

DAVID RAYMOND KEARNEY

Senior Data Scientist

🌐 davidraymondkearney.com/

@ david.raymond.kearney@gmail.com

in davidrkearney

🔗 davidrkearney

📍 Chicago, IL

SKILLS

- **Languages:** Python, SQL, R, Unix/Bash, Git, JSON
- **Machine Learning Techniques:** Pyspark, Dask, XGBoost, Random Forest, h2o, NLP, LDA, Hyperparameter tuning, Cross-Validation, exposure to Keras, TensorFlow, Pytorch, Prophet, Sparkling Water, DataRobot.
- **Statistics:** Experimental Design, Randomized Control Trials, Bootstrapping, ANOVA, t-tests
- **Tools:** Spark, Hadoop, Hive, Pandas, Jupyter-Lab, NumPy, SciPy, Plotly, Seaborn, Matplotlib, Scikit-Learn

EXPERIENCE

Senior Data Scientist

CVS Health

📅 2019-

📍 Greater Chicago Area

- Created automated member and provider engagement campaigns that leverage e-mail, IVR, SMS and live calls to increase medication adherence.
- Designed experiments and employed randomized control trials to consistently improve campaigns by identifying the most effective campaign variants. Implemented campaigns with **JSON** and evaluated with **bootstrapping**, **ANOVA** & **t-tests**.
- Productionized machine learning models to identify members at risk of behaving in a way that contributes to poor health outcomes and most likely to benefit from campaigns.
- This includes training, **hyperparameter-tuning** and deploying adherence models with **3B+** observations to predict medication adherence using **Pyspark**, **h2o**, **XGBoost** & **Hive** on **Hadoop** clusters.

Insight Data Science Fellow

Fenix International Consulting Project

📅 2019

📍 Remote (Durham, NC)

- Queried **PostgreSQL** 350GB **Amazon Redshift** database for 50K+ accounts and 80K+ GSM enabled solar kits. Developed a predictive classification model for Fenix International to predict probability of missing first loan payment using **Python**, **NumPy**, **Pandas**, **Jupyter Notebooks**, **scikit-learn**, **Matplotlib**, **Seaborn** & **Plotly**.
- Trained **Random Forest Classifier** validated model with **k-fold cross-validation**. Improved previous accuracy baseline by **25%** and previous precision baseline by **40%**.
- Further trained and validated **Random Forest Regression** model to predict the amount repaid on loan at 30, 60, and 90 days.

Ph.D. Candidate and Researcher

Duke University

📅 2013 - 2019

📍 Durham, NC

- Developed and presented **NLP** and **LDA** analysis of the manuscript texts of New Faces Conference from 2000 to 2018 using **Python**, **NumPy**, **Pandas**, **Seaborn**, **Matplotlib** & **Jupyter Notebooks** as well as **n-grams** and **named-entity recognition**. Delivered analytical report based on the analysis to the 2019 New Faces Conference.

EDUCATION

Ph.D. in Political Science

Duke University

📅 May 2019

📍 Durham, NC

M.A. in Political Science

Duke University

📅 May 2016

📍 Durham, NC

B.A. in Political Science & International Studies

Summa Cum Laude

Iowa State University

📅 May 2012

📍 Ames, IA

PROJECTS

Analysis of Chinese Economic and Fiscal Data (Dissertation)

- Investigated the relationship between political connections & the distribution of billions of USD in yearly fiscal transfers.
- Retrieved & cleaned biographical, fiscal & economic data & stored them in a database with 35K+ observations & 100+ features.
- Employed linear & logistic regression analysis to evaluate the impact of political connections on the distribution of fiscal transfers **Python**, **NumPy**, **Pandas**, **Jupyter Notebooks**, **Plotly**, **R**, & **dplyr**.
- Trained **Random Forest Regression** model to predict the distribution of fiscal transfers. Validated model with **k-fold cross-validation**.

Analysis of Chinese Development Assistance to Africa

- Employed **Python** and **R** to engage in linear & logistic regression analysis on database of 1.9K+ Chinese development finance projects in 3.5K+ locations in 50 African states between 2000-2011.
- Combined & analyzed with public opinion metrics derived from representative sample of African citizens covering the same period.
- Implemented data cleaning, management of multiple data sources & visualization with **GIS** tools & **ggplot2**.

HONORS & AWARDS

- Emerging Leaders Institute (Selective Professional Development Program), Duke Graduate School, 2018
- Society of Duke Fellows (Selective Honor Society for Duke Graduate Students), 2016