

Greg Faletto

DATA SCIENTIST · PH.D. STUDENT

| ✉ gregory.faletto@marshall.usc.edu | 🏠 gregoryfaletto.com/ | 🐙 github.com/gregfaletto | 🔗 <https://www.linkedin.com/in/gregfaletto/>

Employment History (Summary)

Google

Los Angeles, CA

DATA SCIENTIST INTERN

May 2021 - Aug. 2021

- Designed framework and wrote Python code for simulation studies to quantify flaw in prior method for estimating treatment effects (experimental change in metric) from A/B test data.
- Developed a new method for treatment effect estimation and wrote Python code implementing method.
- Simulation studies showed that new method reduced bias and mean squared error of treatment effect estimates by well over 99%, while also controlling Type I error rate much better, in settings where the assumptions of the prior method were violated. Delivered similar results even when assumptions of prior method were satisfied.
- Submitted just under 4000 lines of reviewed Python code implementing method and simulation studies. Left detailed documentation (both within code and separately) as well as explanations of areas of improvement for proposed method and suggested next steps.

ZipRecruiter

Santa Monica, CA

DATA ANALYTICS RESEARCH INTERN

Jul. 2017 - Jan. 2018

- Using R, developed a model to predict whether a job seeker will apply to a given job listing to infer preferences for listed perks.
- Accessed data via SQL. Chose observational study design and model form. Conceived outcome metric, constructed from raw data.
- Final deliverable was approximately 3500 lines of R code and a written report.

Live Nation

Los Angeles, CA

MACHINE LEARNING INTERN

Feb. 2017 - May 2017

- Working remotely, created time series algorithm in R to predict future concert set lists of bands from data. See Github [link].

Education

University of Southern California Marshall School of Business

Los Angeles, CA

PH.D. IN DATA SCIENCES AND OPERATIONS (STATISTICS TRACK); ADVISOR: DR. JACOB BIEN

Aug. 2018 -

- Current research: developing a novel feature selection method; developing a recommendation system in with a tech startup.
- Responsibilities as research assistant include developing and writing code for novel methods and simulation studies to test them (write 2000 – 4000 lines of code per semester).
- Coursework includes: machine learning, regression and generalized linear models, causal inference, applied econometrics, time series and panel data modeling, stochastic processes, dynamic programming, mathematics of statistical learning, convex optimization.

Washington University in St. Louis

St. Louis, MO

B.A. IN PHYSICS

Aug. 2006 - Dec. 2009

- Graduated Phi Beta Kappa and with College Honors in Arts and Sciences. Dean's List every semester.

Data Science Competitions

City of Los Angeles and RMDS

2020 COVID-19 COMPUTATIONAL CHALLENGE

Jun. 2020

- Team won second place. Interpreted general goals without specific directions. Presented solution at IM Data Conference 2020.
- Used daily neighborhood test results across Los Angeles to estimate real-time numbers. Forecasted new infections in each neighborhood a day ahead using a GLM panel data model. Finally, converted forecasts to interpretable risk scores.
- Full solution writeup and code available at grmds.org/2020challenge.

Orange County R Users Group

ORANGE COUNTY R USER GROUP HACKATHON 2019

May 2019

- Team won "Best Model." Developed a model associating health outcomes with levels of water pollutants. More details here [link].

Skills

- **Languages:** R (proficient; experience with packages like **tidyverse** collection, **caret**, **Keras**, etc.), Python (proficient, including libraries like **NumPy**, **pandas**, **scikit-learn**, **SciPy**, **matplotlib**, etc.), SQL (intermediate), Matlab (intermediate).