SpotHero

Drew Fustin

data scientist

Collecting Event Data
Build web scrapers to automatically find upcoming event data near our parking garage assets.

Collecting Universal Parking Garage Information
Build web scrapers to identify all parking garages throughout the country with as much data attached as is publicly available.

Custom Data-Driven Pitches for Potential Garage Clients
Generate pitch decks with estimated additional profits for potential parking garage clients.

Parking Garage Selection via Isochrones
Given a driver's destination, allow the selection of only parking garages that are within a certain walking (or CTA) duration. This is accomplished through the calculation of isochrones (contours of constant time) around the destination.

Parking Garage-to-Destination Routing
Push parking garage recommendations to drivers along with transit information to get them to their final destination. Include CTA fees with recommended parking rates when displaying options on a map to the driver.

Predicting Parking Demand Spikes
Use historical data and scraped nearby event times to predict parking demand spikes. Ahead of events, decrease the reactive pricing discounts for drivers going to the event and remove event-adjacent parking garage recommendations for drivers going elsewhere.

Reactive Pricing to Fill Garage Space
Unused parking spots are lost money to the garage operators. By enticing drivers into using empty spots that would have otherwise gone unused by reactively decreasing the prices, we can increase expectation value per unused spot from $0 to the Discounted Price.

Auction-Based Recommendation Rankings
Based on the fee each parking garage pays to us, boost their standing in the sort order or improve their reactive pricing to increase driver likelihood of landing in their garage.

Real-Time Sensor Data
Installing internet-connected in/out traffic counting sensors at gateways to provide real-time data for individual parking garages.
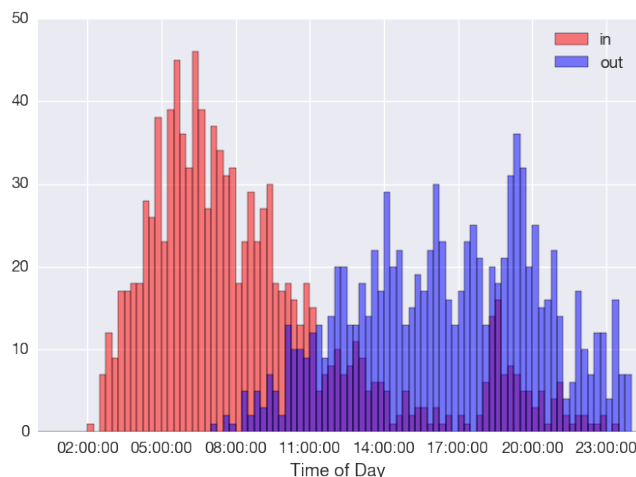
Education

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
    - 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
    - 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
    - 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
    - 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
    - Predicted future restaurant order volume (Python statsmodels - ARIMA) for delivery driver staffing needs
    - 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
    - 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

drewfustin@gmail.com | 773.304.7316

Cover Letter

Hello.

I am completely passionate about creating data-driven product. While I have experience in software development and testing, business intelligence, and data visualization, the ultimate goal of my career is to increasingly influence my company to make data-informed decisions and to provide data-based products.

I have three years experience in data analysis and product development in the tech industry, a year and a half in the financial industry, and five years of full-time data analysis and experimentation in graduate school. I have been a part of a merger between two billion-dollar companies, a very successful IPO, and have just been a part of building a startup to the point of acquisition.

My goal now is to settle into a long-term role where my skills can be honed, and I can transform into a more complete data lead (think chief data officer or head of product, depending on the future company). Ideally, my next role will involve product ownership, management opportunities, and leadership that is passionate about molding my future and expanding my statistical toolset.

Some project examples:

- Restaurant Delivery Boundary Optimization [GrubHub] – *drewfustin.com/isochrones*

  Restaurants often set their delivery boundaries ignorantly – more from a "feel" perspective than anything data-driven. Leveraging GrubHub's geospatial ordering history and the time-cost of delivery drivers, I was able to determine more ideal delivery boundaries. The initial delivery boundaries were set using isochrone contours generated using the Google Maps distance matrix API to perform radial binary searches uniformly distributed around the restaurant. These boundaries were then warped to include areas of dense historical order volume if the expected order volume increase outweighs the cost to send the driver a further distance. These insights allowed GrubHub to scale a delivery experiment into new markets without requiring local geographical and traffic knowledge.

- Hydrocarbon Production Forecasting [Digital H2O]

  The time series of production curves for oil/gas wells typically follow regular decay patterns, with (probably) Poisson distributed shocks. Eliminating the shocks, the regular functional nature of these production curves allowed me to develop exceptionally well-fit projections generated by Bayesian Markov Chain Monte Carlo models that are often more predictive than typical autoregressive methods. Individual well production forecasts like these are completely unique within the competitive product space.

I'd love to hear from you on any ways you think I can help you succeed.

References

Forrest Webb – Lead Data Scientist at Digital H2O
  Supervisor and colleague at Digital H2O who can speak on technical qualifications.
      773.899.6667
      fwebb@digitalh2o.com

Abby Hunt – Director of Public Relations at GrubHub Inc.
  Supervisor at GrubHub who can speak on non-technical qualifications.
      630.464.9798
      ahunt@grubhub.com

Tom Hayden – Senior Financial Analyst at GrubHub Inc.
  Colleague at GrubHub who can speak on technical qualifications.
      313.247.9453
      thayden@grubhub.com

Phil Lynch – Head Trader at DRW Trading Group
  Colleague at Susquehanna who can speak on technical qualifications.
      773.573.7815
      plynch@drw.com