Robert Brown

⌂ elreb.me • ⬡ elreb • ▣ rbrtbrwn
✉ rebrown@berkeley.edu • ▢ 510.xxx.xxxx

SKILLS

Programming Languages: Python • R • SQL • SAS • LaTeX • Bash • HTML/CSS • JavaScript

Tools: Apache Airflow • PySpark • Pandas • NumPy • Scikit-learn • TensorFlow • AWS • D3.js • Tableau • Flask • Shiny

Machine Learning/Statistics: experimentation/AB testing • time-series analysis • hypothesis testing • regression • ensemble methods • dimensionality reduction • regularization • clustering • survivor analysis • NLP • neural networks

EXPERIENCE

DATA ENGINEER | Berlin Brands Group October 2021 - Present • San Francisco, CA
- Created and implemented computing, reporting, and ETL infrastructure.
- Developed multiple Django applications using dedicated servers and AWS cloud products. Applications include a prospecting tool for brand acquisition, monitoring product purchase behavior and predict future revenue growth using data ingested from data mining, operational and third-party data API end points.

DATA SCIENTIST | Los Angeles Co. Public Health Dept. January - October 2021 • Los Angeles, CA
- Created multiple interactive internal and public facing dashboards displaying infectious disease data for LA County (R and Python).
- Developed a whole genome sequencing (WGS) database and implemented ETL pipelines, implemented in Airflow, processing tens of inputs and 100s of GBs of epidemiologic, genomic sequence data and associated metadata.

HEALTH DATA SCIENCE FELLOW | Insight Fellows September - November 2020 • San Francisco, CA
- Consulted with an organization to provide a churn model using client records (Python).
- Used advanced SQL querying to aggregate together 10+ tables, and used contextual knowledge to engineer features such as an engagement score to increase model performance.
- Used survivor analysis to generate a probability of of client churn over time, as well as used a binary classifier using logistic regression with lasso regularization to achieve 78% ROC AUC.

EPIDEMIOLOGIST II | Alameda Co. Public Health Dept. August 2018 - August 2021 • Oakland, CA
- Lead analyses and surveillance reporting of tuberculosis and over 70 other notifiable acute communicable diseases such as the measles and COVID-19 (R and SAS).
- Served as Data Chief for the Data Branch for the counties COVID-19 ICS response, managing up to 25 data entry staff, program staff and other epidemiologists leading data system development and analytical capacity.
- Developed and maintained ETL pipelines, and 8 SQL Server databases, and complex VBA featured Access applications with PowerShell installation and versioning.

DATA ANALYTICS INSTRUCTOR & TA | UC Berkeley Extension July 2019 - August 2020 • Berkeley, CA
- Taught fundamentals of computer science, programming, Git, Bash, data munging, visualization, web development, project management, and machine learning in six-month long UC Berkeley Data Analytics Boot Camp.

RESIDENT DATA SCIENTIST | Method Data Science Nov 2018 - April 2019 • Remote Work
- Created an ensemble machine learning algorithm (R) to predict if and when a hip replacement surgery will be needed using EHR and insurance data for a Bay Area biotech start-up with 77% AUC.
- Developed a Dash (Python) dashboard using KPI's and unsupervised learning to segment customers to create targeted marketing resulting in 13% increase in customer base within 6 months.

EPIDEMIOLOGIST | Ca. Rural Indian Health Board October 2017 - August 2018 • Sacramento, CA
- Designed state-wide disease surveillance data collection tools, as well as provided IRB and data management, and wrote grants to secure future funding.
- Developed a data warehouse to store internal data and various state data sources, and provided various automated reporting for various grant tracking and for various leadership staff.

GRADUATE STUDENT RESEARCHER | UC Berkeley August 2016 - March 2017 • Berkeley, CA
- Conducted analysis on the accuracy of population estimates of those with a genetic disorder or genetic determinants (e.g. cancer, congenital malformation, etc.) with those of elusive populations (e.g. drug addiction) using Bayesian analyses (R).

PUBLIC HEALTH DATA SPECIALIST | Google May - December 2016 • Mountain View, CA
- Curated data, & provided data analysis (Python) of public health & medical research to improve the information in the health knowledge graph & related algorithm performance.

EDUCATION

UC BERKELEY | MPH in Biostatistics & Epidemiology May 2017 | Berkeley, CA

UC DAVIS | BS in Neurobiology, Physiology, & Behavior, BA in Psychology May 2011 | Davis, CA