

Matt D'Souza

3A Software Engineering

✉ matt.dsouza@edu.uwaterloo.ca

🌐 mattdsouza.com

🐙 [DSouzaM](https://github.com/DSouzaM)

Skills

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

Experience

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

- Wrote Luigi pipelines and cron jobs in Python and SQL to generate data essential to product features; reduced the runtime of an existing job from over 20 hours to less than 30 minutes
- Developed a versioning scheme for Redis to enable consistent data updates across all Redis nodes
- Using AWS and Prometheus, implemented backup, restore, and monitoring of graph databases
- Created a daily performance monitor for machine learning models; resolved a crucial bug to double one model's performance

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

- Maintained and improved ETL systems written in JRuby, Scala, Java, and Python
- Reduced overlap between incremental extraction jobs by 95% after using Spark to investigate late-arriving data from historical logs
- Built new connectors to incrementally extract data from Google Cloud and third-party APIs
- Improved developer efficiency by automating environment setup and project build process

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

- Developed enhancements and fixes for single-page web app using Java, SQL, and HTML/CSS/JS
- Built a Spring service to asynchronously send backend metrics to Google Analytics
- Identified shortcomings with front-end features and took ownership of developing solutions

Projects

2017 **Language Analysis** | 🐙 [DSouzaM/LanguageAnalysis](https://github.com/DSouzaM/LanguageAnalysis)

Comparison of GitHub projects by programming language

- Built a Scala wrapper for the GitHub API to scrape data, using Slick to interface with MySQL
- Implemented a simple tokenizer and used TextRank to extract keywords from READMEs
- Used K-means clustering to group similar repositories and look for language trends

2016 **Employment Stats** | 🌐 mattdsouza.com/EmploymentStats

Visualization of UWaterloo co-op employment rates

- Developed a Python scraper to incrementally extract information about co-op employment
- Designed a web app to visualize and explore time-series data with Flask and Highcharts

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