

# Logan Cooper

Portfolio: [logan-cooper.com](http://logan-cooper.com)

Github: [github.com/ldtcooper](https://github.com/ldtcooper)

LinkedIn: [linkedin.com/in/ldtcooper/](https://linkedin.com/in/ldtcooper/)

Email: [LDTCoop@gmail.com](mailto:LDTCoop@gmail.com)

Phone: (323) 698-6487

## SKILLS

---

- **Languages:** Python, SQL, JavaScript
- **Frameworks:** Pandas, scikit-learn, numpy, scipy, PyTorch, spacy, Flask, BeautifulSoup, D3.js, Node, Puppeteer.js
- **Tools:** Git, PostgreSQL, MongoDB, Amazon Web Services (AWS), Google Cloud Platform (GCP), Excel

## EDUCATION

---

### Duke University

Durham, NC

- *M.S. Economics and Computer Science*

August 2021 - May 2023 (Expected)

**Selected Coursework:** Urban Economics, Machine Learning, Natural Language Processing, Databases, Topics in Data Science

### U.C. Santa Cruz

Santa Cruz, CA

- *B.A. Global Economics (cum laude)*

September 2012 - December 2016

## EXPERIENCE

---

### Duke University

Durham, NC

- *Graduate Research Assistant*

Feb 2022 - May 2022, Sep 2022 - Present

- Created a novel research dataset by computationally scraping and cleaning 80GB of news broadcast transcripts with BS4 and Puppeteer.js.
- Identified subsets of data relevant to specific research questions with BERT-based keyword extraction and topic modelling algorithms.

### Pearson

Durham, NC

- *Data Scientist Intern*

May 2022 - August 2022

- Developed Python data pipelines to download, clean, and transcribe ~1.5 million text, audio, and video documents from internal APIs and store them in Google BigQuery.
- Prototyped a rank-based recommendation system to match learner queries to paragraphs and video from 1,200 e-textbooks using Pandas, FastText, Gensim, 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. Resulted in access time for this data dropping from hours to minutes.

### Forio

San Francisco, CA

- *Software Developer*

Jan 2018 - Jun 2020

- Built React/Redux frontends for 9 educational simulations collectively worth over \$990,000 in revenue, including data visualizations 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.

## PROJECTS

---

- **PollyDarton – NLP-driven Data Curation for Polymer Nanocomposite Data Resources:** Group project. NLP-based classification of sentences in published materials science papers. Improved F1 score over current models by 0.05 points. Tech: Python, PyTorch, scikit-learn, NLTK, Gensim, RoBERTa
- **Obesity and Corruption in the Former USSR:** Working research paper with Prof. Charlie Becker which corrects for omitted variables in a 2021 paper which uses machine learning estimates of politicians' BMIs as a proxy variable for corruption. Finds that upon correcting for age and sex, the corruption-BMI relationship is weaker but still significant between the 5% and 10% levels. Tech: Python, pandas, Statsmodels
- **Crypto Currency Arbitrage DB:** Group project. Arbitrage tool using Python data scrapers scheduled hourly feeding data from multiple crypto exchanges to a PostgreSQL database and visualizing it with React. Tech: Python, Flask, PostgreSQL, React
- **The Vocabulary of Buck 65, Quantified:** Blog post with analysis of rapper Buck 65's vocabulary inspired by Matt Daniels' Raper Vocab Chart. Finds that Buck 65 should be the third highest vocabulary in hip-hop. Tech: Python, Pandas.
- **Linear Regression API:** API which can perform and visualize simple OLS linear regressions. Tech: Python, Flask, JavaScript, D3.js, AWS Lambda