# **Katherine Wood**

Seattle, WA

☑ katherinemillerwood@gmail.com | ★ katherinemwood.github.io | ♠ katherinemwood | in katherine-m-wood

Skills \_

Languages Python, R, JavaScript, HTML, CSS, SQL

**Tools** Mechanical Turk, Presto, Jupyter Notebooks, RStudio

Statistics Bootstrapping, model selection, null hypothesis significance testing, Bayesian hypothesis testing and

model fitting, regression, machine learning

### Experience \_

## **Insight Data Science**

Seattle, WA

DATA SCIENCE FELLOW January 2020 - present

- Developed an app with Streamlit and deployed on AWS to predict product recalls from consumer complaints.
- Extracted meaningful information from unstructured text with natural language processing to pair complaints with recall announcements using Python and NLTK.
- Used TF-IDF encoding and logistic regression with scikit-learn to classify consumer complaints that referred to a recalled product, achieving .73 AUC.

Facebook Menlo Park, CA

**CORE DATA SCIENCE INTERN** 

May 2018 - August 2018

- Mined and cleaned 5TB+ data sets with SQL and Presto to create in-depth analyses and visualizations.
- Trained and validated gradient-boosted decision tree models to predict future user behavior.
- Established a pipeline to collect, clean, and organize daily incoming data using Hive and Spark.
- Deployed a survey of 5,000 users to inform product decisions.
- Communicated results and insights to product teams and stakeholders.

## **Association for Psychological Science**

Freelance

PROGRAMMER AND STATISTICIAN

2016 - 2018

- Conducted meta-analysis and meta-regression in R for two collaborative international replication efforts.
- Programmed a cross-platform experiment in Python with Psychopy to administer a behavioral experiment to over 2000 participants.
- Created publication-ready data visualizations summarizing thousands of individual data points.

## **University of Illinois at Urbana-Champaign**

Champaign, IL

GRADUATE RESEARCHER

2011 - 2015

- Ran large-scale online experiments with Mechanical Turk using boto3, Javascript, HTML, and CSS, leading to 5
  peer-reviewed publications and 2 conference presentations.
- Built laboratory-based behavioral experiments with Python, leading to a further 2 peer-reviewed publications and 1 conference presentation.
- Analyzed behavioral data in R using Frequentist, Bayesian, and non-parametric methods.
- Developed an open-source Python application that recovers raw data distributions from summary statistics using systems of Diophantine equations. (http://bit.ly/CORVIDS)
- Built an automated Mechanical Turk pipeline for a study on trauma and recovery using boto3, TurkGate, and Qualtrics for the Kinesiology Department.

Teaching Assissment Fall 2016

- Led discussion and laboratory sections for 30 students enrolled in Introduction to Experimental Psychology.
- Designed lectures and lab activities to teach fundamental concepts.

#### **Education** \_

# **University of Illinois at Urbana-Champaign**

PHD IN PSYCHOLOGY

M.S. IN PSYCHOLOGY

2015 - 2020

2015 - 2017

#### **UC Berkeley**

B.A. IN PSYCHOLOGY 2011 - 2015