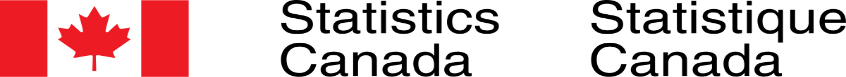
**Freelance Software Developer** *Nov 2023 – Present (8 months)*

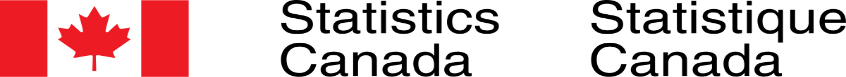
* **Developed Front-End and Full Stack Website for Clients**
  + Developed websites for clients, whether that be front-end or full-stack applications primarily using **Javascript, Node, Html, CSS, SQL, MongoDB**



**Statistics Canada**

**Software Developer** *March 2021 – Jan 2023 (1 year 11 months)*

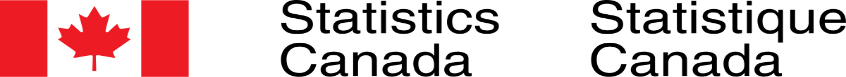
* **Semi-structured Excel Workbook to structured CSV files**
  + Developed an end-to-end continuous deployment (CD) pipeline to automate converting Semi-Structured Excel Workbooks to structured CSV files by leveraging **Python**, **Docker**, and **ArgoWorkflows** to streamline the process
* **Human Resources Robotic Process Automation (RPA) Application**
  + Used **Automation Anywhere** software to interface with **Microsoft Access** Forms, databases, emails, etc. to automate the assistance of the HR registration processes
  + **Trained** clients to use RPA software through creating video tutorials on practical use cases
  + **Presented** findings to directors and the Assistant Chief Statistician of Canada
* **Machine Learning Auto-Encoding PSIS Project**
  + Implemented scripts in **Jupyter Notebook** using the **fastText ML Library** to create text classification models on business surveys to predict a 6-digit statistical code



**Statistics Canada**

**Data Scientist (Co-op)** *June 2020 – Dec 2020 (7 months)*

* **Proof of Concept Machine Learning Alternative to SQL-Rule Based Record Linkage**
  + Used **PySpark** and **MLlib** to implement a Locality-Sensitive Hashing based model for record linkage across multiple Excel datasets



**Statistics Canada**

**Data Engineer (Co-op)** *Jan* *2019 – Dec 2019 (1 year)*

* **Extract, Transform, Load (ETL) Census Data using Apache Ecosystem**
  + Used Income Census Data to find the median income per person for each gender, categorized by their provinces
  + Introduced Apache Ecosystem to reduce department costs by introducing open source technologies
  + Optimized parallelizing Datasets **55% faster** than the existing process
  + Leveraged **Apache Spark** (PySpark) to ingest CSV data using **Apache Hive** within the **Hadoop file system** (HDFS), and **Apache Zeppelin** for data visualization (in particular, **Visualization Drill Downs** and interactive charts)