

Joel Therrien (MSc, BSc)

(604)-700-6824, joel@joeltherrien.ca
<https://www.linkedin.com/in/joel-therrien-280ab7a8/>

Joel is a Data Scientist who combines statistics and software engineering to create unique and powerful solutions for complex problems. He love statistics because it equips him to learn about the world; and loves software engineering because he can build unique and novel solutions. This drive of learning & self-sufficiency bleeds into his work where he excels at learning new skills, and into his hobbies where he self-hosts web services and experiments with making ice cream.

Technical Skills

- Designed, tested, and deployed a custom random forest model for valuing homes in British Columbia. Initially was prototyped in Julia, later implemented in Rust and R.
- Built an internal business intelligence (BI) tool in Python, Typescript, and React where an interactive one-file HTML site could be emailed to clients.
- Built an ETL process to migrate production data into an analytics database. Customized Looker, a BI tool, on top of this.
- Reverse-engineered, documented, and improved numerous legacy processes.
- Wrote and designed SQL reporting.
- Successfully explained advanced statistical concepts to clients and stakeholders.
- Actively maintain a Linux server self-hosting several web services.

Work Experience

March 2024 - Present

Imperva

Data Scientist

I work on the Advanced Bot Protection team, where I research new techniques to distinguish bot traffic from human traffic in order to develop new mitigations.

February 2021 - March 2024

Landcor

Data Scientist

Context – Landcor uses models to automatically provide valuations on homes in BC using data from BC Assessment, as well as fulfilling custom data requests. Their customers are typically banks & realtors. During my time there I've:

- Designed, implemented, and deployed a new machine learning home valuation model, which significantly outperformed the previous model in terms of accuracy, flexibility, coverage, and computational time. This is Landcor's core product!
 - Prototyped in Julia, implemented in Rust & R.
- Used Python & Typescript to create internal websites & tools to allow non-IT team members to access data.
- Used Python + SQL to automate customer-facing reporting, creating interactive HTML files with the relevant summary statistics embedded for offline access. Used Python, React, Typescript, and Plotly.
- Reverse-engineered undocumented legacy processes, improving, documenting, and rewriting them into modern languages while fixing previous bugs.

- Designed several functional prototype product ideas, usually involving Python for the backend server and Typescript & React for the front-end, using a custom built REST API.
- Mentored, guided, and taught a junior Data Scientist.

September 2022 - December 2022 Klue Contract Work

Developed a Python tool that integrated into BigQuery that allowed staff to interactively track customer retention over time using a variety of metrics. Worked closely with a product manager there as we built out requirements. Worked concurrently with my job at Landcor.

January 2019 - August 2019 SFU Library Graduate R Data Peer

Worked for the SFU library to provide R software support to other graduate students for their research, as well as to create and teach workshops (ranging from 3 hour to 2 days) to teach introductions to R and Python.

November 2015 - August 2017 Grow Financial Data Engineer

Set up a data pipeline from Google Datastore on the Google Cloud Platform into a self-hosted PostgreSQL database. Configured and maintained Looker, a business intelligence tool, with all of our business definitions.

Designed and ran custom SQL reporting, and assisted the back-end developers with new features in the Java codebase.

May 2014 - September 2014 UBC Okanagan Student Researcher

Developed a model that connected breeding bird populations across the United States to land-cover zones (using the National Land Cover Database). Bayesian methods (JAGS) was used to fit the model.

Responsibilities included collecting the data, cleaning the data, putting the data into a database format, learning R, and experimenting with various modeling techniques.

Education

2017 - 2019 SFU Master of Science in Statistics

Studied both statistical theory and practical applications. For my master's project I wrote a software package in R & Java to train competing risks random forests (a machine learning model) for large datasets, which was then used to analyze a dataset of loan prepayment. This was of personal interest as it solved a similar problem previously encountered at Grow Financial.

2011 - 2015 UBC Okanagan Bachelor of Science Honours in Mathematics

Degree earned at the UBC Okanagan campus, majoring in math with a concentration in statistics. Transcript average was 96.5%.

Volunteering

2010 - 2015 Kelowna Gospel Mission Volunteer Cook

Once a week helped prepare and serve meals to Kelowna's homeless community.