Jake Barnabe *Data Scientist*

✉ jake.barnabe@queensu.ca 📞 (613) 898-7113 in LinkedIn GitHub 🔗 Website

PROFILE

Creative, problem-solving Data Scientist interested in 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
UBC MDS Capstone Project with Vancouver Whitecaps FC
05/2024 – 06/2024
- Constructed a data pipeline using Python, R, and SQL that queries new data from Snowflake, preprocesses it, feeds it to a series of linear mixed-effects models to predict first-team player performance, and saves the predictions back to Snowflake as a new table.

Dreamhouse - Real Estate Dashboard 🔗
UBC MDS Group Project
04/2024 – 04/2024
- 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 🔗
UBC MDS Group Project
11/2023 – 12/2023
- Constructed a classification model using an RBF SVM classifier algorithm which uses FIFA22 player attribute ratings to classify players' potential with target classes "Low", "Medium", "Good", and "Great".

Quanteda - Python package 🔗
UBC MDS Group Project
01/2024 – 02/2024
- Created a Python package of functions that generate charts and metrics to simplify exploratory data analysis of quantitative financial data.

PROFESSIONAL EXPERIENCE

Analyst
Statistics Canada - Canadian Center for Education Statistics
01/2022 – 08/2023
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
RCMP - Federal Policing
08/2020 – 12/2021
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, 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, etc)