Michael G. Winsor

mgwinsor@gmail.com | github.com/mgwinsor | mgwinsor.dev

EXPERIENCE

Software Engineer — May 2024 – Present
Siemens Mobility — *Singapore*
- Implemented a thee-tier web architecture for deployment of a web application to AWS ECS
- Leveraged NGINX for improved security and delivery of static frontend files
- Created datawarehouse with AWS Athena and S3 for analytic models to read data

Data Engineer — Nov. 2021 – Apr. 2024
ST Engineering — *Singapore*
- Developed batch and streaming ETL pipelines with PySpark on Hadoop and AWS
- Created a GitLab CI/CD pipeline to test, build, and deploy docker containers to Azure
- Maintained data engineering code git repository and oversaw production deployments
- Improved Hive query efficiency by 40% through optimization of file storage across several TBs of data in HDFS
- Realized $100,000 savings via PySpark memory optimization, avoiding expansion of on-prem infrastructure

Software Engineer — Sept. 2020 – Nov. 2021
TUMCREATE — *Singapore*
- Implemented a driving simulation model of a road network to facilitate public transport research
- Lead efforts for a proof-of-concept project to test the real-world application of new public transport concept
- Used Python to collect and analyze vehicle trajectory data via AI to provide insights on the effectiveness of the proof-of-concept project

TECHNICAL SKILLS

Languages: Python, Go, SQL (Postgres, Hive), Bash, HTML
Libraries/Frameworks: Spark, pandas, Polars, NumPy, Matplotlib
Developer Tools: Git, Docker, Kubernetes, GitLab CI, AWS, Azure, Linux, NeoVim

PROJECTS

MRT Station Crowd Density Model | *Python, Pandas, PostgreSQL* — Dec. 2023 – Apr. 2024
- Developed a model to process raw data streamed from all MRT stations in Singapore
- Implemented forecasting to predict future station crowding
- Output from this model is available to commuters through MyTransport.SG app
- Model results were sent to downstream servers as JSON lines via SFTP

SAP Financial Data Pipeline | *PySpark, Bash, NiFi, HDFS* — Jan. 2022 – Nov. 2022
- Built an ETL pipeline with PySpark and NiFi to map SAP financial data with other data sources
- Pipeline allowed downstream applications to unlock cost analysis and forecasting capabilities
- Successfully delivered the pipeline under tight deadlines and maintained it throughout its life cycle
- Implemented AES encryption at rest to protect sensitive financial information

CERTIFICATIONS

AWS Solution Architect – Professional (SAP-C02) | AWS	Oct. 2024
AWS SysOps Administrator – Associate (SOA-C02) | AWS	Aug. 2024
AWS Developer – Associate (DVA-C02) | AWS	July 2024
AWS Solutions Architect – Associate (SAA-C03) | AWS	Nov. 2023
Azure Data Engineer Associate (DP-203) | Microsoft	Aug. 2024

EDUCATION

Master of Science in Transportation Systems — Oct. 2017 – June 2020
Technical University of Munich — *Munich, Germany*

Bachelor of Science in Mechanical Engineering — Sept. 2012 – June 2017
Drexel University — *Philadelphia, PA, USA*