

# Jake Barnabe *Data Scientist*

✉ jake.barnabe@queensu.ca

☎ (613) 898-7113

🌐 LinkedIn

🐙 GitHub

🌐 Website

## PROFILE

Creative, problem-solving Data Scientist interested in predictive machine learning and statistical modelling. Experienced building data pipelines. Skilled with Python, R and SQL. Thrives working as part of a team. Strong written and oral communicator.

## EDUCATION

### Master of Data Science

University of British Columbia

09/2023 – 06/2024

Vancouver, Canada

- Relevant courses: Statistics, Data Wrangling, Data Visualization, Predictive Modelling, Machine Learning, Databases, Algorithms

### Bachelor of Arts (Honours) - Applied Economics

Queen's University

09/2016 – 04/2021

Kingston, Canada

## PROJECTS

### Predicting Player Performance

05/2024 – 06/2024

UBC MDS Capstone Project with Vancouver Whitecaps FC

- Constructed an end-to-end pipeline using Python, R, and SQL that features a series of linear mixed-effects models to predict first-team player performance from match data
- Improved prediction accuracy by 19.5%
- Voted "Best Capstone Presentation" by the class

### Dreamhouse - Real Estate Dashboard [🔗](#)

04/2024 – 04/2024

UBC MDS Group Project

- Built a dashboard using Dash to present an interactive visualization of housing prices across the top 50 most populous cities in the United States of America.
- The dashboard enables users to filter and explore the housing market across various dimensions including state, city, house type, square footage, number of bedrooms and bathrooms, and more.

### Predicting high-potential FIFA players using individual performance data [🔗](#)

11/2023 – 12/2023

UBC MDS Group Project

- Constructed a classification model using an RBF SVM classifier algorithm which uses FIFA22 player attribute ratings to predict players' potential and then classify them into target classes "Low", "Medium", "Good", and "Great".

### Quanteda - Python package [🔗](#)

01/2024 – 02/2024

UBC MDS Group Project

- Created a Python package of functions that generate visualizations and metrics to simplify exploratory data analysis of quantitative financial data.

## PROFESSIONAL EXPERIENCE

### Analyst

01/2022 – 08/2023

Statistics Canada - Canadian Center for Education Statistics

Ottawa, Canada

- Coordinated data collection for the 2022 Elementary-Secondary Education Survey (ESES) by corresponding with provincial ministries of education.
- Led data validation to identify errors and missing values by comparing the new data to trends from the past 5 years.
- Conducted exploratory data analysis and identified significant impacts of the Covid-19 pandemic on elementary and secondary enrolment statistics.
- Co-authored the ESES survey release article [🔗](#) to communicate analysis results in a manner that is easily understood by Canadians.
- Maintained the Public Service Employee Survey dashboard in Power BI to visualize data for clients and upper management.

### Project Coordinator

08/2020 – 12/2021

RCMP - Federal Policing

Ottawa, Canada

- Conducted data validation of records of all staffed and vacant positions within Federal Policing.
- Optimized the data validation process and revised standard operating procedures documents to reflect new procedure as well as conducted training sessions with team members to teach them the new optimized method

- Using Insperity OrgPlus, visualized the validated records of Federal Policing's positional data for upper management to facilitate strategic decision-making.

## **SKILLS**

---

### **Coding Languages**

Python, R, SQL

### **Software & Tools**

GitHub, Spark, TensorFlow, AWS, MongoDB, Snowflake, Docker

### **Statistics**

Predictive and statistical modelling, A/B testing, Hypothesis testing, Time-series analysis, Bayesian inference, Data visualization

### **Machine Learning**

- Supervised learning (Regression, Classification, Neural networks, Gradient descent models)
- Unsupervised learning (Clustering, Dimensionality reduction, Natural Language Processing, etc)