# Lead Data Scientist with Capital One

drewfustin@gmail.com | 773.304.7316

#### **Education**

PhD Physics (2012) - The University of Chicago [Experimental Dark Matter Physics - Juan Collar, advisor]

MS Physics (2005) – The University of Chicago

BS Physics, Mathematics, Astronomy (2004) - Drake University

## **Skills**

Python (pandas, matplotlib, scientific stack, etc.), Tableau, MATLAB, Excel/VBA, R, d3.js, Google Maps, Mapbox/turf.js MongoDB, SQL (MySQL, SQL Server, PostgreSQL), AWS/EMR

Statistics (linear/logistic regressions, machine learning, time series analysis, Bayesian statistics, etc.)

Writing production code, product ownership, experience at tech companies going through a merger, IPO, and acquisition

# **Experience**

## Digital H2O

Data Scientist (2014-present)

- Developed the individual well hydrocarbon production allocation algorithm, quantifying assets within the product
  - o Built a unique offering within the oilfield intelligence software space by utilizing data from a variety of sources to determine the most likely contributions each well makes to its member lease's production
    - Wrote unit tests within Python backend to provide QA during ETL and ensure complete data coverage
- Created an industry-unique method to determine water production from individual well assets
  - o Collected water-to-hydrocarbon ratios from governmental filings and projected the trajectory of this ratio over time, giving an extremely accurate estimation of monthly water production
- Eliminated lag in production data from delayed governmental filings by forecasting monthly hydrocarbon and water production values to present day and into the future (post-ETL, pre-production automated Python task)
  - Utilized both Bayesian Markov Chain Monte Carlo (Python emcee) and ARIMA methods (R forecasting package through Python RPy) to project the trajectory of production values across a given time series
  - o Created an Information Criterion used to select a winning forecasting model on an individual well basis
- Designed data-driven presentations (Tableau/PowerPoint) for the CEO to deliver at national conferences, resulting in multiple sales leads converting into site licenses, making up nearly half of recent company revenue

### GrubHub Inc.

Insights Analyst/Data Scientist (2013-2014)

- Built efficiency maximization models for an experimental delivery product launch which went into full-scale production, doubling this project's restaurant coverage in Chicago and allowing for expansion into Los Angeles
  - o Optimized restaurant delivery boundaries using isochrone contours (self-created Python module) utilizing the Google Maps API to maximize delivery driver efficiency/cost by time of day
  - o Predicted future restaurant order volume (Python statsmodels ARIMA) for delivery driver staffing needs
  - o Built the economic model for this initiative to prove project profitability and sustainability
- Co-created the News Bureau program a PR initiative to generate gH data-intensive stories, providing insights to media/trade publications and establishing gH as the thought-leader in the industry
  - o Increased the data-driven story rate six-fold in 2013 compared to 2012, contributing to a near tripling of total media mentions of the gH brands, thus improving gH's SEO value
- Generated the core content for an industry white paper, surveying the effects of delivery fees and minimums on revenue generation for both gH and our restaurant customers
- Developed methods to indirectly identify and cluster previously unknown diner demographics (college students, office workers, hotel travelers, and more) using clustering algorithms and text analysis in Python/Excel

#### The University of Chicago

Graduate Research Assistant/Data Analyst (2008-2012) [also 2004-2007]

- Collaborated on an experiment that set best-in-world dark matter limits
- Characterized background rates and detector efficiencies by comparing (gigabyte-scale) experimental data with multi-interaction neutron scattering Monte Carlo simulations (Los Alamos MCNP package and MATLAB)
- Wrote real-time diagnostic software using PID loops in LabVIEW to monitor and control temperature and pressure settings of vital experimental components to ensure safe remote detector operation

## Susquehanna International Group

Assistant Options Trader (2007-2008)

- Built BI tools for senior trading staff, interfacing Excel through VBA to Bloomberg and in-house tools
- Worked closely with traders and market-makers in an apprenticeship role

