Seth Saperstein

www.sethsaperstein.com

Email : sethsaper@gmail.com
Mobile : +1-248-310-2020

Skills

- **Languages**: Python, C++, SQL, Scala, ReactJS, Java, Swift
- **Technologies**: AWS, Spark, Kafka, Airflow, Terraform, Cloudformation, Troposphere, Airbyte, Hudi, Docker, Jenkins
- **Certifications**: AWS Certified Solutions Architect - Associate, AWS Certified Data Analytics - Specialty

Experience

- **Rocket Homes**   Detroit, MI
  Big Data Engineer   *Jan 2020 - Present*
  - **Data Platform**: Architected and implemented multiple data pipelines, a data platform, and a data lake on AWS to normalize, centralize, and democratize 200+ real estate related source systems as well as serve a site of 1M+ users, scaling coverage to 50 states.
  - **Team Development**: Founded Rocket Homes Data University to encourage technical skill development, providing structure, relevant certifications, and training material to get a team of 10+ engineers AWS Solution Architect certified by Jan 2021.
  - **Orchestration**: Automated data engineering and data analyst processes via configuring Airflow and authoring a series of DAG's, saving 4500+ engineering hours in the next year.
  - **Test Automation**: Architected and implemented end to end test orchestration of data pipelines using snapshot, approval, and load testing methodologies to reduce bugs in production and minimize risk during releases.

- **University of Michigan - Office of Academic Innovation**   Ann Arbor, MI
  Data Science Fellow   *May 2019 - Aug 2019*
  - **Competency Modeling**: Created a feature for a university-led exam preparation platform to give a user their percent competency in a course or subject based on a series of questions using item-response theory and computer adaptive testing models.

- **X-Mode Social**   Reston, VA
  Data Science Intern   *Jun 2018 - Aug 2018*
  - **Algorithm Optimization**: Improved accuracy of a reverse geocoding visitation algorithm on mobile user location data by 30% through iOS SDK modifications for data enhancement, conducting field tests for labeled data, and feature engineering via Pandas.
  - **Data Cleaning**: Automated location data cleansing processes that led to a reduction of 2000 hours per year of analyst time.
  - **Backtesting**: Created a backtesting system to validate returns of stock trading strategies using models powered by location data.

Projects

- **Real-Time Sports Betting Exchange System & Web App**: Implemented a sports betting exchange ReactJS site and cloud system allowing for P2P bets in real-time, leveraging serverless AWS architecture (API Gateway, Kinesis, Lambda, S3, Athena, Glue, Redis, ECS, Elasticsearch, RDS) to execute, store, and process orders.
- **Location Based Ad-Tech Data Pipeline, SDK, & Web App**: Implemented a location collection SDK for iOS in Swift with serverless AWS architecture (API Gateway, Kinesis, Lambda, S3, Athena, and Glue) to ingest, process, store, and query mass amounts of streaming data from iOS devices, visualized in a ReactJS admin dashboard web-application.

Education

- **University of Michigan**   Ann Arbor, MI
  Bachelor of Science: Data Science, Economics, and Mathematics   *Sep. 2015 – Dec. 2019*

Leadership

- **Michigan Data Science Team - VP of Communications**: Hosted the university's first ever data hackathon, MDataHack, consisting of 100+ people, 25+ teams, and 5+ sponsors.