

Matthew Bain, MSc

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.



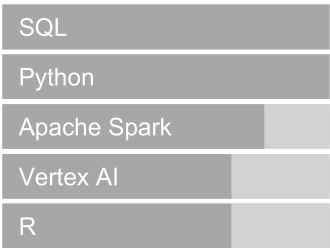
Professional Experience

Present Jan 2024	Open Source Developer Independent • Developed a software package [] simplifying data model design and ETL workflows (Python, Pandas) • Built a data management package [] automating job application workflows (R, Excel) • Implemented a data app [] displaying geospatial KPIs in space and time (Streamlit, Plotly)	Toronto, ON
Oct 2023 Jun 2023	Data Scientist VocaLinks Inc. • Created and communicated market breakdowns and technical tutorials (R Markdown, Python) on 4 AI-assisted assistive technology products, facilitating data-driven discussions among industry leaders • Distilled key statistical insights in hearing research into visually compelling reports (ggplot (R), Seaborn), driving a 100% increase in sales of an assistive listening product • Designed the data layer of an Azure speech recognition data center, enabling rapid delivery of meaningful analytics to 200+ customers via the cloud	Toronto, ON
Aug 2022 Sep 2021	Research Analyst University of Western Ontario • Designed scalable online experiments (AWS, Amazon MTurk, JavaScript) to expedite data collection for 3 auditory neuroscience projects, yielding 2 publications and a robust cloud infrastructure • Provided multivariate statistical modeling support for research on neural signatures of epilepsy, laying the groundwork for automated labeling of 10+ hours of multimedia stimuli (Python, GCP)	London, ON
Aug 2021 Jul 2018	Computational Cognitive Neuroscientist University of Western Ontario • Designed and deployed 11 experiments assessing neural speech processing in 300+ research subjects • Developed a time series analysis toolbox (R, MATLAB) to measure hearing impairment through brain imaging, enhancing diagnostic sensitivity by a factor of 2	London, ON
Apr 2021 Sep 2018	Graduate Teaching Assistant University of Western Ontario • Directed 4 weekly research methods and statistics labs for 100+ undergraduate neuroscience students, leading to a nomination recognizing excellence in teaching and mentorship	London, ON

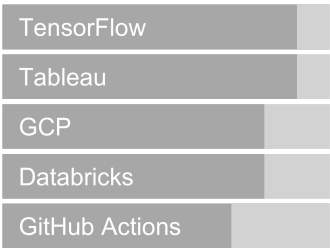
Contact

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

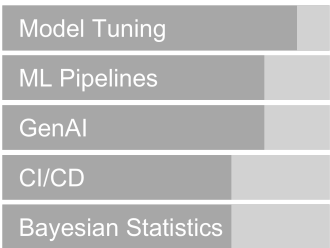
Core tools



Platforms



Key competencies



Last updated on 2025-04-10.

Apr 2018 Apr 2015	<div>●</div> <div> Head of Finance Lorne Park Landscaping <div> <div>📍 Mississauga, ON</div> <ul style="list-style-type: none"> • 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 </div> </div>
---------------------------	--



Education

Dec 2023 Sep 2021	<div>●</div> <div> BSc (Honours), Mathematics and Statistics McMaster University <div> <div>📍 Hamilton, ON</div> </div> </div>
Aug 2020 Sep 2018	<div>●</div> <div> MSc, Neuroscience [Thesis] University of Western Ontario <div> <div>📍 London, ON</div> </div> </div>
Apr 2018 Sep 2014	<div>●</div> <div> BSc (Honours), Neuroscience McMaster University <div> <div>📍 Hamilton, ON</div> </div> </div>



Certifications

Mar 2027 Mar 2025	<div>●</div> <div> Databricks Certified Data Engineer Associate [Certificate] Databricks </div>
Feb 2027 Feb 2025	<div>●</div> <div> Tableau Certified Data Analyst [Certificate] Tableau </div>
Expected May 2025	<div>●</div> <div> Professional Machine Learning Engineer Google Cloud Platform (GCP) </div>
Expected May 2025	<div>●</div> <div> Machine Learning Specialization [Certificate] Stanford University </div>
Expected May 2025	<div>●</div> <div> SQL Basics for Data Science Specialization UC Davis </div>



Selected Projects

Present Jan 2024	<div>●</div> <div> datopy (Data Management Python Package) [Documentation] <ul style="list-style-type: none"> • 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) </div>
Present Apr 2024	<div>●</div> <div> mlvizz (Machine Learning Data Visualization Package) <ul style="list-style-type: none"> • 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) </div>
Present Apr 2024	<div>●</div> <div> statvizz (Statistical Data Visualization Package) <ul style="list-style-type: none"> • 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) </div>

