

JUAN SHISHIDO

Data Scientist

EDUCATION

University of California, Berkeley

School of Information

Master of Information Management and Systems, 2016

College of Natural Resources

Bachelor of Science, Environmental Economics and Policy, 2009

PROFESSIONAL EXPERIENCE

Graduate Student Researcher, D-Lab

Berkeley, CA — 2014-present

Senior Analyst, EnerNOC

Walnut Creek, CA — 2010-2014

Assessed program impacts and established causality using regression models and algorithm-based methods. Created matching algorithms to select non-participants with observable characteristics similar to program participants. Developed customer-specific energy analytics, comparing each customer's cost and consumption to similar households identified using hierarchical clustering. Contributed to strategic design and implementation decisions for this data product, including touchpoint frequency, behavioral-based messaging, comparison group metrics, and geographic expansion. SQL joins and queries of large electric utility tables for identifying duplicates. Designed and selected samples, both simple and stratified.

Research Assistant, Berkeley Policy Associates

Oakland, CA — 2009-2010

Data munging with Stata: merging, reshaping, reconciling, developing descriptive statistics, identifying outliers, and creating new features.

PROJECTS

Dynamic Hexbin Map [Python, D3]

Kaggle: Titanic Survivors [Python]

Dbot [Python]

Boaltscape [Python]

Iterative Proportional Fitting [Python]

Portfolio [HTML, CSS, JavaScript]

D-LAB WORKSHOPS

Geocoding

Pandas

Twitter API/Bots

Regular Expression

LANGUAGES

Python · R · SAS · SQL · Stata

INTERESTS

Causal inference

Machine learning

Data visualization

Open data and APIs

Experimental design

RECENT READS

Python for Data Analysis

Data Science at the Command Line