

Francis Tao

8688 Hazelbridge Way, Richmond, BC V6X0R6 | P: 778-681-6828 | E: ty199695@gmail.com | [Linkedin](#) | [Github](#)

Professional Summary

A diligent and ambitious professional with expertise in data analytics, passion in machine learning, exceptional communication skills and multilingual ability. A vast skill set of statistical inferences, time series analysis, predictive analysis and data visualization..

Skills

Database: PostgreSQL, MySQL, T-SQL, SSMS

Data Visualization: Tableau, Power BI, Matplotlib, Seaborn

Others: Excel, Access, Outlook, Slack, Jira

Statistical Programming: Python, SAS, RStudio

Data Analytics and Science: Pandas, Numpy, Sklearn

Work History

RecycleSmart

Richmond, BC

Data Coordinator

11/2021 – Present

- Pull the data from SQL Server, generate waste diversion reports monthly in Excel and give insights of recycling data based on calculations and Power BI dashboard.
- Clean and transform over 6100 rows of customers/vendors' addresses with Pandas, build Python validation tools to automate the process by utilizing CanadaPost/BingMap API, computing textual similarity using Levenshtein Distance with 74.8% accuracy, converting kmz file to json and creating region/neighborhood codes based on Polygon geospatial analysis.
- Write sql query and utilize Power Query/Power BI to transform the data and create visualizations for various projects including SmartTrash pickup predictions and Vendor invoice gap amount/difference.
- Reduce the monthly waste diversion number manual check workload by an average of 47.2% by building a Python tool and utilizing libraries such as re, pandas, dateutil.parser, datetime and calendar.
- Automate over 3000+ vendor purchase contract lines' unique key, unit price, region code and product id in Python(lambda, pandas)

ehsAI

Vancouver, BC

Data Labelling Analyst

02/2021 – 10/2021

- Process EHS documents with citations, language labelling, splitting, etc.
- Develop internal QA survey, evaluate the employee performance by trending prediction based on visualization through Power BI/Excel Query, Pivot, Charts, Python Pandas, Matplotlib, Numpy, Seaborn and improve the task accuracy by 10% within 6mo.

Providence Health Care

Vancouver, BC

Data Quality Clerk

01/2020 – 06/2020

- Assist end-users with registration-related processes, liaise with various departments and escalate problems.
- Extract, sort and aggregate data using SQL and verify them from client registry/clinical information reports to maintain data integrity and to ensure compliance with high accuracy of above 85%.

ECCO Shoes Canada

Richmond, BC

Brand Ambassador

11/2019 – 10/2021

- Greet customers, help to find the wanted shoes, fix the displays and resolve the return and exchange issues.
- Generate more than \$1500 in weekly sales with successful marketing, sales and customer relations approaches.

Quest Food Exchange

Vancouver, BC

Office Assistant Volunteer

05/2019 – 07/2019

- Excel data entry and update referral clients' information into spreadsheet over 2800 rows for 2000+ organizations.

Education

The University of British Columbia

Vancouver, BC

Bachelor of Science, Statistics

05/2019

Dean's List 2016 – 2017; Member of Golden Key Society: Top 15% in faculty; Graduated with 82%

Personal Projects

Bike Share Demand Prediction ([Link](#)) – 0.4075 RMSLE(301/3243 on Kaggle, top 9%); Data manipulation with pandas, numpy and datetime; Histogram, scatterplot and bars visualization with matplotlib; Modelling and training with Train_Test_Split, LinearRegression, RandomForestRegressor and XgBoost.

Heart Disease Prediction ([Link](#)) – GridSearchCV Logistic Regression with 88.5% accuracy; 303 rows of data; Barplot, scatterplot and heatmap by matplotlib and seaborn; Modelling and training with LogisticRegression, KNeighborsClassifier, RandomForestClassifier and DecisionTreeClassifier; Hypertuning using RandomizeSearchVC and GridSearchCV

Certification

- (1). [MCSA: SQL 2016 Database Development](#)
- (2). [Tableau Desktop Specialist](#)
- (3). [SAS Certified Specialist: Base Programming](#)
- (4). [Data Science Professional Certificate](#)