　　　**Sept 2024 - Dec 2024**

- Compiled Python scripts for traditional NLP models, such as topic modeling, decision trees, etc.
- Implemented various Large Language Models (e.g., Llama) and Machine Learning for text data analysis
- Supported interdisciplinary application of computational social science techniques

Data Analyst, *Half-time Research Internship*　　　　**University of California, Berkeley**
PI: Prof. Keith Johnson, Department of Linguistics　　　　**Feb 2023 - Dec 2023**

- Led analysis and management team of 2 for large scale database of audio and visual data
- Responsible for the entire statistical modelling and documentation using GAMM, LMER, SEM, etc.
- Implemented Computer Vision models for automated processing of multimodal data

Project Leader　　　　**The University of Hong Kong**
Department of Politics and Public Administration　　　　**Sept 2021 - Dec 2021**

- Analysis of the relationship between China's direct aids/investments overseas and diplomatic closeness
- Scraped, managed, analyzed, and modeled HTML database
- Tracked global aid projects and coordinated project documentation and replicability.

Data Analyst and Programmer, *Part-time Internship*　　　　**The University of Hong Kong**
PI: Prof. Youngah Do, Department of Linguistics　　　　**Sept 2021 - Sept 2022**

- Developed web-based experiments with user-friendly GUI to pre-schoolers
- Participated in automated standardization, visualization, and interpretation of behavioral data
- Communicated findings of cognitive and statistical learning in leading journals.

SKILLS

Languages	Mandarin (Native), English (Proficient), Cantonese (Professional)
Skills	Data Analytics and Database Management: Python, R, SQL, Visualization: ggplot, Seaborn, PowerBI, Tableau; MS Office, Praat, LaTeX, Markdown