

## Nathan Chan

nathanjchan2@gmail.com – 650-430-3780 – www.nathanjchan.com – www.linkedin.com/in/nathanjchan

---

### EDUCATION

**Bachelor of Science, Statistics** (Statistical Data Science), *University of California, Davis* Jun 2020 (expected)  
Minor in Political Science, Dean's Honor List, Judicial Council Member (student government), GPA 3.7/4.0

### SKILLS

**Languages:** Python, R, SQL

**Other:** Excel, Stata, SPSS, C++, Unix, HTML, CSS

**Data Science:** Jupyter notebooks, visualizations (ggplot2, seaborn, pandas, numpy), web scraping, machine learning, big data processing, data structures, data cleaning

### EXPERIENCE

**Data Analyst**, *Hibel Lab and HERD Lab, Center for Mind and Brain, UC Davis* May 2019 - Present  
Write R scripts to clean hormone, heartbeat, and survey data for social science researchers. Built Linux pipeline to format brain imaging data. Redesign website for California Families Project ([www.californiafamiliesproject.org](http://www.californiafamiliesproject.org)).

**Course Assistant**, *Statistics 103 and 106, UC Davis* Apr 2019 - Present  
Grade and record homework from 150 students in upper-division statistics courses.

**Data Assistant**, *Alumni Relations, UC Davis* Jan 2019 - Oct 2019  
Adjust and add entity information in database of 1+ million people. Verify donation and gift amounts and upload receipts to database. Sort through and shred confidential financial information.

**Research Assistant**, *Asian American Center on Disparities Research, Psychology, UC Davis* Sep 2017 - Jun 2018  
Designed, wrote, presented research poster at Undergraduate Research Conference at UC Davis. Performed data analysis with sample size of 425 students (ANOVAs, multiple regression in SPSS).

### PROJECTS

**Scraping and Analyzing the SF Chronicle in Python** Feb 2019  
In Python, scrape sfchronicle.com for article information, like title, author, and date using html requests. Analyze website sections for most-used words using nltk for natural language processing and matplotlib for visualizations.

**Querying and Analyzing Baseball Payrolls in SQL and Python** Feb 2019  
In SQL and Python, query data from database, create visualizations, and analyze. Used numpy, pandas, seaborn, and matplotlib for Python. Found that payroll increases as performance, win-rate, and home game attendance increases.

**Predict Handwritten Numbers in R** Dec 2018  
A set of R scripts that predicts handwritten numbers. Uses k-nearest-neighbors machine learning algorithm coded from scratch to predict numbers with 95% accuracy.

**College Scorecard Data Exploration and Analysis in R** Oct 2018  
Explored and analyzed data on US colleges using R and ggplot2. Found differences in test scores between public and private schools, the impact of test scores on earnings, and differences in in-state and out-of-state tuition across states.

**Analyzing Apartment Listings from Craigslist in R** Oct 2018 - Nov 2018  
In R, clean, explore, and analyze apartment listings web-scraped from Craigslist. Used ggplot2, ggmap, and stringr. Found different apartment costs in LA, Sacramento, and SF and how apartments differ depending on price and location.

### AWARDS

**Certificate** for outstanding achievement in presentation of research, UC Davis Apr 2019

**Eagle Scout**, Boy Scouts of America May 2016