KEVIN DANG

☏ (647) 909-9810 | ✉ dang.kevin@protonmail.com | 🏠 dang-kevin.github.io | ○ dang-kevin | in dang-kevin

EDUCATION

University of Toronto

MASTER OF SCIENCE, BIOSTATISTICS *2021 - 2022*

- Emphasis in Artificial Intelligence and Data Science Option

HONOURS BACHELOR OF SCIENCE, APPLIED STATISTICS & MATHEMATICS *2016 - 2020*

EXPERIENCE

St. Michael's Hospital *Toronto, ON*

JUNIOR BIOSTATISTICIAN *Oct. 2021 - Present*

- Building **linear mixed-effects regression** and piecewise linear regression models while incorporating **machine learning** methods such as bootstrapping, cross-validation and LASSO regularization
- Using **Multivariate Imputation by Chained Equations** to create multiple imputations for multivariate missing data, recovering **20%** of lost information
- Working with the **TARGetKids!** research group to study COVID-19 public health measure adherence in the Greater Toronto Area, expected to culminate in a **manuscript**

University of Toronto *Toronto, ON*

TEACHING ASSISTANT *Sep. 2021 - Present*

- **Leading tutorials** and office hours for **graduate** & undergraduate courses in statistics & data science to reinforce students' understanding of analytical concepts and programming skills in **R & Python**
- Managing the **MarkUs** grading platform and writing **test cases** with **pytest** to streamline the grading process for TAs

Export Development Canada *Ottawa, ON*

BUSINESS INTELLIGENCE ANALYST *Jun. 2020 - Dec. 2021*

- Implemented **LASSO-penalized logistic regression** and **random forest classifiers** with the **caret** package in **R** to predict client retention and conversion rates for EDC insurance products within a specified timeframe
- Utilized the **tidyverse** to perform **exploratory data analysis** on customer churn and conversion rates, created data visualizations and tables for **PowerPoint** presentations
- Developed and maintained dashboards in **PowerBI** and **Excel** for senior management and clients, automated multiple steps by using macros, functions (vlookup, index & match) and running **SAS EG** projects & process flows

GlobalData Plc *Toronto, ON*

MEDICAL DATA ANALYST *Sep. 2020 - Sep. 2021*

- **Automated** multiple steps for a global medical device pricing tool via an **ETL pipeline** with **Pandas dataframes** in **Python** and **pivot tables & macros** in **Excel**, reducing analyst hours by more than **50%**
- Contributed a new methodology to the team and incorporated concepts involving data cleaning and **winsorization**, **doubling production speed** of numerous medical device markets in the pricing tool
- Performed **data quality checks** on the **Medical Intelligence Centre** and core databases to ensure clients have accurate & up-to-date information, wrote articles & press releases to capture the attention of news outlets

University of Toronto *Toronto, ON*

DATA ANALYST, STATISTICAL CONSULTANT, RESEARCH ASSISTANT *May 2018 - Jun. 2020*

- Performed data cleaning, wrangling and visualization on millions of rows of data using **dplry, tidyr, ggplot2** and wrote reports in **R Markdown** explaining analytical concepts for non-technical clients
- Analyzed differences between groups using **ANOVA** and **t-tests** to investigate pay equity among agriculture workers and used **linear regression** models to answer business problems regarding a consulting firm's online platform
- Accessed the **Scopus RESTful API** using **pybliometrics** (Python-based API-Wrapper), built a **data pipeline** in Jupyter Notebook using **Pandas dataframes** and extracted datasets after reshaping and joining dataframes

SKILLS

Languages   R, Python, SQL, SAS, Stata, VBA, HTML, CSS
Software/Tools   PowerBI, Tableau, Excel, GitHub, Git, LaTeX, ArcGIS