

Samantha Roberts

saroberts5@gmail.com · (603) 443-0086 · linkedin.com/in/samantha-roberts-profile · samantha-roberts.github.io

Core Competencies and Skills

| **Programming:** Python (Matplotlib, NLTK, Numpy, Pandas, PySpark, Scikit-Learn), R (Dplyr, GGplot, Leaflet, Lubridate, Shiny, Stringr, TidyR, Web Scraping), SAS (Programming 1: Essentials, and Programming 2: Data Manipulation Techniques), SQL, HTML, MATLAB | **Machine Learning/Statistics:** Classification, Regression, Clustering, Decision Trees, Deep Learning, Dimension Reduction, Feature Engineering, Hypothesis Testing, Modeling, Neural Networks, Predictive Analytics, Random Forests, Statistical Analysis, Supervised/Unsupervised Learning, Time Series Analysis | **Software:** AWS, Hadoop, JMP, MariaDB/HeidiSQL, Microsoft Office, MongoDB, Power BI, QGIS, Tableau | **Capabilities:** LEAN Six Sigma Green Belt Certification, Data Management, Data Pipelines, Strong Communication Skills, Leadership, Organization Skills, Problem Solving, Team Collaboration, Time Management

Projects

- Amazon Customer Reviews: Predicting 5-star review rating using Natural Language Processing and sentiment analysis. Included Document-Term Matrices, TextBlob, VADER, dimension-reduction techniques (UMAP, PCA) and machine learning algorithms (KNN, Random Forest, Gradient Boost, XGBoost)
 - Analytics Summer Practicum Project: Analyze income inequality using Tableau, Python (classification regression) and R
 - Lending Club: Analyze data on Lending Club's financial scandal using Python (Numpy and Pandas) to clean, standardize, and perform unsupervised learning techniques (Principal Component Analysis and k-Means Clustering)
 - Senior Capstone: Collect data from UNH students and analyze, using JMP Statistical Software and R, how various habits and lifestyle trends affect their GPA
-

Work Experience

360 Intel | Manchester, NH

Data Science Consultant

October 2019 – May 2020

- Build a clearer and more communicative dashboard for clients (Visualizations)
- Explore how to quantify a survey question's importance to the overall business report (Data Cleaning, One-Hot Encoding)
- Use Natural Language Processing and variable weighting (Tokenization, Stemming, Polarity Scores using VADER)

The Music Hall | Portsmouth, NH

Data Science Consultant

October 2019 – May 2020

- Identify artists that are ideal (successful) candidates for this venue
- Use web scraping and modeling (Linear Regression, Train/Test Sets, Random Forests, Feature Importance) to build an algorithm to automate the process of identifying these musicians

S.A.J. Dawgs | Lebanon, NH

Proprietor

June 2016 – September 2018

- Small business owner of food cart selling hot dogs to the public
- Balance profits while controlling expenses through Excel (started with no profit the first year to turning ~\$3,000 profit by the end of third summer)
- Marketing of products and service through social media (Facebook), newspaper and local events

StartDate Labs | 10 Water Street, Lebanon, NH 03766

Intern

June 2015 – August 2016

- Projects at talent acquisition company
 - Salesforce and Excel: change and consolidate salesman districts
 - Customer service: respond to emails, go into accounts to fix customer problems
-

Education

University of New Hampshire, Durham, NH

Master of Science in Analytics and Data Science

May 2019 – May 2020

- President of Data Driven Women

University of New Hampshire, Durham, NH

Bachelor of Science in Statistics with a minor in Spanish

August 2015 – May 2019

- Pi Mu Epsilon (Math Honor Society) Member
- Treasurer and Player of the Women's Club Ice Hockey team: Budget presentations, collect dues
- Nourish UNH member: Provide educational programming and outreach about eating well to peers
- Buddies Without Borders member: Pairs international and domestic students to create relationships

Lebanon High School, Lebanon, NH

High School Diploma

August 2015 – May 2019