

Harrick Cheong

harrickcheong@gmail.com
harrick.me linkedin github

Education

University of Toronto

Graduated: Aug. 2023

Honours Bachelor of Science, Statistics Co-op (Statistical Machine Learning and Data Mining Stream)

Experience

Loblaw Companies Limited | Data Engineer

Sept. 2023 – Present

- Designed and implemented a workflow tool that profiles warehouse tables to identify outdated store-level data and automates backup and cleanup in Cloud Storage and BigQuery with Airflow and Datastore
- Migrated data lineage from Cloud SQL to Datastore with Java and created test cases using JUnit and Mockito
- Developed a monitoring service for cloud file ingestion with Composer and Python, optimizing pipeline reliability by automating notifications to support teams of missing files or delayed files
- Utilized Apache Beam to setup initial phase of file tracing in batch data pipelines and added metadata of raw files
- Configured pipelines with Apache Airflow to ingest data from CSV, JSON, and parquet files into BigQuery tables

PepsiCo | Data Science Intern

May 2022 – Aug. 2022

- Developed an end-to-end dashboard tool for sales team to derive key performance indicators from past campaigns, using Python and Pandas in Azure Databricks and visualized in Microsoft Power BI
- Designed and optimized a recommendation model to identify most valuable stores and products, facilitating targeted market expansion, obtained from statistical analysis on consumption data, using PySpark and Spark SQL

Royal Canin | Data Science Intern

May 2021 – Aug. 2021

- Constructed generalized linear models with glm package to obtain insight on the efficacy of marketing promotions
- Documented E2E flow of inventory allocation to customers to gain an overview of the distribution model

Certifications

Google Cloud Platform | Professional Data Engineer

Issued: Dec. 2024

- Expiration Date: Dec 30, 2026

Projects

Obelisk | React, PostgreSQL, Prisma, NextJS

June 2023

- A web app used for social media that features personalized feeds, tag-based searching, and interactive discussion

PrimeCart | Java, Firebase, Android Studio

Feb. 2022

- An android app used for online grocery delivery, enabling real-time database updates for customers and owners

Local Network Database | JavaScript, MySQL, Flask

Apr. 2021

- A webserver with session management capabilities, allowing users to create and update their profile page

Extracurriculars

Software for Love | Software Developer

May. 2022 – Dec. 2022

- Implemented a SSO login and session component for volunteers coordinating events using React and NextJS
- Designed a PostgreSQL database using UML diagrams for relational mapping, and setup Prisma for ORM capabilities

Technical Skills

Languages: Python, Java, SQL, R, JavaScript

Frameworks: NextJS, Android Studio, Flask

Developer Tools: Airflow, Git, Google Cloud Platform, Firebase, Jira, Confluence

Libraries: pandas, numpy, great-expectations, google-cloud, tidyverse, openpyxl