

# Ruhi Mahendra

 (647) 393-5890  [ruhima@hotmail.com](mailto:ruhima@hotmail.com)  <https://ruhimahendra24.github.io/>  <https://www.linkedin.com/in/ruhi/>

## DATA ENGINEERING PROJECTS

### Academic Analytics Consulting Project with Quebec Health

- Designed an ELT pipeline using Azure and Databricks to process semi-structured data from 7+ data sources to be used for a predictive employee churn model implemented on the Quebec Health system

### Academic Analytics Consulting Project with BDC

- Automated the extraction and transformation of data from Azure DevOps Kanban Boards using Microsoft PowerShell and Python in order to develop productivity insights from Kanban board tickets

### Strava Data Pipeline

- Developed a job orchestration pipeline using Apache Airflow to extract activities from the Strava API, transform the data and load the data on an SQLite Database

### Football Player Valuation Model Deployment

- Deployed the football player valuation prediction model on Heroku with Docker through the creation of a Python Flask App
- Implemented a CI/CD pipeline using GitHub Actions to monitor changes within the main branch of the repository to update the model on Heroku

## PROFESSIONAL EXPERIENCE

### Data Analyst

Vancouver, Canada

Taking Root

May 2021- July 2022

- Built an automated reporting and verification infrastructure through Cron jobs using Python and SQL to process tree sampling data provided by clients to improve data quality and saving program managers 5 hours a week through report automation
- Provided optimal sampling solutions to clients to increase carbon removal precision estimates by 30%
- Prepared an easy-to-use interactive R Shiny dashboard with key insights for program managers to visualize and understand client data trends

### Research Assistant

Waterloo, Canada

University of Waterloo

April 2021 – February 2022

- Worked under the supervision of Dr. Kalyaanamoorthy contributing to the development of a Machine Learning model for repurposing therapeutics to target viral diseases
- Developed Bash and Python scripts to increase virtual screening efficiency by 100-fold to collect ligand-protein complex data for use in the classification model
- Conducted an analysis for an external research project which lead to a published scientific paper identifying gene expression trends based on their functions within biological samples using R

### Biodiversity Data Scientist

Ottawa, Canada

Agriculture and Agri-Food Canada

May 2020 – August 2020

- Developed R scripts to perform text mining to reduce redundancy between databases containing data on the flora of North America resulting in one robust relational database
- Designed and implemented interactive R shiny dashboards used by higher management to visualize data trends on the flora of North America

### Scientific Data Analyst

Toronto, Canada

Environment and Climate Change Canada

January 2020 – April 2020

- Wrote Shell and Python scripts to process large satellite-derived remote sensing datasets and utilize numerical models to predict global albedo shifts impacted through climate change
- Created visually appealing geographical maps using Python's Matplotlib and satellite data to explain the variability in landcover due to the rise in atmospheric temperatures

## EDUCATION

**McGill University**, Desautels Faculty of Management

Montreal, Canada

Master of Management in Analytics

July 2022 – May 2023

**University of Waterloo**

Waterloo, Canada

Bachelor of Science (Co-op), Specialization in Bioinformatics

September 2016 – April 2021

## TECHNICAL SKILLS

- Tools & Software:** SQL, NoSQL, Neo4j, Python, R, Power BI, Bash, Git, Azure, Command Line, Docker, Airflow, Databricks
- Skills:** Data Visualization, Data Wrangling, Data Structures and Algorithms, Web Scraping, Object-Oriented programming, ETL, ELT, Machine Learning, Data Science, Data Modelling, Data Analytics, Statistics, Data Warehousing, Data extraction with APIs, CI/CD Pipeline