**JENNIFER BROWN**

Chicago, IL • 419.494.5592 • jabrown5@gmail.com​ ​• [/in/jennifer-brown-data-scientist](http://www.linkedin.com/in/jennifer-brown-data-scientist) • [/github.com/jabrown5](https://github.com/jabrown5?tab=repositories)

**STATEMENT**

I am an experienced Data Scientist with a background in public health, personal wellness and survey research methodologies – skills which have allowed a deep understanding of the data cycle especially as it relates to health and wellness. I have worked for major government clients using my expertise in programming, data collection, analytics and cleaning, data visualization, and research presentation to produce quality insights from a multitude of data. I aspire to continue working to improve the health and wellness of individuals through insights found from a variety of data - whether it’s been scraped or drawn from a survey.

**SKILLS | TECHNOLOGIES**

**Technologies​:** Python | SQL | Tableau | HTML | CSS | JavaScript | Jupyter Notebooks | Github | Excel | Word | PowerPoint | SAS | D3.js | AWS | PySpark

**Skills​:** Web Scraping | Data Cleaning | Supervised and Unsupervised Machine Learning (ML) | Predictive Modeling | Data Visualization | Data Analysis | Natural Language Processing | Neural Nets | Big Data | Survey Research Methodology | Project Management |   
**Python libraries​**: Pandas | NumPy | SciPy | Matplotlib | Seaborn | Scikit-learn | Natural Language Toolkit | Beautiful Soup | ScraPy | Tensorflow | Keras

**DATA SCIENCE PROJECTS** February 2021 - Present

**Using Tweets from the 115th and 116th Congresses to identify political affiliations**

* Scraped XXXX tweets from members of the 115th and 116th Congresses.
* Created an unstructured dataset by scraping 40,000 submissions posted by users in the Body Weight Fitness (bodyweightfitness) and the Weightlifting subreddits.
* Cleaned and transformed data in preparation for modeling using feature engineering, natural language processing, and SpaCy
* *Identified topics associated with Body Weight Fitness and Weightlifting using a logistic regression classifier*
* *Achieved a recall score of .6944 by using SMOTE to address the class imbalance in the dataset*
* *Utilized Latent Dirichlet Allocation to identify topics discussed in each group*

**Classification model with NLP to learn about users from the Body Weight Fitness and Weightlifting Subreddits**

* Created an unstructured dataset by scraping 40,000 submissions posted by users in the Body Weight Fitness (bodyweightfitness) and the Weightlifting subreddits.
* Transformed data in preparation for modeling using feature engineering and natural language processing
* Identified topics associated with Body Weight Fitness and Weightlifting using a logistic regression classifier
* Achieved an accuracy score of 0.88

**Using Regression models to predict home sale prices in Ames, IA**

* Explored a Kaggle dataset of home sales from 2006 to 2010 provided by the Ames’s Assessor’s office
* Used data visualization and feature engineering techniques to analyze and manipulate data into a format usable for modeling
* Evaluated a number of models using multiple linear regression and LASSO or Ridge regularization and K-Nearest-Neighbors
* Developed a model that was able to predict housing prices with an R2 score of 0.922

**SAT score comparisons between states with Compulsory Requirements (2018 - 2019)**

* Examined SAT score data for states / territories with compulsory requirements to determine what, if any, differences exist
* Descriptive statistics indicate that some states have higher average scores for both years and some states are consistently lower scores

**Detecting Drowsiness using image data and a Convolutional Neural Network**

* Created a dataset of open and closed eyes using Kaggle images and ‘Faces In The Wild’ images
* Created and refined a Convolutional Neural Network using grid searching to achieve a PR AUC score of 0.9814
* Applied model to image capture using computer vision to identify and sound an alarm when eyes are closed for more than 1.5 seconds

**PROFESSIONAL DATA SCIENCE EXPERIENCE**  
**NORC at the University of Chicago –** Chicago, Illinois January 2015 - February 2016

**Principal Research Analyst, Senior Data Analyst**

* Delivered clean datasets and reports to government clients for actionable policy decisions.
* Resolved data collection system errors to ensure data accuracy.
* Implemented data collection and cleaning system enhancements to streamline the data cleaning and delivery process.
* Collaborated across teams to win major grants and contracts
* Presented research results to client stakeholders and at professional conferences.

**ADDITIONAL PROFESSIONAL EXPERIENCE**

**Freelance –** Chicago, Illinois December 2016 - February 2021

**Data Analyst and Visualization Architect**

* Consulted with clients on JavaScript language programming questions.
* Developed JavaScript based visualizations for clients using D3.js and Tableau.

**General Assembly –** Chicago, Illinois January 2017 - June 2017

**Full Stack Web Development Teaching Assistant**

* Taught front-end web development skills and techniques to immersive students.
* Assessed students on class assignments and class projects.

**NORC at the University of Chicago –** Chicago, Illinois May 2009 - December 2014

**Survey Director, Survey Specialist**

* Budgeted project funding and managed project resources.
* Ensured confidentiality of respondent data during data collection by ensuring project staff compliance to security measures.
* Created survey questionnaires and tested instrument on focus groups and in cognitive interviews.
* Designed data collection and sampling plans for large-scale data collection efforts.

**EDUCATION**

**General Assembly**

Data Science Immersive May 2021

Web Development Immersive December 2016

**Bowling Green State University, Bowling Green, Ohio**

ABD Doctor of Philosophy in Sociology

Quantitative Coursework:​ Probability and Statistics, Quantitative Survey Research

**Bowling Green State University, Bowling Green, Ohio**

Master of Arts in Sociology

Quantitative Coursework:​ Probability and Statistics, Quantitative Survey Research

**The Pennsylvania State University, World Campus**

Graduate Certificate in Applied Statistics

Quantitative Coursework:​ Probability and Statistics

**Central Michigan University, Mt. Pleasant, Michigan**

Bachelor of Science in Sociology and Psychology

Quantitative Coursework:​ Probability and Statistics, Calculus