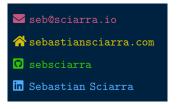
Sebastian Sciarra, PhD

Data Scientist



並 Education

PhD | Industrial-Organizational **Psychology**

University of Guelph Sep. 2018-May 2023

MSc | Cognitive Psychology McMaster University

Sep. 2016-June 2018

Honours BSc | Psychology, Neuroscience & Behaviour

McMaster University Sep. 2012-June 2016

Selected whitepapers

The Game of Supervised **Machine Learning:** Understanding the Setup, Players, and Rules

Published 10 August 2023

The Expectation-Maximization Algorithm: A Method for **Modelling Mixtures of Distributions**

Published 28 April 2023

Probability, Likelihood, and **Maximum Likelihood Estimation**

Published 19 March 2023

Profile

Passionate about coding, machine learning, and statistics. Completed my PhD dissertation at the intersection of these fields to address a practical problem in Industrial-Organizational psychology and received the 2022/2023 Canadian Psychological Association Certificate of Academic Excellence for this work. In my dissertation, I coded and ran Monte Carlo simulations on an AWS instance to evaluate the performance of nonlinear longitudinal models. Writes white papers on machine learning topics at sebastiansciarra.com.

‡ Skills

Coding languages

- Python
- R
- SQL
- LATEX
- Javascript
- HTML
- CSS

- IDEs/platforms Technical skills PyCharm
 - Data visualization (ggplot2, plotnine)
 - Data cleaning (tidyverse, pandas, numpy)
 - Machine learning (e.g., regularized regression, decision trees, random forests, mixture models)
 - Statistics (e.g., latent variable models, factor analysis, multilevel modelling)

Employment experience

Teaching Assistant

University of Guelph

Sep. 2018-May 2023

Sep. 2020-Apr. 2021

- Created R scripts for assignments and taught labs for the following courses in measurement and statistics:
 - PSYC 3290 (Conducting Statistical Analyses in Psychology)
 - PSYC 3250 (Psychological Measurement)
 - PSYC 6060 (Research Design and Statistics)

RStudio

MySQL

AWS

Git

- PSYC 6380 (Psychological Applications of Multivariate Analysis)
- Taught a variety of topics in methods and statistics (e.g., regression with continuous and categorical [i.e., ANOVA] variables, p values, p hacking, hierarchical linear modelling, factor analysis, latent variable modelling, etc.)

Graduate Research Assistant

University of Guelph (Part-Time)

■ Used R to clean data, compute descriptive statistics, and run regression analyses (with categorical and/or continuous variables) for organizational data on turnover, downsizing, and growth

Consultant Sep. 2020-Dec. 2020

Geosyntec (Part-Time)

- Worked with a team of graduate students to improve the interview procedure
- Developed customized recommendations to structure the interview procedure so that adverse hiring outcomes were reduced and skills were more rigorously evaluated

Consultant Jan. 2020-Apr. 2020

Schema App (Part-Time)

- Worked with a team of graduate students to improve the onboarding of new employees
- Synthesized customized recommendations by using literatures on realistic job previews, goal setting, and

Data science experience

smltheory

Aug. 2023

Python package

- Package contains nine modules and 30 functions
- Functions within package simulate data sets and demonstrate propositions of supervised machinelearning propositions (e.g., bias-variance tradeoff, excess risk decomposition)

cobaltResume May 2023

R Package

- Automates generation of resumes and cover letters within RStudio
- A template and class file were created (~700 lines of LATEXcode) to specify a styling template that draws inspiration from the cobalt theme in the RStudio IDE
- R functions were created to easily generate resume entries and merge resume and cover letters into one PDF file

August 2023 Citizenship: Canadian

<u>m</u> Education

PhD | Industrial-Organizational **Psychology**

University of Guelph Sep. 2018-May 2023

MSc | Cognitive Psychology McMaster University Sep. 2016-June 2018

Honours BSc | Psychology, **Neuroscience & Behaviour**

McMaster University Sep. 2012-June 2016

Selected whitepapers

The Game of Supervised Machine Learning: Understanding the Setup, Players, and Rules

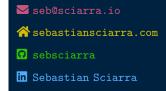
Published 10 August 2023

The Expectation-Maximization Algorithm: A Method for **Modelling Mixtures of Distributions**

Published 28 April 2023

Probability, Likelihood, and **Maximum Likelihood Estimation**

Published 19 March 2023



sebastiansciarra.com Mar. 2023

Personal website

• Used HTML, JavaScript, and CSS to create a personal website for writing white papers

- White papers focus on statistics, machine learning, and coding by explaining technical details, providing demonstrations, and conducting simulation experiments
- White papers use code from a variety of languages to explain content. As an example, my post titled "Coding and Visualizing the Expectation-Maximization Algorithm" used R, Python, and CSS code

guelphdown R Package Mar. 2023

- Created an R package that automates the generation of theses according to the University of Guelph formatting requirements
- A template and class file were created (~1400 lines of LATEXcode) to specify formattings for the preamble, body, references, and appendices
- An example of the formatting can be seen in my thesis

nonlinSimsAnalysis Mar. 2022

R Package

- Package contains 105 functions
- Functions automate the cleaning, analysis, and visualization of large data sets (e.g., 40 000+ rows) for my doctoral
- The creation of several different types of tables and figures were automated by this package

nonlinSims Jan. 2022

R Package

- Package contains 30 functions
- Functions run the simulation experiments of my doctoral dissertation
- The peformance of nonlinear longitudinal models are evaluated (e.g., structured latent growth curve models) are evaluated under several conditions

Learning SQL Mar. 2021

Project

- Went through 16 of 18 chapters from Alan Beaulieu's Learning SQL
- Topics include filtering, querying multiple tables, sets, grouping and aggregates, subqueries, joins, transactions