Steven Chu

161 Lundy's Lane, Apt 1 | San Francisco, CA 94110
stevenc1@gmail.com | 919-637-6189
https://github.com/chusteven

EDUCATION	**MS IN ANALYTICS** (Summer 2014 – July 2015 expected)	**SAN FRANCISCO, CA**

University of San Francisco

BS IN MATHEMATICS; SECOND MAJOR IN ENGLISH (Fall 2006 – Spring 2010) **CHAPEL HILL, NC**
University of North Carolina at Chapel Hill

RELEVANT COURSES

Statistics: Advanced Machine Learning, Linear Regression Analysis, Time Series Analysis, Multivariate Statistics
Data Skills: Data Acquisition, Relational Databases, NoSQL Databases, Distributed Computing, Data Visualization
Business Skills: Business Strategies, Business Communications, Marketing Analytics, Web Analytics
Languages: Python, R, MySQL, PostgreSQL, NoSQL, SAS, JavaScript (for visualization in D3)

SELECTED PROJECTS

USF INTERNSHIP: FANDOR (Fall 2014 – present) **SAN FRANCISCO, CA**
Apply statistical techniques such as machine learning, hypothesis testing, and cohort analysis to answer important business questions. Data gathered and cleaned with MySQL queries; analyses performed in Python and R.
- Defined and calculated metric for understanding efficacy of current film recommendation system
- Defined and calculated lifetime value of a Fandor user under various business models
- Developed framework for performing user path analysis on Fandor's web application
- Defined and calculated metric for understanding the value of an individual film

USF COURSEWORK: MACHINE LEARNING PREDICTION (Spring 2015) **SAN FRANCISCO, CA**
Predicted how a user would rate an artist-track pair, achieving R^2 value of 0.505 with random forest algorithm (Kaggle competition). Predictions based on demographic and musical content features. Compare to Kaggle leader of $R^2 = 0.65$.
- Performed data exploration, munging, feature engineering, and analyses in Python and R
- Created class object in Python for implementing particular framework for training prediction algorithms

USF COURSEWORK: TIME SERIES ANALYSIS (Fall 2014) **SAN FRANCISCO, CA**
Forecasted US non-agriculture monthly workforce employment numbers, achieving mean absolute deviation value of 173.04 on test set forecasts using differenced time-series SARIMA model.
- Created model and tested for plausible stationarity, white noise processes, ARCH, and GARCH effects using econometric analyses and tools in R
- Forecasted future months' employment numbers with 95% confidence bands

USF COURSEWORK: LINKEDIN DATA ACQUISITION (Fall 2014) **SAN FRANCISCO, CA**
Identified skills market on LinkedIn – measured by probability using Beautiful Soup and RESTful APIs in Python.
- Accessed companies and employees of interest with vetted API services as well as web-scraping
- Compiled most employed skills by calculating frequencies of skills for employees of interest

USF COURSEWORK: SENTIMENT ANALYSIS (Summer 2014) **SAN FRANCISCO, CA**
Classified movie reviews as either positive or negative, achieving 80% accuracy on test sets.
- Cleaned individual text files, created dictionary of positive and negative works, and computed probabilities of each class using Naïve-Bayes algorithm in Python

WORK EXPERIENCE

SAINT ANN'S SCHOOL (Fall 2011 – Summer 2014) **BROOKLYN, NY**
Saint Ann's is an independent, non-sectarian private school. The school eschews traditional grading in favor of holistic narrative reports. Teachers design flexible curriculums around student interests.
Math Teacher, Independent Research Mentor, Assistant Basketball Coach
- Planned and administered daily lessons for 3rd grade, 6th grade, Algebra I/II, Geometry, and Calculus classes
- Collaborated with other teachers to generate curricular standards, class rosters, and departmental structure

MATCH CHARTER PUBLIC HIGH SCHOOL (Fall 2010 – Summer 2011) **BOSTON, MA**
The MATCH School works to close the achievement gap through a "No Excuses" approach.
MATCH Corps Tutor, Administrative Assistant to the Director of Recruiting, English 12 TA
- Worked with Recruiting Department to hire MATCH Corps tutors
- Tutored freshmen, sophomores and juniors in Math, English, History, Chemistry and Spanish