# ERIC BOXER

ecboxer.github.io

Email : ecb2198@columbia.edu
Mobile : (347) 525-3516


## Education

- **Columbia University** — New York, NY
  *MS in Data Science; GPA: 3.8 / 4.0* — *Aug 2018 – Dec 2019*
  *Coursework: Applied Machine Learning, Databases, Machine Learning Pipelines, Algorithms, Network Analysis, Causal Inference, Statistical Inference and Modeling, Data Visualization*

- **SUNY Stony Brook** — Stony Brook, NY
  *BS in Mathematics and Economics, Minor in Philosophy; GPA: 3.9 / 4.0; Phi Beta Kappa* — *Aug 2013 – Dec 2016*
  *Coursework: Symbolic Logic, Differential Equations, Topology and Geometry, Game Theory*

## Professional Experience

- **Swiss Re** *Data Science Intern* — May 2019 – Present

  **Early Warning Signs:** Developed a novel outlier detection method for news articles, to identify potentially critical information for decision makers. Viewing the corpus as a network (with keywords as nodes and co-occurrence of keywords in an article as edges), we are able to identify important articles as those which form links between otherwise distant terms.

  Performed exploratory data analysis, predictive modeling and risk scoring for marine insurer CV Starr. With financial and geospatial data from the client, we explored drivers of attritional and named-event loss. Then, we incorporated in-house natural catastrophe models and publicly available fire incident data to create and validate policy- and broker-level risk scores. Recommendations were presented to the client, who expressed interest in continued engagement with the analytics team and a reinsurance renewal with Swiss Re.

- **Frac.tl** *Data Visualization Developer* — March 2019 – June 2019

  Investigated medical student demographics and MLB data. Scraped datasets for visualizations and article content.

  Created static and interactive data visualizations in RShiny and ggplot2. For example, an interactive map of MLB ticket prices over the period 1950 to 2017, with icon size mapped to ticket prices and icon placement according to team location.

- **WS Group, LLC** *Accountant* — Jan 2017 – Aug 2018

  Compiled balance sheets and profit and loss reports for tax and audit preparation for corporate clients. Performed data entry for one hundred corporate clients. Built Excel UserForms and worksheet templates to simplify expenditure analysis.

## Data Science Projects

- **Unlikely Words** *Columbia University* — Dec 2018 – Jan 2019

  Creation of a novel metric intended to aid comparison of RNN architectures. The metric examines the difference between n-gram probabilities assigned by a language model and an empirical probability distribution from a corpus. The project was motivated by the goal of a metric more interpretable than loss scores and more quantitative than "readability".

  **Github Repo** https://bit.ly/2InNqKk

- **Pandas Contributor** *Columbia University* — Sep 2018 – Present

  Building documentation locally and pushing commits to the pandas-dev Github repo to be discussed with core contributors and merged into the master branch.

  **Example pull request** https://bit.ly/2DSbqjg

- **Inkrement Author** *Columbia University* — Sep 2018 – Dec 2018

  Standard data science packages lack tools for incremental data visualization, which will grow in importance as datasets increase beyond the scope of standard visualization methods. The inkrement package is a Matplotlib wrapper, available on PyPI, for visualizing samples of datasets during exploratory analysis.

  **Github Repo** https://github.com/Ecboxer/inkrement

## Programming Skills

- **Python**: Numpy, Pandas, PyTorch, Keras, Scikit-Learn, nltk, spaCy, Matplotlib, Scrapy
- **R**: Tidyverse, Rmarkdown, base, stats, plsRglm
- SQL, Git, AWS, Elasticsearch, Jupyter, Emacs, Linux, LaTeX, Javascript, D3