

# Dan Loew Data Scientist & Survey Researcher

Montclair, NJ • 773-310-2385 • drloew77@gmail.com • [linkedin.com/in/danielloew](https://www.linkedin.com/in/danielloew) • [github.com/danloew](https://github.com/danloew)

---

## PROFESSIONAL PROFILE

Data science professional with extensive experience designing, managing, and analyzing scientific surveys in the fields of public health and market research. Transitioning into data science with formal training in Python, data cleaning/wrangling, statistical analysis, and machine learning. Passionate about leveraging the power of data science in service of the public good.

## SKILLS

**DATA SCIENCE:** Machine Learning, Linear, Logistic, and Multivariate Regression, Classification, Random Forest, Data Cleaning/Wrangling, Descriptive and Predictive Statistics, Python, SQL, Pandas, NumPy, SciKit-Learn, SciPy, SPSS, Excel, Relational Database Queries, Hypothesis Testing, Sampling, Data Visualization, familiarity with NLP, Data Clustering, SVM, Naïve Bayes, kNN, Gradient Boosting, Network Analysis, Deep Learning

**SURVEY RESEARCH:** Survey and Questionnaire Design, Data Collection, Data Processing and Delivery, Cognitive Testing, Survey Methodology, Client Presentations

**LEADERSHIP:** Full Life Cycle Development, Project Management, Client Services, Staff Management, Process Improvement, Training, Coaching, Mentoring, Team Development, Transparent Communication, Budgeting

## EXPERIENCE

### Data Scientist, Freelance, Springboard Data Science Career Track

August 2019 – Current

#### Predicting Cancer Mortality in United States Counties: 2015 (<https://github.com/danloew/Springboard-2nd-Capstone>)

- Built a regression model predicting cancer mortality rates in 97% of all U.S. counties for the year 2015, based on a feature set of socioeconomic, geographic, environmental and dietary features.
- Merged datasets from the Census, clinicaltrials.gov, cancer.gov, the EPA and the USDA, and utilized Python, Excel, NumPy, Pandas, SciKit-Learn, OLS linear regression, Ridge and LASSO regression, Random Forest, Stochastic Gradient Descent Regression, and Kernel Ridge Regression.
- Identified a series of Pearson's correlations and salient predictor features from a Ridge Regression model that had an accuracy of 0.64 and a RMSE of 16; improved the model from an initial accuracy of 0.49 through merging multiple datasets, the calculation of geographic distance variables, and the logarithmic and exponential transformations of key features.

#### Targets: Using Machine Learning Classification Models to Identify Salient Predictors of Cannabis Arrests in New York City, 2006-2018 ([github.com/danloew/SpringboardFirstCapstone](https://github.com/danloew/SpringboardFirstCapstone))

- Built a series of predictive models that differentiate cannabis crime (and its subtypes) from all other crimes in NYC between 2006 and 2018 via machine learning binary classification, thereby identifying the features that have the strongest relationships with cannabis crime.
- Used NYC Open Data set, Python, Excel, NumPy, Pandas, SciKit-Learn, Logistic Regression, Random Forest, Hyperparameter Tuning, Draw.io
- Identified a series of salient predictor features from skillful binary Logistic Regression models that had accuracy scores ranging from 0.68 to 0.84, F1 scores ranging from 0.77 to 0.87, and ROC AUCs ranging from 0.67 to 0.91, all while taking measured steps to eliminate data leakage

### Senior Analyst, Abt Associates, New York, NY

April 2017 – March 2019

- Continued with the work of my Senior Survey Director and Research Analyst roles at Abt SRBI
- Took on a self-directed curriculum of study in data science supported by Abt's tuition reimbursement and learning programs, resulting in a certification in Applied Data Science from Columbia University and two certifications in Statistics Foundations from LinkedIn Learning
- Mined Abt's data stores to create a summary database of all of Abt's historical and current projects in the Food Security and Agriculture subject area, providing a resource of Abt's capabilities for future research proposals
- Co-managed the Building Bridges and Bonds (B3) survey, which was a fully in-person survey of recipients of parenting and employment social service programs in five areas of the United States, valued at \$3.5 million.

**Senior Survey Director, Abt SRBI, New York, NY**

July 2012 – March 2017

- Continued with the work of my Research Analyst role at Abt SRBI
- Took on directorship of a portfolio of surveys valued at an average of \$5 million annually
- Coordinated, managed and trained large teams of field interviewers, data processing staff, CATI and web survey programmers, sampling specialists, and junior project managers to manage complex research projects to successful conclusion
- Designed questionnaires to accurately capture data in service of the client's research goals
- Wrote methodology and summary reports of survey results, analyzed results and presented them to clients
- Interfaced with clients to ensure project success and satisfaction
- Managed sample to maintain validity and reliability of data, including list, RDD and address-based sample frames
- Conducted statistical weighting to national targets, making sample data representative of target populations
- Practiced cognitive interviewing, including analyzing open-ended answers, writing cognitive reports and amending survey questionnaires based on the results
- Wrote and monitored budgets to enhance profitability of projects

**Research Analyst, Abt SRBI, New York, NY**

June 2008 – June 2012

- Directed phone and web surveys of all scales in concert with the Senior Vice President of the Health Department, which served as the basis for public policy development, basic research and corporate marketing strategies
- Wrote custom SPSS code to check data simulations to ensure that the survey program was collecting data correctly, assigning data to the correct variables, and that the data map was set up correctly for the client's needs
- Screen-tested survey programs in all modes to ensure precise data collection, directly providing quality control to survey programmers
- Designed and maintained quality of cross-tabulations of data to accurately illustrate survey data sets with perfect precision, customized to client specifications

**Survey Specialist I & II, NORC at the University of Chicago, Chicago, IL**

June 2004 – February 2008

- Managed data collection for the National Science Foundation's Survey of Earned Doctorates (SED) valued at \$3.5 million annually, resulting in a consistent response rate of 92%, an exceedingly high response rate for any survey
- Co-managed data collection for the Survey of Small Business Finance, a Federal Reserve Board survey valued at \$2.7 million, of how small businesses in the United States used credit
- Promoted to level II in June 2006
- As part of a team, wrote the only perfectly scored contract renewal proposal in all of NORC's history
- Managed focus groups, paradata collection efforts, report writing, public relations, and new tech initiatives

**EDUCATION AND CERTIFICATION**

***Data Science Career Track, Springboard***, 550+ hours of hands-on curriculum, with 1:1 industry expert mentor oversight, and completion of two in-depth capstone projects. Mastering skills in Python, SQL, data analysis, data visualization, hypothesis testing, and machine learning.

***DataCamp and LinkedIn Learning***, 20+ certifications in data science and statistics, available on LinkedIn profile

***Applied Data Science Certification, Columbia University Engineering/Emeritus***, 12-week certification course in basic and advanced Python, SQL, using Python to query SQL data frames, statistical distributions, sampling, hypothesis testing, data visualization, data cleaning/wrangling, text mining (using Natural Language Processing), and machine learning for regression, classification, decision trees, and data clustering

***Master of Arts, Social Sciences, University of Chicago, Chicago, IL***

Relevant Courses: Statistics, Survey and Questionnaire Design, Linguistics

***Bachelor of Arts, Psychology with Honors, University of Wisconsin-Madison, Madison, WI***

Relevant Courses: Psychometric Methods, Experimental Psychology, Psychopharmacology

**PUBLICATIONS**

Full list of 14 publications and presentations available on LinkedIn profile, including articles in the Journal of Personality and Social Psychology, the Journal of Elder Abuse and Neglect, and presentations at the national AAPOR conference.