Marshall Krassenstein

📞 (724) 591-4260 | ✉ mpkrass@gmail.com | 🏠 marshallpk.com | ⌗ mpkrass7 | in marshall-krassenstein/

📖 Education

Georgia Institute of Technology - College of Computing | Engineering *Atlanta, GA*
M.S. Computational Data Analytics 3.92/4.00

Northwestern University - College of Arts | Sciences *Evanston, IL*
B.A. Economics 3.70/4.00

🚀 Skills

Software Python, R (**Shiny**, Tidyverse, sparklyR), SQL (Hadoop, OLAP, OLTP), **Tableau**, Alteryx, Git, LaTeX
Analytics Tree-Based Algorithms, Regularized Linear Models, K-Means Clustering, KNN, PCA, Linear Optimization, Explainable AI
Broad Skill Interactive Visualization, Databases, Machine Learning, Chatbots, Flashy POCs, Teaching, Public Speaking

🔧 Work Experience

Data Robot

SENIOR DATA SCIENCE ENGINEER *2023 - PRESENT*

- Scope and create repeatable machine learning and generative AI pipelines with open source frameworks including Kedro, Airflow and Langchain
- Built out, released, and currently maintain **drx**, a library of usable high level abstractions for interacting with the DataRobot platform
- Worked on DataRobot's first "Generative AI Tiger Team" researching novels ways to use Generative AI to support customers and win sales
- Led team to second prize in company hackathon for building the **RFP Monster**, a web application and accompanying Slackbot that uses Generative AI to answer RFP questionairres

SENIOR CUSTOMER FACING DATA SCIENTIST *2021 - 2023*

- Led technical presales and post sales motion for dozens of prospects and existing clients across the United States; Worked internally and externally to successfully run numerous proof of value engagements acting as a Data Science expert in company AutoML products
- Created data pipelines and logistic regression models in large case control study to assist the CDC in pediatric vaccine effectiveness assessment
- Used Pyspark and various packages to create and automate Covid-19 reporting as part of a $10 million contract with the HHS
- Employed webscraping tools (Selenium, BeautifulSoup) to build a database of sitemaps and pricing data for hospitals around the country
- Built and released **Shmapy**, an open source Python package for easily constructing hex and facet maps of the United States
- Reformulated the 'Supply Exchange' into a linear program that helped to match over 200,000 Covid-19 tests to at need organizations

PNC Bank *Pittsburgh, PA*

DATA SCIENTIST, TREASURY MANAGEMENT *2019 - 2021*

- Product owner with 15 developers on Agile team that created and productionized an award-winning portal with over 500 active users; Responsibilities included product development, release planning, data modeling, large group demonstrations and performance tracking
- Performed network analysis and built a payments web application in R Shiny to deliver previously inaccessible transaction detail to end users in a novel format; optimized network to live query a Hadoop database ingesting 20 million transactions daily in a matter of seconds
- Created an XGBoost machine learning model and built an accompanying UI to forecast withdrawals across 7,000 ATMs in company hackathon
- Improved business process by building a CRUD application for tracking sales implementations currently housing $35 million in booked sales

DATA ANALYST *2017 - 2019*

- Completed initiatives from start to finish to create and automate business analytics reports across enterprise including project scoping, query development in SQL, and constructing interactive dashboards in Tableau, which still save hundreds of hours per week
- United disparate data sources to create TMO 360, a one-stop shop providing information on revenue, balances, opportunities, selling wins, and calling habits which is now used as the single source of truth for measuring salesperson effectiveness across the line of business
- Scoped and built out dashboard used by the head of pricing on over 1,000 separate occasions (I checked her view statistics) which continually informs product offerings and individual discount decisions often worth over $10,000
- Designed curriculum for PNC's first analytics development program and lead training in SQL, Tableau, and R for new analysts

MacGregor Global Investments *Chicago, IL*

MARKET ANALYST *2016 - 2017*

🏆 Interests, Awards, and Side Projects

What I do for fun: Running, Crossfit, chess, practicing Spanish, making smoothies, reading fiction and books about data
Pats on the back: PNC Performance excellence award 2019 and 2020. 3rd and 2nd place in DataRobot hackathons 2022 and **2023**
Selected projects: **img2cmap** (package/webapp), Visualizing **human migration** (webapp), **nested cross validation** (blogpost)