CAITLIN LOFTUS

Email: cloftus@uchicago.edu
Github: https://github.com/ceilingloft

EXPERIENCE

Research Engineer, Mansueto Institute for Urban Innovation — Jan 2021 - present, Chicago, IL

Responsible for building a robust ETL pipeline to download, process and analyse Census Household Pulse Survey data as part of the Urban Research Corps' housing project. Currently leading work to automate our workflow using Google Cloud Platform (GCP) products such as Cloud Scheduler, Pub/Sub, and Compute Engine.

Graduate Research Assistant, Mansueto Institute for Urban Innovation — May 2020 - Sep 2020, Chicago, IL

Worked for Dr Luis Bettencourt on the Adaptive Control for COVID-19 project.

- Built a robust ETL pipeline to clean and standardise the messy, crowdsourced data, and made improvements to the codebase that increased its efficiency.
- Researched and implemented econometrics model to explore impact of lockdown policies on COVID-19 Rt values in the US.

Data Scientist, Springer Nature — Apr 2017 - Aug 2019, London, UK

- Instrumental in creating a new event-based data pipeline to record user events across Springer Nature's sites, using Google Cloud Platform (BigQuery, DataFlow, DataStudio).
- Ran analyses of user interaction data (using Python and BigQuery SQL) to help the department make informed decisions about which improvements would best benefit users.
- Transformed A/B product development testing & data analysis practices, which increased end-user engagement and improved data literacy across the department.

EDUCATION

Master of Science in Computational Analysis and Public Policy — Sep 2019 - June 2021 (expected), Chicago, IL

The University of Chicago, Harris School of Public Policy & the Department of Computer Science

Coursework and projects in: Python programming; database modelling and design (PostgreSQL, POST/GET requests, JSON); big data application architecture (Hive, Hbase, Spark); data visualisation (python packages such as Altair & Seaborn; D3); machine learning (scikit-learn, PyTorch, spaCy); regression modeling (R).

Master of Science with Distinction in Genes, Environment and Development — Sep 2015 - Sep 2016, London, UK

King's College London, Social, Genetic & Developmental Psychiatry Centre

An interdisciplinary programme in behavioural genetics, focusing on how genes and environment shape the development of human behaviour and psychiatric disorders.

Independent research project: used structural equation modelling in R and OpenMx to explore potential interactions between genotype and family environment in increasing risk to adolescent antisocial behaviour (supervisor: Dr Tom McAdams).

Bachelor of Science (hons) in Psychology — Sep 2019 - Jun 2021, London, UK
University College London, Department of Psychology
Independent research project: impact of the use of Compassionate Imagery Training as an intervention for highly self-critical individuals (supervisor: Dr Caroline Falconer).

ADDITIONAL EXPERIENCE

Product Analyst, The Thinking Partnership — Feb 2015 - Mar 2017, Oxford, UK
Worked extensively on the development and implementation of CharacterScope— a leadership development software package—helping it move from initial design concept to a fully launched product. Ran statistical modeling analyses to underpin the development of CharacterScope and improve the user experience, which increased user engagement.

Volunteer mentor, **ReachOut Club, ReachOut** — Sep 2018 - Jun 2019
Worked as a mentor with children in disadvantaged communities to help them grow in confidence and educational attainment.

Volunteer instructor, Python Beginners Course, CodeFirst:Girls — Sep - Dec 2018
Taught a group of eight students a beginners Python course.

TECHNICAL SKILLS

Programming languages	Proficient: Python, SQL, R. Familiar with Java, Stata
Database technologies	PostgreSQL; Apache Big Data software (Hive, Hbase, Spark)
Cloud computing	Google Cloud Platform products, including Cloud Scheduler, Pub/Sub, BigQuery, Compute Engine
Statistical and machine learning methods	- Writing programs in Python to simulate and model data (such as simulating a model of housing segregation) - Building and running machine learning models using scikit learn - Regression modelling in R and Python - Familiar with building natural language based models in PyTorch and spaCy
Web scraping	Using Selenium and Beautiful Soup in Python to scrape web data
Interactive data visualisation	Experience building web page (using NodeJS, HTML, CSS) to host interactive data visualisation created in D3