

# Gregory Faletto

13509 Moorpark St., Apt. 3  
Sherman Oaks, CA 91423

Phone: (973) 349-2153  
Email: [Gregory.Faletto@marshall.usc.edu](mailto:Gregory.Faletto@marshall.usc.edu)  
Website: <https://gregoryfaletto.com/>  
GitHub: <https://github.com/gregfaletto>

## Education

Ph.D. Data Sciences and Operations (Statistics Track), University of Southern California Marshall School of Business (2018 - )

Advisor: Dr. Jacob Bien.

B.A. Physics, Washington University in St. Louis (2009). GPA: 3.85

Graduated Phi Beta Kappa.

Graduated with College Honors in Arts and Sciences.

Dean's List every semester.

Recipient of William George Eliot Scholarship.

Recipient of National Merit Scholarship.

## Research

Feature Selection Research Project (2018 - )

Examining (1) feature selection for linear models with highly correlated covariates and (2) assessment of feature selection methods.

## Coursework

*Graduate Coursework (USC)*

Fall 2019:

[Math 547: "Mathematical Foundations of Statistical Learning Theory"](#) (Instructor: Dr. Steven Heilman).

[Math 541B: "Introduction to Mathematical Statistics"](#) (Instructor: Dr. Stanislav Minsker).

[GSBA 604: "Regression and Generalized Linear Models for Business Applications"](#) (Instructor: Dr. Gourab Mukherjee).

Spring 2019:

[Math 541A: "Introduction to Mathematical Statistics"](#) (Instructor: Dr. Steven Heilman).

[ISE 620: "Foundations of Stochastic Processes"](#) (Instructor: Dr. Sheldon Ross).

[DSO 607: "High Dimensional Statistics and Big Data Problems"](#) (Instructor: Dr. Jinchi Lv).

Fall 2018:

Math 505A: "Applied Probability" (Instructor: Dr. Sergey Lototsky).

Econ 613: "Economic and Financial Time Series I" (Instructor: Dr. Hashem Pesaran).

EE-588: "Optimization for the Information and Data Sciences" (Instructor: Dr. Mahdi Soltanolkotabi).

### *Online Courses (MOOCs)*

"Databases: Relational Algebra" (Widom, Jennifer), Stanford Lagunita. *Currently in progress.*

"Databases: SQL" (Widom, Jennifer), Stanford Lagunita. *Currently in progress.*

"Econometrics: Methods and Applications" (Franses, Philip Hans), Coursera/Erasmus University Rotterdam. *Completed May 2017.*

"Databases: Introduction and Relational Databases" (Widom, Jennifer), Stanford Lagunita. *Completed January 2017.*

"Statistical Learning" (Hastie, Trevor/Tibshirani, Rob), Stanford Lagunita. *Completed December 2016.*

Text: *An Introduction to Statistical Learning with Applications in R* (James/Witten/Hastie/Tibshirani).

"Intro to Python for Data Science" (Schouwenaars, Filip), Datacamp. *Completed December 2016.*

"Data Visualization with ggplot2 (Part 1)" (Scavetta, Rick), Datacamp. *Completed December 2016.*

"Machine Learning" (Ng, Andrew), Coursera/Stanford University. *Completed April 2016.*

"Intermediate R" (Schouwenaars, Filip), Datacamp. *Completed April 2016.*

"Introduction to R" (Cornelissen, Jonathan), Datacamp. *Completed April 2016.*

### *Notable Undergraduate Coursework*

Economics 102: "Macroeconomic Theory" (University of California, Los Angeles)

Math 4650: "Analysis I" [Real Analysis] (California State University Los Angeles)

Econ 451: "Environmental Policy" (Washington University in St. Louis)

MGT 430: "Introduction to Entrepreneurship" (Washington University in St. Louis)

Econ 103B: "Introduction to Political Economy: Microeconomics" (Washington University in St. Louis)

CSE 131: "Computer Science I" (Washington University in St. Louis)

CSE 126: "Introduction to Computer Programming" (Washington University in St. Louis)

## Academic Employment

Washington University in St. Louis (2009)

Research Assistant, Professor Rebecca Messbarger, Department of Romance Languages.

Washington University in St. Louis (2007 - 2008)

Lab Assistant, Professor Ken Kelton, Department of Physics. Assisted on project in hydrogen storage in quasicrystals. Prepared samples of alloys and collected and interpreted data. Cataloged alloys by heating with a DSC and obtaining patterns with x-ray diffractometer. Presented findings to research group.

## Independent Projects

[Orange County R Users Group Hackathon 2019](#) (2019)

Participated in team that won "Best Model" at the Orange County R Users Group Hackathon 2019. Over the course of two days, developed a model associating health outcomes in California counties with levels of water pollutants. More information at [gregoryfaletto.com/2019/05/19/our-entry-in-the-ocrug-hackathon-2019/](https://gregoryfaletto.com/2019/05/19/our-entry-in-the-ocrug-hackathon-2019/).

[Valuing a Digital Firm](#) (2018 - )

Using a methodology developed by scholars from Wharton to value Buffer, an online firm, using publicly disclosed data. Developed a model for customer retention and forecasted residual customer lifetime. More information at [gregoryfaletto.com/blog](https://gregoryfaletto.com/blog).

[Fragile Families Challenge](#) (2017)

Cleaned data and created principal components regression and lasso-penalized linear regression models to predict three different outcome variables in a training set with only 2,121 observations, nearly 13,000 covariates, and a great deal of missing data. All three of my models finished in the top 40% of entries and outperformed a baseline null model. Code and description: <https://github.com/gregfaletto/fragilefamilies>.

## Invited Talks

[Orange County R Users Group](#) (2019)

Presented models that won "Best Model" at the Orange County R Users Group Hackathon 2019. More information at [gregoryfaletto.com/2019/05/19/our-entry-in-the-ocrug-hackathon-2019/](https://gregoryfaletto.com/2019/05/19/our-entry-in-the-ocrug-hackathon-2019/).

## Programming Languages/Software Packages

R (proficient), SQL (intermediate), Matlab (intermediate), Excel (intermediate), Python (some experience), Stata (some experience).

## Industry Employment

### [ZipRecruiter](#), Data Analytics Research Intern (2017 - 2018)

Developed an adaptive lasso-penalized logistic regression model to predict whether individual job seekers would apply to a given job listing in order to infer job seekers' revealed preferences for monetary and fringe elements of compensation packages. Model includes features related to compensation package, controls, and interactions between controls and compensation package features in order to understand how different kinds of job seekers value different forms of compensation. Chose research objective, data, and observational study design; made decisions about modeling process; analyzed results, considering applications and possible next steps; and worked independently with only a handful of in-person and teleconference meetings. Paper available at <http://www.latutor.net/research/>.

### [Live Nation](#), Machine Learning Intern (2017)

Working remotely, independently created an algorithm in R to predict future concert set lists of bands based on past set list data. Code and description: <https://github.com/gregfaletto/setlistpredictor>

### [Independent Audio Engineer](#) (2012 - 2017)

Audio engineer mostly working in live sound. Past clients include Stevie Wonder, Lady Gaga, John Mayer, Jackson Browne, Dave Chappelle, Moby, Jack Black, etc. More information at <http://www.gregfmusicproducer.com/>.

### [Private Academic Tutor](#) (2009 -)

Tutor at the college and high school level. Subjects: Calculus, Statistics, Physics, Chemistry, GRE, SAT/ACT. More information at <http://www.latutor.net/>.