KEVIN DANG

📞 (647) 909-9810 | ✉ dang.kevin@outlook.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 | APPLIED HEALTH RESEARCH CENTRE *Oct. 2021 - Present*

- Building **linear mixed-effects regression** and piecewise linear regression models while incorporating machine learning methods such as **bootstrapping, cross-validation** and **regularization with LASSO**
- 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 amongst parents and children in Ontario

Export Development Canada *Ottawa, ON*

BUSINESS INTELLIGENCE ANALYST | BUSINESS PERFORMANCE ANALYTICS *Jun. 2020 - Present*

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

GlobalData Plc *Toronto, ON*

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

- **Automated** multiple steps for a global medical device pricing tool, including data cleaning & **winsorization** in **Python**, **pivot tables & macros** in Excel, reducing analyst hours by more than **50%**
- Utilized **Pandas** in **Jupyter Notebook** to create an **ETL pipeline** for multiple medical device markets in the pricing tool, increased production speed and contributed a new methodology to the team
- Performed frequent healthchecks on **Medical Intelligence Centre** and core databases to ensure clients have correct & up-to-date information, wrote articles & press releases to promote GlobalData - capturing attention of news outlets

University of Toronto *Toronto, ON*

TEACHING ASSISTANT | DEPARTMENT OF STATISTICAL SCIENCES *Sep. 2021 - Present*

- **Leading tutorials** and office hours to solidify students' understanding of analytical concepts and R programming
- Creating **beamer presentations** using **LaTeX** and maintaining personal website with course content

DATA ANALYST & STATISTICAL CONSULTANT | CIRHR *Sep. 2019 - 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 relating to a consulting firm's online platform

RESEARCH ASSISTANT | ROTMAN SCHOOL OF MANAGEMENT *May 2018 - Aug. 2019*

- Obtained access to the **Scopus RESTful API** using **pybliometrics** (a Python-based API-Wrapper), built a **data pipeline** in Jupyter Notebook using **Pandas dataframes** and extracted datasets after reshaping and joining dataframes
- Manipulated, merged and extracted information from large datasets in **Stata** to create new datasets

SKILLS

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