






MUJTABA SIDDIQUI

 mujtabasiddiqui.github.io
 mujtaba.siddiqui@uwaterloo.ca
 linkedin.com/in/mujtaba-siddiqui/
 mujtabasiddiqui
 647-537-4921

Honours Geomatics | [University of Waterloo](#)

SKILLS

Languages Python, SQL, Scala, R, Java, C, Javascript, Go, Git, Bash/Shell
Data & Databases Hadoop, Spark, Hive, Nifi, Kafka, MySQL, PostgreSQL, SQL Server, Oracle, BigQuery, MongoDB
Tools & Technologies Pandas, Tensorflow, Airflow, SiSense, Tableau, Docker, AWS, GCP, Azure, Matillion, Terraform

EXPERIENCE

Data Engineer | ecobee Toronto, ON | Jan 2021 - Present

- Independently created a parameterized ingestion pipeline using **Matillion** saving **20+ hours** from a new pipeline build process
- Developed **ETL** and **ELT** data pipelines for **SQL** and **NoSQL** sources using tools such as Matillion, **Google Dataflow**, and Python
- Managed cloud infrastructure using **Terraform** and scheduling using **Airflow**

Data Engineer | Manulife Waterloo, ON | May 2020 - Aug 2020

- Developed a **Mainframe** ingestion framework with **Spark** and **Nifi**, saving **\$200,000** yearly from third party alternatives
- Independently wrote new **data pipelines** to improve accuracy of key business metrics used by executives, marketing and advanced analytics teams in building machine learning models, strategy planning, and various other **ETL** intensive tasks
- Led a team of co-ops in creating a web app to automate the co-op hiring process saving **5+ hours** a week for hiring committee
- Created a system to monitor **HIVE** process to mask sensitive PII preventing exposure of **1M+ records** of data
- Configured **CI/CD** pipelines using Pytest for **unit testing** and Jenkins to automate versioning and deployment of new packages

Data Analyst | York Region Newmarket, ON | Sept 2019 - Dec 2019

- Created a program to automate georeferencing with **computer vision** and artificial intelligence using Python with **OpenCV**, and **ArcGIS**, saving **15 hours** from the permit approval process
- Developed various **PowerBI** dashboards with **SQL Server** integration entry to visualize data for enterprise-wide consumption
- Automated workflows using **Python**, such as writing a script to create feature for the Region's online Analytics Gallery

PROJECTS

SpotUp 

Python, Javascript, Plotly, Dash, Flask, React, Pandas, Scikit-Learn

SpotUp is a **web app** aiming to bring **advanced analytics** and **spatial analysis** to any NBA fan with limitless customizability, expanding the use of spatial analysis in the NBA through shot charts in **dashboards** built from user input used to pull data from **API endpoints**

L.I.S.A | UW Starter Hacks 2020 

Python, OpenCV, Tensorflow, RaspberryPi

A proof of concept designed to aid caregivers while looking after their patients using **computer vision** and **speech recognition** to monitor their movements and notify of potential hazards when the caregiver is not present

Imperium Solar | UW Energy Hacks 2019 Winner 

HTML, CSS, Javascript, Python, ArcGIS

A conceptual start-up created to lower the cost of solar energy by connecting businesses with investors and providing a ROI report using **deep learning** to perform **time series forecasting** on satellite and electricity usage data, aiming to show financial incentive to environmentally sustainable practices

EDUCATION

University of Waterloo | 3B B.E.S Honours Geomatics with Minor in Computing | Dean's Honours List

- Relevant Courses: Data Types and Structures, Multivariate Statistics, Spatial Analysis, Databases, Linear Algebra, Calculus
- External Courses: NLP with PyTorch (Pluralsight), Python and Spark for Big Data (Udemy)