Cameron D. Appel

+44 7 435 00 3495 | cam@appel.co.uk | www.appel.co.uk

Summary

I am a data scientist with 7,000+ academic citations and 5+ years of experience in data analysis and visualization, specializing in the healthcare and public health sectors to develop actionable insights and improve health outcomes

Experience

Our World in Data

Data Scientist   June 2019 – July 2022

- Early hire at a Y Combinator startup - in collaboration with Oxford University - contributing to the platform's growth to nearly 100 million visitors and 300 million pageviews annually, with content referenced in nearly 50,000 media articles, including major outlets like The Guardian, BBC, and The New York Times
- Collaborated with senior public health officials around the world to provide accurate and timely data for decision-making during the COVID-19 pandemic. My work was demonstrably fundamental to global efforts to prevent the spread of the pandemic; it was frequently referenced by world leaders in televised interviews, including the sitting US President
- Built data pipelines to collect, transform and visualize 250K+ data points on COVID-19 statistics in real-time, using Python & R libraries such as BeautifulSoup, rvest, and Selenium for web scraping and API integration, building skills in data engineering and feature extraction
- Resolved data quality issues and implemented user feedback using GitHub and Twitter

Environmental Data Science Book

Data Scientist   July 2023 – Present

- Collaborated with researchers from the Alan Turing Institute and the University of Florida on an open-source environmental data science project, utilizing GitHub, Hugging Face and Binder to improve reproducibility and transparency
- Trained a PyTorch CNN model on labelled livestock images, improving box recall from 0.4405 to 0.9535, precision from 0.5826 to 0.8587, and mean IoU from 0.3135 to 0.6571

Health Equity Evidence Centre

Data Scientist   Oct 2023 – Present

- Provided advanced data analysis for NHS primary care, identifying structural healthcare inequalities and driving actionable insights to improve equitable service delivery across a system with an annual operational expenditure of ~$22bn and 60 million patients
- Developed data models that uncovered disparities in healthcare access and outcomes, enabling NHS executives to allocate resources more effectively and optimize care for underserved populations, influencing strategic initiatives to reduce health inequality

Chartbook of Economic Inequality

Data Scientist   June 2021 – July 2022

- Conducted comprehensive research on long-term economic inequality, collecting data on five key inequality measures across 25 countries over a century, contributing to the development of robust, cross-country datasets
- Developed detailed data documentation and created interactive visualizations using D3.js, improving user engagement and understanding of complex inequality data while contributing to blog posts and metadata for accessible, data-driven insights

Education

Imperial College   *London, UK*

Master of Science, Health Data Analytics and Machine   *Sept. 2023*
Learning

- Advanced quantitative methods (statistical modelling, machine learning) to analyze medical and public health data

- Research project on machine learning models (random forest, convolutional neural network) to classify disease status using large cough audio dataset

Columbia University   *New York, NY*
Bachelor of Arts, Economics   *Sept. 2020*
- Highly quantitative coursework focusing on econometrics, statistical methods and financial economics
- Dissertation seminar on ESG (Environmental, Social, and Governance) finance and investing

Skills

Python, R, SQL, JavaScript, HTML/CSS, Pandas, NumPy, Matplotlib, Seaborn, D3.js, Scikit-learn, TensorFlow, PyTorch, Random Forests, CNNs, XGBoost, Logistic Regression, Bayesian Statistics, Time Series Analysis, Web Scraping (BeautifulSoup, rvest, Selenium), Feature Engineering, Azure, Git, GitHub, RESTful APIs, Jupyter Notebooks, Google Colab, Slack, Academic Writing, Policy Reporting, Open Data Practices