**Seth Billiau**

sethbilliau@college.harvard.edu • [sethbilliau.com](http://sethbilliau.com/) • <https://github.com/sethbilliau> • (616) 443-3973

 **Technical Skills & Projects**

**Programming Languages:** Python, R, SQL, HTML, CSS, JavaScript

**Relevant Libraries:** numpy, pandas, Tensorflow, Keras, scikit-learn, statsmodels, Beautiful Soup, Selenium

**Visualization Tools:** GIS, RShiny, Tidyverse, plotly, D3, Tableau, matplotlib, seaborn

**Visualizing Ballot Roll-Off in Kent County, MI** June 2020 – August 2020

Built a web scraper to collect election results using publicly available data sources. Cleaned and aggregated data using tidyverse. Produced an interactive GIS webmap of election results in Kent County, MI, to highlight precincts susceptible to down-ballot roll-off. Created for use by 7 Kent County Commission candidates in the 2020 August Primary Election.

**Predicting pCR in Breast Cancer Patients Using Gene Expression** March 2020 – May 2020

Created ANN, Random Forest, and XGBoost models to predict pathological complete response (pCR) from gene expression in breast cancer \_patients. Outperformed proprietary Prosigna benchmark in terms of test AUC in a sample of 1,309 patients collected across 18 studies.

**Generating Spotify Playlists from a Cold Start** September 2019 – December 2019

Used content-based recommendation and k-Means clustering to generate Spotify playlists of an arbitrary size from a single input song. Developed an objective function to train the model on 1 million real Spotify playlists. Reported results in an HTML write-up with matplotlib visualizations.

**Work Experience**

**Booz Allen Hamilton** Washington, DC **Data Scientist, Consultant**  July 2021 - Present

Hired as a member of the Strategic Innovation Group to develop algorithms, write scripts, use automation, and find answers to help clients make important decisions. Active Secret clearance.

**Michigan House of Representatives** Lansing, MI **Legislative Aide**  May 2020 – September 2020

Lobbied state lawmakers in favor of restorative justice legislation including reforms to expungement of public criminal records, drug sentencing guidelines, and cash bail by interpreting quantitative results and analyses. Problem-solved during the COVID-19 crisis by connecting constituents to services like the MI eviction diversion and food assistance programs.

**Harvard University** Cambridge, MA

**Teaching Fellow**, **Statistics Department** December 2019 – May 2020, September 2020 – May 2021

**Course Assistant, Mathematics Department**  September 2018 – May 2019

Taught advanced topics in probability theory, Markov chains, Monte Carlo methods (Statistics Department), linear algebra, real analysis, multivariable calculus, and vector calculus (Math Department). Courses included examples of applied data science and statistical inference in R.Awarded for Distinction and Excellence in Teaching.

**Education**

**Harvard University**  Cambridge, MA

A.B. in Statistics (Data Science), minor in Government August 2017 - May 2021

Relevant Coursework: Data Science I & II, Generalized Linear Models, Statistical Machine Learning

**Leadership & Interests**

**Harvard Open Data Project**  Cambridge, MA

**Vice President, Treasurer** August 2018 – May 2020

Oversaw the financial and creative direction of Harvard’s premier club for data scientists, data journalists, and open data enthusiasts. Taught data science bootcamps in R and Python on numpy, pandas, and tidyverse. Published journalism and data visualizations in the [Harvard Political Review](https://harvardpolitics.com/harvard/student-group-financing/) and with the [Harvard Open Data Project](https://medium.com/harvard-open-data-project/harvard-student-opinion-on-the-second-democratic-debate-6a2818506b01).

**Interests:** Avid sports fan, choral music enthusiast, 2-time state champion public speaker, backpacker