**ANDREW SIQUEIRA**

andsiq@yahoo.ca | 905-921-2139 | [https://www.linkedin.com/in/andrew-siqueira](about:blank) | <https://github.com/AndSiq97> | <https://andsiq97.github.io/>

# PROFESSIONAL PROFILE: Data analytics professional with a strong background in descriptive analysis, predictive analytics, machine learning, deep learning, and AI. Very adept at managing relationships at all levels. Ability to provide support and work in partnership as well as work independently with minimum supervision.

# PROFESSIONAL EDUCATION AND PROFICIENCIES:

**Bachelor of Applied Science**

**Major in Applied Statistics, Minor in Economics and Mathematics/CS**

*University of Toronto*

* **Relevant Coursework:** Regression Analysis, Machine Learning, Financial Economics, Data Structures and Analysis, Experimental Design, Linear Algebra, Java Software Design/Development, Object Oriented Programming, Natural Language Processing and Text Analytics, Time Series Forecasting
* **Data Analysis and Research:** Former data analyst at Bank of Canada. Executed research projects from start to finish including data cleanup, visualization, statistical analysis, and report building
* **Communication:** Experience presenting the findings of my projects in both formal and informal settings
* **Coding:** Professional proficiency in R, Python, SQL, Hadoop/Hive/Spark, Azure/Docker Java, Power BI, Tableau

# WORK EXPERIENCE

**Data Analyst (Contract)** July 2020 – December 2020

*Bank of Canada*

* + Utilized internal and publicly available data to identify key areas in which the BoC could improve operational efficiency
    - Analyzed Canadian COVID data to provide recommendations on WFH and return-to-work policies
    - Developed a visualization tool in Power BI to track the financial performance of the BoC during the pandemic
  + Created reports using R and Python and SQL to depict trends in COVID-19 cases as well as departmental productivity

**Machine Learning Projects** November 2018 - Present

*Data Science Projects using Datasets from Kaggle and Online*

* Apply statistical and regression/classification machine learning techniques to minimize forecast error in predicting a dataset (in Python and R along with SQL)
* Focus on Feature Engineering, Model Training and Model Evaluation in addition to data munging/wrangling
* Utilize Natural Language Processing and text analytics for sentiment analysis and topic modeling in text-based documents/datasets
* Incorporate deep learning time series forecasting models to forecast future values in the dataset
* Utilized various Python libraries such as Pandas, NumPy, SciPy, Matplotlib, Stats Models, Scikit-learn, Sklearn, TensorFlow/Keras/PyTorch, NLTK, SpaCy
* Identify right algorithms and statistical techniques and provided justification

***Heart Failure Classification***

* Applied classification machine learning algorithms to classify whether patients have heart disease given risk factors such as age, gender, cholesterol, fasting blood sugar, etc.
* Conducted plots and statistical metrics visualize relationship between each factor and its effect on heart failure

***Stock Price/Return Time Series Forecasting***

* Applied time series statistical model (ARIMA) and deep learning algorithms (LSTM) to forecast price/returns of any given stock on Yahoo Finance
* Utilized statistical plots and metrics and applied optimized portfolio diversification

***Apple Twitter Sentiment Classification***

* Applied NLP techniques using machine learning algorithms and transformers to classify Apple tweets as either positive, neutral, or negative sentiment
* Cleaned and preprocessed tweets for analysis

**Research Assistant** October 2019 – February 2020

*University of Toronto* *– Centre of Student Engagement*

* Analyzed data (both quantitative and qualitative) from Ontario high school students to develop a recommendation for a program that would ease transition of students to the post-secondary level
* Applied regression techniques and hypothesis testing (in R) to determine which group of students would benefit most from the program based on socioeconomic, geographic, and cultural characteristics

# ADDITIONAL PROJECTS:

**Pedagogic inquiry into active learning spaces** May 2020 – August 2020

*Research Opportunity Program,* *University of Toronto*

* + Cleaned anonymized survey data from students to investigate effects of active learning spaces on performance and learning experience
  + Conducted statistical analysis in Python/R using regression, documenting findings in an academic research paper

**Quantitative Analyst** January 2020 – April 2020

*Algo Trade (UTM Society of Algorithmic Modelling),* *University of Toronto*

* Collaborated with my team to build a fully automated stock trading algorithm in R and Python
* Using publicly available financial data, developed a model that achieved a success rate of greater than 50% in predicting stock prices

**Evaluation on Paper Airplanes** March 2019

*Experimental Design Research Experiment, University of Toronto*

* Designed a research experiment testing the impact of airplane design and texture on flying distance
* Created summary and visual reports using R to communicate findings to student peers

# CERTIFICATIONS:

**IBM Data Science Professional Certificate** In Progress

*Coursera*

# INTERESTS:

Data Analysis and Data Science, Finance and Banking, Fitness and Music, Mathematics, Statistics, and Programming

References Available Upon Request