SAMUEL LOGAN BELL

https://medium.com/@sambellsoup | https://github.com/sambellsoup
https://www.linkedin.com/in/sam--bell/
sambellsoup@gmail.com | 757-652-7433 | Brooklyn, NY 11216

Data Scientist and machine learning engineer with a penchant for visualizations and making data interactive. I have experience in data acquisition and data modeling, statistical analysis, and machine learning with a background in Psychology and Education. I bring a unique skill set to team building and project management that helps companies realize their potential with data.

TECHNICAL SKILLS

AWS (S3, EC2), Bokeh, Github, Heroku, Linux, NumPy, Python, Pandas, scikit-learn, SQL, Tableau

EMPLOYMENT HISTORY

RookiePlay New York, NY
Data Science and Engineering Consultant, May 2020-Present

- Scripted **ETL** pipeline using **NLP** techniques for topic modeling and data analysis
- Created dashboards using **Python's Dash** and **Tableau**
- Engineered job recommendation system

Boy Scouts of America New York, NY
Exploring Executive, May 2014 - August 2017

- Increased membership for Exploring by **16.5%** through managing several marketing outlets
- Established **9** sustainable after school programs hosted by local businesses

TECHNICAL PROJECTS

Video Game Recommendation System - https://github.com/sambellsoup/VGR2
A system that recommends users new games based on their previous ratings

- Scrape and populated a database of user reviews from Metacritic using **Python's BeautifulSoup** package and standard library regular expressions
- Program with **Pandas** to process and analyze data
- Use **Singular Value Decomposition(SVD)** algorithm and clustering to predict user ratings and generate recommendation
- Deployed as a web application written in Flask backed by a Postgres database

Global Voter Turnout - https://github.com/sambellsoup/Voter_Turnout
Evaluating the most influential predictor of voter turnout by combining UN and IDEA datasets

- Use **Bokeh** and **Tableau** to create interactive geographic visualizations of a large dataset
- Find the best model for predicting voter turnout on a global scale and for individual countries
- K fold cross validation and modeling with **Gradient Boost** and **Random Forest**

Tree Coverage - https://github.com/sambellsoup/Trees
Predicting tree species growth in Colorado with Kaggle dataset.

- Utilize **XGBoost**, **Decision Trees**, and **Random Forest** to create the best model for prediction
- Evaluate accuracy and precision on a multiclass model using scikit-learn
- Manipulate features with **Principal Component Analysis** and then **Lasso** for feature selection

EDUCATION

DataCamp - AWS Boto in Python, Intro to Scala, Fall 2019
Flatiron School, August 2019 New York, NY

George Mason University Fairfax, VA
B.S. in Psychology, May 2010