

SUE-ELLEN DUFFY

Boston, MA | 617-631-8909 | duffysueellen@gmail.com | [linkedIn.com/in/sue-ellenduffy](https://www.linkedin.com/in/sue-ellenduffy)

Education

M.Sc.in Data Analytics and Computational Social Science, GPA: 4.0

Feb 2025

University of Massachusetts Amherst

Relevant Courses: Machine learning (R, Python), advanced data-driven storytelling (R), text as data (R), social networks (R), program and policy evaluation, independent study: argument analysis (R, Python), quantitative analysis (R, Python), data science fundamentals (R), research design (R), survey research methods (R), mapping geographic data with ArcGIS (ArcGIS, R)

Certificate: Social Justice Mediation

Dec 2020

Vital Village Social Justice Mediation Institute

B.Sc. Public Health: Community Engagement & Education, GPA: 3.7

May 2013

University of Massachusetts Amherst

Skills and Credentials

Programming Languages: R, Shiny, LaTeX, ArcGIS, Python, SQL

Analysis Skills: Data Cleaning and Preparation, Qualitative and Quantitative Data Analysis, Text Mining, Survey Analysis, Machine Learning, Data Visualizations, Statistical Modeling (Regression, Causal Inference), Experimental Design, Predictive Modeling, Data Storytelling, Workforce Analytics, Health Informatics Analytics

Work Experience

State of Maine, Office of Behavioral Health

Feb - Oct 24

Data Analyst and Project Coordinator

- Led a cross-functional team of analysts, grant managers, and developers to pilot a statewide patient matching system, successfully matching 100,000+ patients across four disparate datasets.
- Spearheaded ETL workflows to integrate data for substance use treatment programs, enhancing accessibility and enabling data-driven policy decisions (R, Oracle).
- Partnered with data owners to create comprehensive data dictionaries and asset inventories to improve cross-agency data sharing and governance.
- Developed and implemented data governance strategies ensuring accuracy, security, and usability of integrated health records.

Abt Associates, Cambridge, MA

Sep 18 - Dec 23

Freelance Data Collector & Assessor

- Conducted 100+ structured classroom observations to collect behavioral and educational data on preschool-aged children.
- Scheduled and facilitated 150+ assessments with K-5 students across Massachusetts, adhering to rigorous protocols to maintain coding reliability and data integrity.

Jewish Vocational Services (JVS), Boston, MA

Feb - Dec 19

Academic Services Researcher

- Researched causes of community college “stop-out” among non-traditional students through surveys and 20 in-depth interviews.
- Identified key themes including unclear academic goals, financial hardship, choosing jobs over education, and family commitments among other reasons.
- Conducted literature review on college persistence among non-native English-speaking populations.
- Presented research findings to executive leadership and funders, influencing updates to curricula and program structure.

Harvard Graduate School of Education, Cambridge, MA

Sep - Dec 18

Research Assistant

- Administered 50+ structured assessments in Boston Public Schools to evaluate cognitive and behavioral development in early learners.
- Collected, cleaned, and submitted data to lead researchers for longitudinal analysis.

Fair Foods, Boston, MA |

Feb 16 - Sep 2018

Data Collector & Community Engagement Lead

- Designed and executed data collection strategies to assess community food access needs.
- Conducted interviews and surveys to guide improvements in food distribution initiatives.
- Managed volunteer teams and coordinated logistics while fostering culturally responsive community relationships.

UMass Alliance for Community Transformation, Amherst, MA |

Jun 13 - Dec 14

Facilitation Trainer

- Co-developed and delivered facilitator training for seven peer-led university courses and workshops.
- Trained 60+ students using anti-oppression frameworks and social justice pedagogy.
- Conducted qualitative analysis research of partner organizations to improve future training programs.

Projects

Predicting Stroke with Synthetic EHR Data

Fall 24

- Applied machine learning models to synthetic electronic health records (EHR) to predict stroke risk among Massachusetts residents aged 30–70.
- Developed logistic regression models to predict stroke risk using synthetic health records, achieving 87% accuracy with known risk factors and 95% without, revealing potential issues of model overfitting and spurious patterns.
- Explored ethical and analytical implications of using black-box machine learning techniques in healthcare settings, particularly around model interpretability and clinical trust.
(R, Tidyverse, caret)

Syringe Service Program Evaluation Plan

Spring 24

- Developed a data-driven evaluation framework for assessing the impact of Syringe Service Programs (SSPs) on HIV transmission and health outcomes among people who use drugs (PWUD).

Healthcare Network Analysis of People Who Use Drugs (PWUD) Using Synthetic EHR Data

Spring 24

- Mapped healthcare journeys of PWUD using electronic health records (EHR), highlighting treatment fragmentation and provider centralities. (R, Synthea)

Glassdoor Reviews Sentiment vs. Star Rating: Unlimited PTO Policies

Fall 23

- Conducted natural language processing (NLP) and sentiment analysis on Glassdoor reviews to examine employee sentiment around unlimited PTO policies.
- Mined 65,000+ Glassdoor reviews using RSelenium.
- Matched review sentiment with corresponding star ratings to uncover discrepancies between written narratives and overall satisfaction.