Daniel Chen

Email: danielmc4102@gmail.com
Mobile: (469) 212-4058

EDUCATION

- **The University of Texas at Austin** — Austin, TX
  B.S. Computer Science, B.S. Mathematics; GPA: 3.85 — *Aug. 2018 – Dec. 2022*

EXPERIENCE

- **Affirm** — San Francisco, CA
  Software Engineer — *Jun. 2022 – Feb. 2023*
  - **CI Tests**: Designed Buildkite CI test in GitHub to identify PII usage across 1000+ Dbt models. Written in Python to parse YAML Datahub metadata and SQL text from models, and deployed using Airflow and K8s operators.
  - **SLAs**: Developed Airflow jobs gathering freshness information of Dbt models and source data using Snowflake metadata to provide analysts and BI teams clarity on data freshness, as well as provide a reference for past data inconsistencies to decrease on-call resolution times.
  - **Warehouse Optimizations**: Performed cost analysis on company's Dbt usage and investigated benefits of horizontal/vertical scaling within Snowflake. Findings lead to millions of dollars saved for analysts and our teams spending, and to spread awareness to users on the importance of efficient design and maintenance of data models.
  - **Dbt Project Migration**: Prepared new region infrastructure in Snowflake, involving the configuration of Snowflake roles, databases, tables, and warehouses using Terraform, +1000 Dbt model migration, and coordination of analysts and engineering teams across the company to determine prioritization of the given tasks.
  - **Data Processing**: Created multiple ELT processes using Dbt by request for analysts across the company, improving data transparency, processing efficiency, and availability of company analytics.
  - **Looker Dashboards**: Designed and troubleshooted Looker dashboards for analysts to understand company data. Maintained existing dashboards and corresponding email alerts for inter-team efficiency.

- **Integra FEC** — Austin, TX
  Data Analyst — *Jan. 2022 – Feb. 2022*
  - **Blockchain Analysis**: Designed algorithm with Python scripts to trace blockchain transactions and attribute them to individuals for gathering evidence towards fraud investigations provided by clients. Generated graphs and charts within Jupyter Notebooks using Matplotlib and Seaborn visualization packages.
  - **Municipal Bonds Analysis**: Refactored and provided documentation for research-grade C# project used to generate mark-up and mark-down pricing data of municipal bond sales from financial institutions.

- **H-E-B** — Austin, TX
  Data Engineer Intern — *Summer 2021*
  - **SLA Dashboard**: Worked on React dashboard using PostgreSQL databases to display ETL job statuses, user database permissions, and service-level agreements (SLAs) using Kubernetes, Docker, and Datadog

- **The University of Texas at Austin** — Austin, TX
  Teaching Assistant — *Jan. 2020 – Dec. 2021*
  - **(CS 439) Principles of Computer Systems**: Provided assistance teaching about topics for approximately 450 students over four semesters relating to the x86-64 Linux operating system as well as the C language. Covered topics over CPU scheduling, processes, synchronization, virtual memory, file systems, and networking.

PROJECTS

- **Distributed Systems**: Recreated popular distributed systems technologies, MapReduce and Raft, using Golang after self-teaching course material provided by MIT's graduate distributed systems course 6.5840. Acquired understanding of distributed systems principles in tangent with existing knowledge of computer systems to reproduce MapReduce and Raft.
- **PintOS**: Gained extensive knowledge of components in x86-64 operating systems through designing the components of the PintOS model operating system. Designed the virtual memory management, process priority scheduling, and file system.

PROGRAMMING SKILLS

- **Languages**: C++/C, Python, Golang, Java, SQL
- **Technologies**: Linux, Shell Scripting, Snowflake, Dbt, Airflow, Git, CI/CD, Docker, Looker, Terraform
- **Relevant Coursework**: Distributed Systems, Algorithms, Operating Systems, Data Structures, Machine Learning