Natalie Lubich

nlubich@ucsb.edu | github.com/nlubich

Education

University of California, Santa Barbara — Santa Barbara, CA
Bachelor of Statistics and Data Science — *June 2025*

- **GPA:** 3.7
- **Relevant Coursework:**
  Intermediate Python, Probability and Statistics, Mathematical Staistics, Differential Equations, Regression Analysis, Machine Learning, Time Series, Principles of Data Science with R, Linear Algebra, Calculus II, Calculus III, Discrete Math, DS Concepts and Analysis, SAS Base Programming

Technical Skills

Languages	Python, R, SQL, Java, C Swift, SAS, JavaScript, HTML, CSS, Bash
Libraries	Requests, Tableau, Pandas, Numpy, BeautifulSoup, Selenium, Bootstrap 5
Frameworks	Flask, Spring Boot, PyTest, JUnit
Tools	Git, Jupyter, Github Actions, Xcode, MongoDB, SQLite, Figma
Platforms	AWS, SQL Server, Splunk, Github, BitBucket, AWS, Linux, Windows, macOS
Skills	Datamining, Data Analysis, REST APIs, Software Development, QA, Test Automation, Regression Testing, DevOps, CI/CD, Agile, Clean Code, Problem Solving, Teaching, Leadership, Collaboration, Documentation

Work Experience

Data Analyst Intern
LeanTaaS — *June 2024 - Aug 2024*

- Contributed to a growth-stage company transforming hospital and infusion center operations using lean principles, predictive and prescriptive analytics, and machine learning
- Managed and communicated complex data integration processes, including predictive and prescriptive analytics workflows, with clients in the healthcare sector, communicating technical concepts to non-technical roles within the company
- Refactored, documented, and unit tested over 4,500 lines of legacy Python ELT preprocessing code using Git version control to manage code changes, enhancing data management, logic, and readability for sensitive hospital patient data, significantly benefiting 3 key customers
- Collaborated with the Marketing team to analyze 2023 and 2024 campaign performance, focusing on conversion timelines across channels, products, and partners, using Python to create visualizations that uncovered insights and optimized strategies to meet organizational KPIs

AI Research Intern
Keywords Studios — *Jun 2023 - Mar 2024*

- Focused on and conducted research in providing AI-powered natural language processing solutions
- Orchestrated the production of an extensive dataset comprising of over 3,000 recordings, essential for training and fine-tuning AI
- Conducted rigorous Quality Assurance (QA) processes by meticulously reviewing recordings and tests, ensuring the precision and reliability of AI
- Thoroughly documented research findings and development processes to facilitate knowledge transfer
- Thrived in a fast-paced and fully-remote work environment, adapting to rapidly-changing project requirements
- Effectively communicated and collaborated with multidisciplinary team members, sharing innovative ideas, answering questions, and promptly addressing project-related issues to support team success

Data Processing Intern
SF Pathways to Citizenship — *Jan 2020 - May 2020*

- Boosted runtime of automation tests of core survey platform by 200%
- Manipulated user data in SQL database
- Assisted others with filling out citizenship forms using Excel Spreadsheets, by utilizing macro formulas
- Taught QA principles to fellow interns, allowing for more automation to be developed
- Generated a statistics report using Pandas and made an interactive dashboard in Tableau

STEM Subject Tutor
The Village Project SF — *Jan 2019 - Mar 2020*

- Empowered students from SF marginalized communities to tackle STEM subjects after school
- Mentored the class in subjects including Math and Computer Science, fostering technical skills, providing study strategies, and supporting their overall learning journey

Projects

GradeBook | *Grade Manager w/ Auto Grouping and Stats*

- Automatically calculates statistics and categorizes course data by coursename, subject etc...
- Allows user to add/update/delete entries in database
- Command line interface that allows the users to interact with the system
- Persistent storage using serialization
- Used PyTest to develop a comprehensive suit of tests for the program

MazeSolver | *Generates and solves (1000x1000)+ mazes*

- Used recursive algorithms and data structures like stacks to solve mazes correctly
- Data structure optimization allowed me to solve mazes as large as 7400x7400 quickly!
- Allows user to store grades and grade with various difficulty settings
- Implemented pre-order traversal to label nodes as discovered using BFS
- Used PyTest to develop a comprehensive suit of tests for the program

Volunteering

Tutoring STEM Tutor

- **Mentored Students** Teaching Data Science principles and Structures and Algorithms
- **Helping empower women** Taught women and people of color STEM and increasing diversity in tech
- **Boosting Leadership Skills** Helping students gain valuable leadership and public speaking skills