KEVIN DANG

☎ (647) 909-9810 | ✉ dang.kevin@proton.me | ⌂ 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

College of Nurses of Ontario *Toronto, ON*

DATA ANALYST *Apr. 2022 - Present*

- Produced public **HTML reports** containing statistical analyses of nurses in Ontario using **R Markdown**
- Extracted large datasets through **Microsoft SQL Server** to increase loading times significantly, then performed data cleaning, wrangling and visualization with **tidyverse** to streamline the workflow

St. Michael's Hospital *Toronto, ON*

STATISTICIAN *Oct. 2021 - Present*

- Built linear mixed-effects regression and piecewise linear regression models and incorporated **machine learning** methods such as bootstrapping, cross-validation and LASSO regularization
- Utilized **Multivariate Imputation by Chained Equations** to create multiple imputations for multivariate missing data, recovering **20%** of lost information
- Worked with the **TARGet Kids!** 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 - Aug. 2022*

- **Led 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**
- Managed 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
- Developed and maintained dashboards in **PowerBI** and **Excel** for stakeholders, 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** production of 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%**
- 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*

- Analyzed pay equity among agriculture workers groups with **ANOVA** and **t-tests** and answered business problems for a consulting firm's online platform with **linear regression** models
- Accessed the **Scopus RESTful API** using **pybliometrics** (Python-based API-Wrapper), built a **data pipeline** in **Jupyter Notebook** using Pandas dataframes, and exported datasets after reshaping and joining dataframes

SKILLS

Languages   Python, R, SQL, SAS, Stata, HTML, CSS, VBA
Software/Tools   PowerBI, Tableau, Excel, Google BigQuery, Azure DevOps, GitHub, Git