# Matt D'Souza

✉ matt.dsouza@edu.uwaterloo.ca
🌐 mattdsouza.com
�containing DSouzaM

## Skills

- Fluent in Java, Python, Scala, C++, Ruby, SQL
- Data engineering experience using Hive, Presto, MongoDB, Redis, Hadoop, and Spark
- Infrastructure experience using AWS, Chef, Prometheus and Google Analytics

## Experience

**May–Aug 2018** — **Software Engineering Intern** | *Snowflake Computing* | San Mateo, California

- Optimized the SQL compiler to re-route large quantities of bind variables, protecting the database engine from heavy load and out of memory errors
- Implemented compiler and client support for multiple SQL statements in a single query
- Improved instrumentation of SQL clients to discover performance bottlenecks and user workflows

**Sep–Dec 2017** — **Relevancy Engineering Intern** | *Wish* | Toronto, Ontario

- Reduced the time to build product image similarity graph from 20 hours to 30 minutes
- Developed a versioning scheme for Redis to enable consistent data updates across all Redis nodes
- Implemented backup, restore, and monitoring of graph databases using AWS and Prometheus
- Created a system to monitor changes in the performance of machine learning models
- Wrote Luigi pipelines and cron jobs in Python and SQL to generate data for product features

**Jan–Apr 2017** — **Data Platform Engineering Intern** | *Shopify* | Toronto, Ontario

- Built and maintained connectors to extract data from Google Cloud and other third-party APIs
- Reduced overlap between incremental extraction jobs by 95% after using Spark to investigate late-arriving data from historical logs
- Automated environment setup and project build process to improve developer efficiency

**May–Aug 2016** — **Software Engineering Intern** | *Veeva Systems* | Toronto, Ontario

- Built a Spring service to asynchronously send backend metrics to Google Analytics
- Spearheaded and took ownership of rewriting of the application's breadcrumb navigation system

## Projects

**2017** — **Language Analysis** | ⌗ DSouzaM/LanguageAnalysis
Comparison of GitHub projects by programming language

- Built a Scala wrapper for the GitHub API to build a data set, using Slick to interface with MySQL
- Implemented tokenization and keyword extraction algorithms to generate features from READMEs
- Experimented with clustering to group repositories and look for programming language trends

## Education

**2015–2020** — **Candidate for Bachelor of Software Engineering, 4.0 GPA** | *University of Waterloo*

- Undergraduate Research Assistant (2016): Experimented with Bazel and Java 9 compatibility for an annotation-based type-checking framework

## Interests

- Court & beach volleyball
- Functional programming, programming languages
- Acoustic guitar, electric keyboard