

CHUCHU CHENG

781-956-2369 ◇ cccbnu@gmail.com ◇ Sunnyvale, CA ◇ [Linkedin.com/in/chuchucheng/](https://www.linkedin.com/in/chuchucheng/)

EDUCATION

Boston College , PhD in Quantitative Psychology, Statistics concentration	08/2015-05/2020
Brandeis University , MA in Psychology, Statistics concentration	2012-2014
Beijing Normal University , BS in Psychology	2008-2012

TECHNICAL SKILLS

Languages	R, Python, SQL, SAS, Stata
Packages	sklearn, seaborn, nltk, re, BeautifulSoup, Selenium; caret, tidyverse
Tools	Jupyter, Flask, AWS, Markdown, Tableau, Shiny, LaTeX, SPSS, Qualtrics
Methods	Multivariate Statistics, Predictive Modeling, Hypothesis Testing, Experiment Design

EXPERIENCE

Insight Data Science	01/2020 - Present
<i>Data Science Fellow</i>	<i>San Francisco, CA</i>

- Developed a Flask app *SmartShopper* on AWS EC2 to predict popularity of cosmetics discount sales on dealmoon.com and provide purchase suggestions.
- Scraped 4k+ web pages with *Selenium* and extracted features using regex, bigrams, tf-idf with *re*, *nltk*.
- Built regression tree, random forest, gradient boosting, xgboost models with *sklearn*, *xgboost* in Python and optimized RMSE by 30%.

Boston College	09/2015 - Present
<i>Graduate Researcher</i>	<i>Chestnut Hill, MA</i>

- Proposed methods to estimate bootstrap CIs of model fit in SEM, evaluated with parallel simulations.
- Developed iterative algorithm using Newton's method and bootstrap functions in R package *lavaan*.
- Increased the accuracy of interval converge by 10-40% and reduced computation time by 20%.

Statistical Consultant, Teaching Assistant

- Advised and mentored university researchers with data analysis in a series of cross-cultural studies.
- Taught statistics & research methods sessions, provided training of statistical packages to 500+ students.

Hay Group	05/2013-08/2013
<i>Analyst Intern</i>	<i>Beijing, China</i>

- Queried from relational database and conducted data mining using SQL for marketing presentations.
- Analyzed 3 years salary data of insurance companies and visualized data in segment reporting.
- Collaborated with consultants for projects under tight deadlines, contributed to 9 contracts.

RELEVANT PROJECTS

Factors Associated with Eviction Rate in Seattle, WA	05/2019
---	---------

- Extract key factors of eviction: rent, population, income, using linear regression and random forest.
- Presented interactive visualization for geographic distributions of variables with Tableau.
- Received Best Visualization Award in the data hackathon of SDSS 2019.
- Provided actionable solutions for eviction problems to local agencies and government stakeholder.

Airbnb's Booking Rate	09/2018
------------------------------	---------

- Performed data cleaning, missing data imputation and feature engineering on 20K+ data points.
- Built classification models and performed hyper-parameter tuning to optimize the model.
- Provided advanced insights, product recommendations and marketing suggestions.