Cayman Simpson

Address: 625 Sky Hy Circle, Lafayette, CA 94549 | (925) 818-3550 | cayman@stanford.edu

Objective: To work with passionate people on world-changing problems. Looking for full-time in Data Science or SWE.
LinkedIn: http://www.linkedin.com/in/caymansimpson

EDUCATION

Stanford University, Stanford, CA (GPA: 3.54) September 2012 – present
Expected BS in Mathematical and Computational Science: June 2016
Expected Minor in Sociology: June 2016

COMPUTER SKILLS

Languages: Proficient in **Java** and **Object-Oriented Programming**. Designed and developed a database, OOP structure and corresponding API for a quiz social media website, in addition to having developped a multi-threaded WebCrawler. In spare time, implemented network analysis algorithms, pythonic dictionaries and arrays (supporting Lambda functions) and a music player.

Proficient in **Python**. Specifically ML and structuring data. Wrote scripts to create, maintain and export databases. Have done many text, network and predictive analyses in Python, too.

Proficient in **Tableau**. Both academic and industry experience. Became Tableau point for my industry teams. Have worked closely with Tableau Product Teams to detail more customized visualizations.

Experience in **HIVE/SQL**. Worked with trillion-row tables in Hive to extract meaning from data. Pioneered new tables and pipelines at Facebook using many custon UDFs, UDAFs and UDTFs.

Experience in **Javascript**. Designed and developed a dashboard ready for deployment using primarily Javascript, in addition to designing, creating and implementing my own custom heatmap and tooltip libraries. Personal segregation visualization project also written primarily in Javascript and D3.

Experience in **Map-Reduce**. Extended an API with a function calculating the homogeneity of an individual or place given geotemporal data. Also designed an algorithm that optimized the preprocessing of geotemporal data through data compression and I/O reduction.

Familiar with **R** and **Matlab**. Implemented Pagerank, MaxFlow and Stochastic Gradient Descent in Matlab. Analyzed Facebook user behaviors and San Francisco's Housing Industry in R.

WORK EXPERIENCE

Facebook *Data Science Intern* Stanford, CA June 2015 – September 2015
Performed Product Analytics on Facebook's Ecosystem to analyze user's interactions with certain Facebook phenomena, using primarily Hive, R and Tableau. Additionally performed NLP/NLU analyses on cohorts of users under the Care section of Facebook to best understand how to cater and modify Facebook to meet their needs. This effort was due to a large push to make Facebook a safe space for *all* its users.

Stanford Engineering *Computer Science Tutor* Stanford, CA November 2014 – Present
Official CS tutor associated with Stanford Engineering. Specifically helped underrepresented demographics in CS. I teach:
CS106A: Programming Methodologies CS106B: Programming Abstractions
CS107: Computer Organization and Systems CS108: Object-Oriented Systems Design

DataSpark *Data Analyst* Singapore, Singapore June 2014 – September 2014
Worked with Javascript and Python to build targeted event-based visualizations and dashboards for our geo-analytic products from ground-up. The dashboard was created to be dynamic for easy future use. Used programs such as Tableau for quick visualizations and libraries such as D3 and Leaflet for custom, in-depth visualizations. Also analyzed a month of

call data with 5M+ nodes and 26M+ edges using various scalable algorithms in order to best figure out how Dataspark should target a population to maximize market penetration and catalyze product growth.

-Received the SPOT award, given to the staff member who makes the most significant contribution to the team in that quarter. First and only intern to receive this award.

-Dashboard was sold to the Singaporean government for $80K as an analytical product. Also was used for POC's by multiple departments.

SingTel *Research And Development Intern* Singapore, Singapore June 2013 – September 2013

Worked with many frameworks to profile and understand behavioral patterns of different consumers by analyzing their location patterns. Wrote visualization programs and applications using languages such as Javascript/JQuery, PHP, Java and C++ to both gain better insight into and increase the monetary value of the location data. This work contributed to my team (R&D) branching off of Singtel and becoming DataSpark.

PROJECTS

- **Taking the Critical Reading Section of the SAT**: Worked on a team using principles of both traditional NLP and more experimental NLU (Natural Language Understanding) to take a portion of the SAT, primarily focusing on word disambiguations. Using a combination of techniques including Shallow Neural Networks and VSM's, we scored 500.
- **US Segregation Visualization**: A visualization (in development) using census data over the last 10 years to extremely efficiently highlight segregation in the US based on race, income and educational attainment. Includes interactivity with responsive charts and dynamic SVG Elements. Done primarily in SVG/Javascript/D3 and CSS with Python/Java for data-processing on the backend.
- **Fractal Analysis and Validation of Networks**: Studied fractal nature of nontrivial networks ranging from road networks to power-grid networks to social media networks. Assessed mathematical and biological multi-fractal generating models using an original validating measure (fractal dimension over reductions). Found both unexpected fractality in networks and that no current model can universally and sufficiently model fractality; our validating measure was sufficient at illuminating that fact. (Link: http://web.stanford.edu/class/cs224w/projects/cs224w-70-final.pdf)
- **Dynamic Dashboard**: A generalized dashboard that visualizes geotemporal data using Leaflet and D3 libraries. Implements my original GeoJSON heatmap library and works on a "plug and play" basis with any tabular data.
- **WebSpider**: Multi-threaded exhaustive web crawler written in Java with GUI. Recreates website servers locally.
- **Music Maker**: Developped a Java program from scratch that accepts a String and plays its song.

RELEVANT COURSEWORK

CS103: Mathematical Foundations of Computing
CS106A: Programming Methodologies
CS106B: Programming Abstractions
CS107: Computer Organization and Systems
CS108: Object-Oriented Systems Design
CS124: From Languages to Information
CS 221: Principles in AI*
CS224W: Social and Information Networks
CS224U: Natural Language Understanding
CS246: Mining Massive Datasets
CS246H: Mining Massive Datasets Hadoop Labs
STATS116: Theory of Probability
STATS200: Statistical Inference
MATH51: Linear Algebra and Diff. Calculus
MATH52: Integral Calculus
MATH53: Differential Equations
MATH104: Applied Matrix Theory
CME193: Scientific Computing in Python
MS&E177: Creativity Rules (Design School)*
MS&E211: Linear and Non-Linear Optimization
MS&E221: Stochastic Processes
ENGR350: Stanford Data Lab: Social Impact

*currently enrolled

ADDITIONAL INFORMATION

- I enjoy playing basketball and tenor sax, eating, traveling, and I am a humanitarian, an avid Warriors fan and member of Phi Kappa Psi.
- I've backpacked through Europe, Southeast Asia twice, Central Mexico, Costa Rica, Ecuador and the Amazon.