

COREY DUBIN

dubincorey@gmail.com • 201-410-9093 • github.com/dubincorey • dubincorey.github.io

EXPERIENCE

Caris Life Sciences

New York, NY

Software Engineer

Dec 2022 – Present

- Scale efforts using Python to automate deployment, transforming data into digestible formats
- Initiate and develop tool to validate 325,000+ quarterly file deliveries, flagging 100+ for correction
- Upload 10+ clinical and molecular data studies to AWS hosted platform for visualization and analysis
- Write 20+ extraction scripts for translation and analysis across 2,000,000+ files within ETL pipeline
- Contribute to the open-source genomics platform cBioPortal accessed by 2,000+ users per day to improve usability

Pinata

Remote

Software Development Engineer in Test (SDET)

Jan 2021 – Dec 2022

- Overhauled Django models for property technology startup to improve non-engineering team accessibility
- Reviewed backend unit tests and pull requests for accuracy in documentation and style guidelines
- Collaborated with product development team to ensure consistent project execution and best practices
- Worked with automation team to develop automated test suites employing Selenium Web Driver

AliveAndKickn

Remote

Data Engineer Intern

Jul 2020 – Aug 2020

- Utilized Python to analyze over 120,000+ survey responses for Lynch Syndrome cancer advocacy group
- Visualized survey responses with Pandas and Matplotlib to discover inconsistencies in patient screening
- Published original abstract to improve guidelines regarding cancer screenings and genetic testing

PROJECTS

Home Loan Predictor

May 2022

- Engineered Scikit-Learn algorithms in tandem with GitHub Copilot examining loan eligibility data
- Cleaned, validated, and stored data in Pandas and predicted results via Linear Regression
- Determined 31% decrease in loan eligibility between machine learning algorithm and observed statistics

Whole Foods Web Scraper

Mar 2022 – Apr 2022

- Extrapolated sale prices at Whole Foods through BeautifulSoup and Regular Expressions
- Organized data and estimated value of Prime membership using Pandas

Fitegy

Oct 2021 – Dec 2021

- Led team of 5 students to develop social fitness platform employing MERN stack
- Coded user registration and authentication to be stored in MongoDB
- Built continuous integration workflow through Travis CI and deployed site to DigitalOcean

EDUCATION

New York University

New York, NY

Bachelor of Arts

Aug 2018 – May 2022

- **Individualized Major:** Computer Science | **Minor:** Web Development and Programming | **GPA:** 3.53/4.00

SKILLS

Languages: Python, TypeScript, SQL

Frameworks: Pandas, Node, React

Tools/Processes: AWS, Jupyter, Git

PUBLICATIONS

- "Findings from the HEROIC Registry," *Journal of Clinical Oncology*, October 2020