Logan Cooper

Portfolio: logan-cooper.com
Github: github.com/ldtcooper
LinkedIn: linkedin.com/in/ldtcooper/

Email: LDTCoop@gmail.com
Phone:(323) 698-6487

SKILLS

- **Languages**:   Python, SQL, JavaScript
- **Frameworks**:   Pandas, scikit-learn, Statsmodels, PyTorch, Flask, BeautifulSoup, D3.js, Node, Puppeteer.js
- **Tools**:   Git, PostgreSQL, MongoDB, GLPK

EXPERIENCE

- **Duke University** — Durham, NC
  Research Assistant — *Feb 2022 – May 2022, Sep 2022 – Present*
  - Automated the scraping and cleaning of several million broadcast news transcripts with BS4 and Puppeteer to create a dataset for NLP-based research by professors Pat Bayer and Jonathan Moreno-Medina.
  - Prototyped a data pipeline to run BERT-based keyword extraction, event extraction, and sentiment analysis algorithms on 80GB of unstructured text data.

- **Pearson** — Durham, NC
  Intern: Data Engineering - Predictive Analytics — *May 2022 - August 2022*
  - Demoed a proof-of-concept book search and recommendation engine for Pearson's digital product catalog to CTO and senior VPs as part of a product pitch.
  - Developed a data pipeline and rank-based recommendation system for 11.5 million documents from 1,200 e-books and online texts with BigQuery, Pandas, FastText, BM25, and Flask.

- **Federal Reserve Bank of San Francisco** — San Francisco, CA
  Software Developer — *Jul 2020 - Jul 2021*
  - Created Java ETL tool to transform CSV bank data from the New York Fed into a format our data platform could ingest, allowing economists in the entire Fed system to access the data.

- **Forio** — San Francisco, CA
  Software Developer — *Jan 2018 - Jun 2020*
  - Developed React/Redux frontends for 9 educational simulations collectively worth over $990,000 in revenue, including data visualisations and grade dashboards.
  - Saved 8-16 hours of developer time per project while improving adherence to internal coding standards by creating a Node CLI tool to auto-generate project and component boilerplate.

EDUCATION

- **Duke University** — Durham, NC
  M.S. Economics and Computer Science — *August 2021 - May 2023 (Expected)*
  Selected Coursework: Urban Economics, Computational Microeconomics, Machine Learning, Natural Language Processing, Databases

- **U.C. Santa Cruz** — Santa Cruz, CA
  B.A. Global Economics (cum laude) — *September 2012 - December 2016*

PROJECTS

- **PollyDarton – NLP-driven Data Curation for Polymer Nanocomposite Data Resources**: NLP-based classification of sentences in published materials science papers. Tech: Python, PyTorch, scikit-learn, NLTK, Gensim, RoBERTa

- **Obesity and Corruption in the Former USSR**: In-progress paper co-authored with Prof. Charlie Becker rigorously examining the relationship between politicians' algorithmically-estimated body mass index and corruption indicators posited by Blavatskyy in his paper "Obesity of politicians and corruption in post-Soviet countries". Uses multivariate linear regression trained on health survey data to control for confounding factors. Tech: Python, Statsmodels (Ongoing)

- **The Vocabulary of Buck 65, Quantified**: Blog post with analysis of rapper Buck 65's vocabulary inspired by Matt Daniels' Raper Vocab Chart. Tech: Python, Pandas. (May 2021)

- **Linear Regression API**: API which can perform and visualize simple OLS linear regressions. Tech: Python, Flask, JavaScript, D3.js, AWS Lambda (Aug 2020)

VOLUNTEER EXPERIENCE

- **Code for San Francisco** — San Francisco, CA
  Volunteer — *Mar 2020 - Nov 2020*
  - Built and maintained data scrapers to automate gathering COVID case data from Bay Area public health websites.