

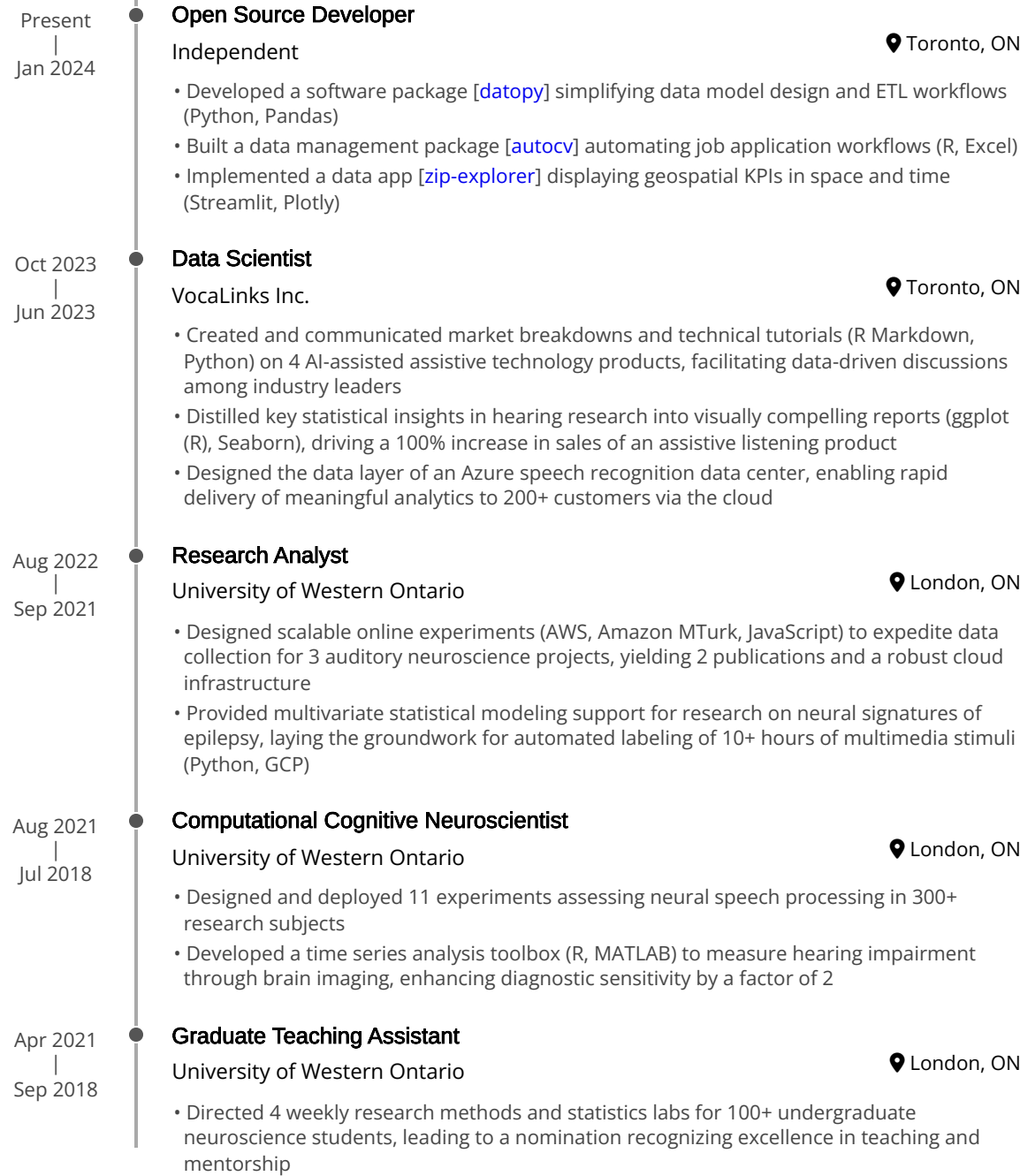
Matthew Bain

Data Scientist | Tableau, Databricks | Neuroscience Researcher Turned AI Developer

Data Scientist with 5+ years experience and certifications in Tableau and Databricks. My work spans neuroscience and assistive tech, with a focus on robust, interpretable ML. I help organizations understand their data and automate the hard stuff. I am seeking a role involving analytics and pipeline development.



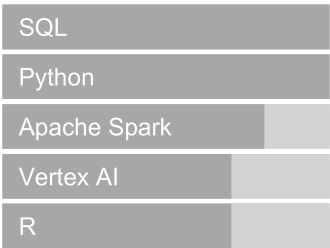
Professional Experience



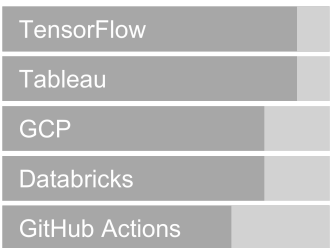
Contact

Ontario
[Personal Website](#)
[GitHub](#)
[LinkedIn](#)

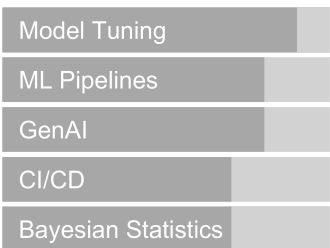
Core tools



Platforms



Key competencies



Last updated on 2025-04-14.

Apr 2018
|
Apr 2015

●

Head of Finance

Lorne Park Landscaping

• Co-founded and managed payroll, client invoicing, and government remittances for a successful partnership specializing in property maintenance, servicing 80+ clients and leading a team of 5 employees

📍 Mississauga, ON



Education

Dec 2023
|
Sep 2021

●

BSc (Honours), Mathematics and Statistics

McMaster University

📍 Hamilton, ON

Aug 2020
|
Sep 2018

●

MSc, Neuroscience [[Thesis](#)]

University of Western Ontario

📍 London, ON

Apr 2018
|
Sep 2014

●

BSc (Honours), Neuroscience

McMaster University

📍 Hamilton, ON



Certifications

Mar 2027
|
Mar 2025

●

Databricks Certified Data Engineer Associate [[Certificate](#)]

Databricks

Feb 2027
|
Feb 2025

●

Tableau Certified Data Analyst [[Certificate](#)]

Tableau

Expected
May
2025

●

Professional Machine Learning Engineer

Google Cloud Platform (GCP)

Dec 2024

●

Machine Learning Specialization [[Certificate](#)]

Stanford University

Selected Projects

Present
|
Jan 2024

●

datopy (Data Management Python Package) [[Documentation](#)]

• Implemented and maintain a package for handling unstructured data, providing a simple interface for data modeling, extraction, validation, and building ETL pipelines (Pandas, PyTest, Pydantic, GitHub Actions)

Present
|
Apr 2024

●

mlvizz (Machine Learning Data Visualization Package)

• Build Python interfaces for efficient, intuitive model selection, tuning, inspection, and ML pipelines through modular, object-oriented designs with built-in data validation (Scikit-learn, TensorFlow, ArviZ, PyMC, SciPy)

Present
|
Apr 2024

●

statvizz (Statistical Data Visualization Package)

• Build Python extensions unifying Pandas data summarization with Seaborn statistical plotting functionality for a unified, intuitive, fully transparent statistical plotting interface (Pandas, Seaborn, SciPy, Matplotlib)

Present
|
Jul 2024

●

mathvizz (Mathematical Visualization Package)

• Build Python interfaces for exploring the geometry of machine learning math, including linear maps, derivatives, statistical distributions, and series approximations (SymPy, SciPy, Seaborn, Matplotlib)

Feb 2025		
Feb 2025		
Dec 2024		
Nov 2024		
Jul 2024		
May 2024		
	●	One-stop Superstore Dashboard [Tableau] <ul style="list-style-type: none"> Designed an interactive dashboard displaying sales in a convenient, unified view, with key KPIs and sales logs visualized in both space and time (SQL, Excel)
	●	Zip Explorer (Geospatial Data Analysis Streamlit App) [Web App] <ul style="list-style-type: none"> Implemented a data application with interactive visualizations of geospatial data in space and time, across various KPIs and at state and county levels (Streamlit, Pandas, Plotly, Altair, GeoPandas)
	●	autocv (Automated CV-generation R Package) [Documentation] <ul style="list-style-type: none"> Built a data management package automating job application workflows using validated data stores and macro-integrated spreadsheets, resulting in an efficient, extensible workflow (dplyr (R), pkgdown, Excel)