Dustin Wicker

919-353-4199
dustinlwicker@gmail.com

Github:
github.com/dustinwicker
LinkedIn:
linkedin.com/in/dustin-wicker

Skills

Python, PySpark, SQL, Amazon Web Services, Tableau, Git, R, Hadoop Ecosystem, Dataiku DSS

Experience

Experienced team leader and data scientist with creative problem-solving mindset and strong interpersonal skills. Proven track record of implementing machine learning models and advanced data mining techniques to turn technically complex challenges into viable solutions. Detail-oriented and innovative critical thinker dedicated to making a difference through data-driven results. Key accomplishments and abilities include:

- Building custom production-level analytics deliverables in workflow process
- Communicating complex concepts to technical and non-technical audiences
- Working collaboratively with diverse sets of professionals to achieve optimal results

Centers for Medicare & Medicaid Services (CMS) / Data Scientist & Team Lead
MARCH 2023 - PRESENT | DENVER, CO (REMOTE) - CONTRACT POSITION

- Leading a team of diverse data professionals in analyzing how initiating and non-initiating parties select independent dispute resolution entities (IDREs) during the dispute process. Responsible for providing CMS with deep insights into the selection process, helping to identify and address inefficiencies, and increasing overall understanding of decision-making patterns.

- Implementing Python-based machine learning algorithms to predict dispute closure durations using a dataset of over 1.6M disputes and 4.3M line items, and conducting comprehensive analysis to dissect the mechanics driving both prolonged and accelerated closures which improves strategic decision-making.

- Developing a script that automates the entire process of creating the Federal Independent Dispute Resolution Public Use File (IDR PUF) and Supplemental Tables using SOQL and Python, ensuring data integrity and providing critical transparency in the implementation of the Federal IDR portal.

- Leveraging SOQL and Python to conduct complex statistical analyses and create compelling data visualizations that respond to ad hoc requests from various CMS teams, providing valuable analytical insights, optimizing data ingestion processes, and informing process improvements.

- Developing interactive Tableau dashboards to empower stakeholders in making data-driven decisions on complex topics by highlighting key trends in a user-friendly format, ensuring critical insights are accessible and actionable.

U.S. Department of State Refugee Processing Center / Senior Data Analyst
JUNE 2021 - MARCH 2023, DENVER, CO (REMOTE) - CONTRACT POSITION

- Built machine learning models and applied natural language processing techniques in Python to understand 15K referred cases' deferral reasons. Used the insights to improve case processing procedures and decrease deferral periods.

- Performed in-depth statistical analyses and analytical techniques on 1M+ rows of data in Python and T-SQL to identify top action items causing delays in refugee processing operations. Visualized the results to upper-level officials through Tableau dashboards.

- Built and redesigned 20+ interactive Tableau dashboards to provide pertinent, time-sensitive information in an easily digestible format to top-level government officials.

- Implemented T-SQL code into 21 Common Data Model tables that allowed for smooth transition of data from initial ingestion into production. Solved crucial issues for processing centers around the world and improved data integrity.

General Dynamics Information Technology / Senior Metrics & Data Analytics Analyst
NOVEMBER 2020 - JUNE 2021, DENVER, CO (REMOTE) - CONTRACT POSITION

- Enhanced and productionalized eight statistical healthcare models in a time-sensitive environment using PySpark, HiveQL (HQL), and Python to support the Centers for Medicare & Medicaid Services (CMS) Center for Program Integrity in its effort to strategically combat fraud, waste, and abuse across Medicare, Managed Care, and Medicaid programs.

- Built data process workflow from 100M+ rows of CMS data for each model using HQL, ensuring the proper population of providers, claims, and payment information was utilized in the data analysis and reporting process to identify fraudulent medical providers and practices.

- Extracted qualitative and quantitative relationships by performing data manipulation and statistical analysis methods which resulted in robust data visualizations, reports, and presentations used by CMS and its partner organizations to further investigate identified cases of fraudulent activity.

Thinkful / Data Science Mentor
NOVEMBER 2018 - AUGUST 2022, DENVER, CO (REMOTE)

- Mentored data science students one-on-one through a six-month online academy specializing in Python, PySpark, and SQL to develop their analytical skills and become proficient in the full data science stack.

- Guided mentees through content such as programming fundamentals, statistical and regression analysis, machine learning, and big data using instructional sessions, code review, and pair programming.

- Provided students with career coaching and industry best practices to assist their transition into the advanced analytics and data science field.

Cognizant / Business Analytics and Insights Senior Associate
AUGUST 2017 - JULY 2019, PARSIPPANY-TROY HILLS, NJ

- Worked with a Fortune 500 rental car company to develop three production-level machine learning models in Python and synthesized the results to predict rental patterns of 4M brand switching customers resulting in a win-back email marketing campaign.

Dustin Wicker

- Developed machine learning models in PySpark using real-time car telemetry data to predict car maintenance failures before they occurred resulting in optimized fleet usage and overall reduction in maintenance costs per vehicle.

- Utilized a variety of text mining techniques in Python on 45K post-rental survey responses to decipher patterns among positive and negative customer experiences.

- Analyzed 25M customers and their 82M rentals through statistical techniques in Python to identify differences between churned and non-churned customers.

- Segmented those 25M customers into five unique clusters using the K-means algorithm in Python to enable the client to send tailored emails and promotions.

- Created bundles of the most popular products purchased together by performing association analysis in R on 35M ancillary product transactions resulting in immediate uplift in online sales.

Northrop Grumman / Graduate School Practicum
SEPTEMBER 2016 - MAY 2017

- Investigated specific questions related to fraud, waste, and abuse such as: DRG upcoding, atypical care patterns, and 30-day readmission rates.

- Implemented techniques including data mining, association analysis, regression analysis, and network analysis using R, SQL, and SAS.

- Analyzed over 12GB of Centers for Medicare and Medicaid Services (CMS) data, containing 2.7 million Medicare claims with approximately 3,400 fields per claim.

Education

Master of Science in Analytics
Institute for Advanced Analytics, North Carolina State University
MAY 2017

Bachelor of Science in Civil Engineering, magna cum laude
North Carolina State University
DECEMBER 2013

Honors: Dean's List seven times, National Collegiate Honor Society
Study Abroad: Università Cattolica del Sacro Cuore