Armaan Thapar

athapar.2017@gmail.com | (914)-486-0735 | github.com/athapar | Atlanta, GA | New York, NY

Skills

Python, SQL, ETL/ELT, AWS (Lambda, Redshift, SQS), Spark, Docker, Kubernetes, pandas, scikit-learn, Power BI, Git, Azure DevOps, PI Historian / AF

Experience

Georgia-Pacific – Senior Data Engineer (Asset Health & Analytics)
Atlanta, GA | 05/2022 - Present

- Owned production data systems for electrical asset health monitoring, acting as the sole technical owner across engineering, operations, and vendor platforms
- Designed and implemented a Python automation pipeline for asset monitoring model creation, integrating request queuing into an existing service to reduce manual effort by 66% and eliminate recurring operational bottlenecks
- Built and operated production data pipelines with AWS Lambda and Redshift to ingest sensor data, perform real-time calculations, and generate forecasts for live monitoring across 4 manufacturing facilities
- Developed a custom Python parser for XML-based process flow diagrams, extracting node metadata and graph relationships to automatically generate downstream material-tracking pipelines
- Led large-scale model synchronization and validation workflows, handling schema mapping, API rate limits, error recovery, and data integrity checks across multiple production systems
- Wrote and optimized SQL pipelines and Redshift schemas supporting analytics use cases and Power BI dashboards used by engineering teams for operational decision-making

Reckitt – Data Analyst: Raw Materials & Quality
Belle Meade, NJ | 01/2021 - 12/2021

- Analyzed spectroscopy and quality data in Excel and maintained raw-materials inventory datasets supporting compliance and operational planning

Diamond S Shipping - Data Analytics Intern: Maritime Operations
Greenwich, CT | 06/2016 - 08/2016

- Analyzed real-time operational data from five crude oil tankers to model fuel consumption as a function of vessel speed and operating conditions
- Built a simulation model incorporating vessel speed, weather, and cost inputs to recommend optimal routing speeds, projecting annual fuel savings of 5–15%

Projects

Master's Thesis | *01/2021 - 12/2021* | [Full Paper] [GitHub Link]

- Developed an object-oriented Python library to simulate interactions between pollutant molecules and analyze molecular cluster dynamics
- Analyzed high-dimensional simulation data to evaluate thermodynamic and spatial properties of molecular systems

Modeling Differential Equations | *10/2018 - 12/2018* | [GitHub Link]

- Modeled chemical reactor dynamics by implementing vectorized numerical solvers (RK4, Euler) in Python
- Visualized and presented reactor behavior through plots using matplotlib and seaborn

Education

The Cooper Union, New York, NY
Master of Engineering, Chemical Engineering — **Dec 2020**
Bachelor of Engineering, Chemical Engineering — **May 2018**