

# Greg Simon, Ph.D.

gregorygsimon@gmail.com      (805) 746-1668

## Professional Summary

I like to solve problems. I teach statistics at Western Governors University. I have over three years of experience programming and automating with Python, primarily data wrangling, data visualization, and analytics for our students' outcomes.

I am experienced with the practical application and theory of state-of-the-art machine learning and statistical methods, with successful projects in forecasting, classification and regression, clustering, natural language processing and image recognition.

Portfolio available at: <https://gregorygsimon.github.io>

## Technical Skills

**Languages:** Python, SQL, Spark, Mathematica, R (some), Julia (some)

**Other:** Amazon Web Services (AWS), Google Script, Linux, Docker

**Certifications:** AWS Certified Machine Learning Specialty (2020, ID#RX11FD2CEB4E1WKF)

## Employment Experience

**Western Governors University - Course Faculty** 2017 – Present

- Teaching applied statistics to undergraduate and graduate students
- Led machine learning project for aligning course practice problems with positive student outcomes
  - Gathered data by web scraping (Python, selenium) and Salesforce API (SOQL)
  - Assessed feature importance of practice problems on assessment outcomes using tree methods (Random Forests, XGBoost) and permutation importance
- Department expert in Data Visualization (in Python)
  - responsible for creating publication quality data visualizations for students using real world data sets (matplotlib, plotly, seaborn)
- Led project creating seven fully animated approximately 10 minute videos with visualizations of complex concepts with real-world datasets
  - animations and data visualizations created in a Java animation language (Processing)
- Created model predicting student success on assessment based on pre-assessment
  - Model used to help prioritize and guide outreach for students at higher risk of failure
- Created and maintained automation pipeline in Python to email customized welcome message to students
  - ETL pipeline in Python using Salesforce API (SOQL) and Pandas and implemented using Microsoft Office Suite API (win32com in Python)
  - Used by department daily for over a year before enterprise solution was piloted and launched across university

**University of Michigan** - *Graduate Student Instructor* 2011-2016, *Lecturer* 2016-2017

- Primary instructor for calculus series, including Differential Calculus, Integral Calculus, Vector Calculus, and Different Equations with Linear Algebra
- Led labs in MATLAB and in Mathematica for Vector Calculus and Differential Equations

## Other Experience

**Image recognition 6-week project with TrainX**

07/2019 – 09/2019

*Data Science consultant*

- Co-lead of cohort of five student practitioners creating minimal viable product for image recognition to predict make and models from webcam car images
- using AWS Sagemaker, deep learning / CNN, wrangling diverse data sources for training images
- Contact: Aubrey Agee, aubrey@trainx.ai

**Forecasting state-wide markets via Covid-19 Twitter sentiment**

- Python data pipeline to gather tweets and geocoder to parse user locations
- using deviations in regional sentiment to predict stock fluctuations in local businesses

**NLP sentiment analysis and topic extraction from online reviews**

- web scraping text reviews and using word embedding (word2vec) and dimensionality reduction to summarize flavor profiles of whiskeys from 34,000 unstructured text reviews

## Education

**PhD in Mathematics**, 2016, University of Michigan. Ann Arbor, MI

- Thesis research in monstrous moonshine, utilizing dimensionality reduction techniques to simplify matrix representations – implemented algorithms in *Mathematica*

**B.A. in Mathematics**, 2010, University of California, Santa Cruz. Santa Cruz, CA

- *cum laude* and *highest honors in the major* (mathematics)

### *Other coursework*

DataCamp - Data Scientist in Python Certificate (2018)

Udemy – AWS Certified Machine Learning Specialty

*updated: June 7, 2020*