Brendon Phillips

LinkedIn | GitHub | Publications

Location: Toronto ON M5A 3X2
Email: brendoncphillips@gmail.com | Mobile: 437-268-1484

DATA SCIENTIST, DATA ANALYST

Over 7 years of experience in **data science**, **modelling** and predictive **analytics** of big data with **Python**, **R**, **C++** and **SQL**.

Flexible scientist proficient with data **visualization**, dashboards, **machine learning** and project management with **Git**.

TECHNICAL SKILLS

Python	:	BeautifulSoup, matplotlib, geopandas, nltk, numpy, pandas, prophet, scikit-learn, seaborn, statsmodels, tensorflow, word2vec
R	:	comprehenr, data.table, dplyr, ggplot2, knitr, orca, plotly, quantreg, reshape2, shiny, stringr, trend
SQL	:	MySQL, BigQuery, DuckDB, PostgreSQL, dbt, ETL/ELT, PL/SQL
Other tools	:	C/C++, Julia, Rust, Shell scripting, Fortran, regex, XML, Microsoft Excel, Google Sheets, Git
Productivity	:	Microsoft Powerpoint, Microsoft Word, Google Slides, Google Docs, LaTeX, Tableau

PROJECTS

Modelling Various COVID-19 Reopening Scenarios in Ontario Schools *C++, R, Google Sheets, Git* Source Code

- Built and parametrised an agent-based disease transmission model for different scenarios using **C++**,
- Gathered insights from roughly 243 scenario combinations in 50GB of data using **R** and **Python**,
- Results were published, covered in the media and presented to the members of two regional school boards.

Data Modelling and Time Series Analysis for Predicting Epidemics *C++, R, Python, Julia, LaTeX* Web Page

- Built a large infectious disease model with social networking and opinion dynamics using **C++** with **SNAP** and **Eigen**,
- Gathered statistical metrics from various fields to **analyze epidemic curves** (ecology, physics, network theory, etc),
- Automated the **time series analysis** responsible for generating early warning signals of epidemics using **R** (368GB),
- Analysis and results published in **peer-reviewed articles 1** , **2** , and **3** and presented at conferences.

Effects of Geographical Remoteness on COVID-19 Infection Spread *R, Python, Julia, Git* Source Code

- Gathered data sets from Statistics Canada, public health units, dashboards and APIs using **R** and **SQL**,
- Cleaned, blended performed regression analyses on the data set using **R** and **Julia**.

PROFESSIONAL EXPERIENCE

Data Analyst (contract) 06/2023-Present
Hospital for Sick Children *Hybrid - Toronto ON, Canada*

Working on a research team to clean and repair medical trial data for statistical analysis and optimize the code review process.

Senior Software Engineer 02/2022 – 02/2023
Liquid Analytics *Remote – Toronto ON, Canada*

Designed and delivered optimized data mining algorithms that:

- assemble complex SQL queries and use NLP and other machine learning techniques in Python to recover features from a large volume of unstructured data,
- achieve average parsing times of under 1 millisecond per entity profile over 97% of data without external API calls,
- score the degrees of similarity between entity profiles using Python and upload scores to BigQuery.

Wrote documentation, status reports and updates, and gave regular non-technical presentations to business stakeholders.

Postdoctoral Fellow (contract) 05/2021 – 02/2022
York University *Hybrid - York, ON, Canada*

Worked individually and in cross-functional (some international) teams to model and simulate disease transmission dynamics, and evaluate the effectiveness of interventions including vaccination with behavioural and social responses with regression and other statistical analyses.

Postdoctoral Fellow (contract) 01/2021 – 04/2021
University of British Columbia *Remote - Vancouver BC, Canada*

Co-developed software with ApexRMS to use their SyncroSim application to compare and chain together COVID-19 models written in R, Python and C#.

Researcher 09/2015 – 11/2020
University of Waterloo *On-site - Waterloo ON, Canada*
Proposed new techniques useful for predicting epidemics and public health crises in a computational disease/behaviour model in R, C++ and Julia using:

- continuous testing algorithms,
- time series analysis,
- machine learning (NLP, regression, cluster analysis, etc),
- extensive statistical and data analysis.

Course Coordinator & Instructor (contract) 09/2018 – 12/2018
University of Waterloo *On-site – Waterloo ON, Canada*

Taught advanced mathematical skills and covered material on acoustics, differential equations, Fourier analysis and sound engineering; organized guest lectures and demonstrations; and supervised term-long research projects.

PEER-REVIEWED PUBLICATIONS

- Network structural metrics as early warning signals of widespread vaccine refusal in social-epidemiological networks. J. Theor. Biol. 531, 110881 (2021).
- Model-based projections for COVID-19 outbreak size and student-days lost to closure in Ontario childcare centres and primary schools. Sci Rep 11, 6402 (2021).
- Spatial early warning signals of social and epidemiological tipping points in a coupled behaviour-disease network. Sci Rep 10, 7611 (2020).
- Early warning indicators of epidemics on a coupled behaviour-disease model with vaccine hesitance and incomplete data. J Dyn. Games, 2023, 10(1): 49-86.

Additional Training

- **Google Data Analytics Professional Certificate** 04/2023

Education

Ph.D. Applied Mathematics Waterloo ON, Canada
University of Waterloo *09/2015 – 12/2020*

M.Sc. Mathematics London ON, Canada
Western University *09/2014 – 08/2015*

B.Sc. Mathematics Cave Hill, St. Michael, Barbados
University of the West Indies *09/2011 – 07/2014*

Languages

English
Native proficiency

German
Intermediate proficiency