William Ma

Ph.D. Student · Research Assistant

wma1017@gmail.com | williamma12 | williamma12 | William Ma

Education

University of Chicago — *Chicago, IL*

Ph. D in Computer Science — *Sept. 2020 - Present*

- Focus: Database

University of California, Berkeley — *Berkeley, CA*

B.A. in Computer Science w/ High Distinction in General Scholarship — *Aug. 2016 - May 2020*

- Dean's List: Fall 2017, Spring 2018
- GPA: 3.83

Publications

- Devin Petersohn, Stephen Macke, Doris Xin, **William Ma**, Doris Lee, Xiangxi Mo, Joseph E. Gonzalez, Joseph M. Hellerstein, Anthony D. Joseph, Aditya Parameswaran, "Towards Scalable Dataframe Systems", *VLDB. August 2020*
- **William Ma**, "Serverless Query Processing on a Budget", *ACM SIGMOD. June 2020*

Presentations

RISE Lab SURF Poster Session — *Berkeley, CA*

The Price is Right: Data Analytics on a Budget — *August 2019*

- Poster presentation showing how we are able to reduce wall-clock time or monetary cost, constraining the other using an ideal serverless platform in comparison to regular serverful platforms.

Cal Day Statistics Undergraduate Research Poster Session — *Berkeley, CA*

Conjoint Analysis — *May 2019*

- Poster presentation showing that Conjoint Analysis, a common survey technique, is "brittle."

Research and Industry Experience

University of California, Berkeley - RISE Lab — *Berkeley, CA*

Research Assistant — *Aug. 2018 - May 2020*

- Mentored by Professors Aditya Parameswaran, Joey Gonzalez, and Ion Stoica and graduate students Devin Petersohn and Robert Nishihara.
- Implemented a gossiping routine for inter-node object transfers, which would reduce broadcast operations to $O(n)$ from $O(n^2)$ in Ray (github.com/ray-project/ray), a parallel task-execution engine.
- Redesigned the backend of Modin (github.com/modin-project/modin) to allow for greater code reuse and various other optimizations (e.g., shuffle sort, broadcast shuffle).
- Demonstrated that serverless can have better performance and allow for greater control of monetary cost and wall-clock time than serverful systems (Presented findings at RISE Lab SURF poster session).
- Helped set future research goals for the Modin project and resulted in the Modin vision paper.
- Developed a rule- and cost-based model for optimal serverless cluster configuration with fixed monetary or wall-clock time budgets, using an idealized version of Spark.

University of California, Berkeley - Statistics — *Berkeley, CA*

Research Assistant — *Jan. 2018 - Present*

- Mentored by Professor Philip Stark.
- Demonstrated that typical applications of Conjoint Analysis violated the underlying assumptions, which leads to erroneous conclusions and biased estimates.
- Implemented Hierarchical Bayesian choice-based Conjoint Analysis, a common survey technique, and a survey simulator.

University of California, Berkeley - Art History — *Berkeley, CA*

Research Assistant — *Jun. 2017 - Dec. 2019*

- Mentored by Professor Diliana Angelova.
- Analyzed Roman Imperial Coinage to show how the trends of certain characteristics (e.g., "divus", "helmet", "radiate") correlate to specific times in Roman history (e.g., 3rd century crisis, rule of Constantine)
- Summarized findings in interactive plots available on github (https://williamma12.github.io/roman_coinage/).

Starbutter AI — *Berkeley, CA*

Software Engineer Intern — *Sep. 2017 - Jan. 2018*

- Built analytics platform to help measure user activity and performance metrics of company chatbots.
- Lead weekly analytics meetings to support data-based decisions.

Lawrence Berkeley National Laboratory — *Berkeley, CA*

Research Intern — *Aug. 2016 - Sep. 2017*

- Mentored by Shirley Ho.
- Used the cosmological redshift from the Sloan Digital Sky Survey to calculate Baryonic Acoustic Oscillation with the two-point correlation function.

Stanford University - Stanford Solar Center — *Palo Alto, CA*

Research Intern — *Jan. 2016 - Aug. 2016*

- Mentored by Professor Philip Scherrer.
- Analyzed solar magnetogram data to look for the twenty-four known solar cycles and demonstrate the existence of the 25th solar cycle.

Honors & Awards

2020 **Crerar Fellowship,** University of Chicago — *Chicago, IL*

2017, 2018 **Dean's List,** University of California, Berkeley — *Berkeley, CA*

Service

Students Mentored

Anthony Tong *(BA CS 2021)* — 2019

Japjot Singh *(BA Data Science 2021)* — 2019

Manan Khattar *(BA CS/Math 2021)* — 2019

Berkeley High School - BRIDGE Program — *Berkeley, CA*

Mentor — *August 2019 - May 2020*

- Mentored underprivileged students about college and tutored them in various topics including algebra and English.

Skills

DevOps	AWS, TravisCI
Technologies	Spark, vim, tmux, git
Programming	Python, SQL, Scala, C, C++, Java, R, MatLab, LaTeX