

Kevin Dang

☎ (647) 909-9810

✉ dang.kevin@outlook.com

🌐 dang-kevin.github.io

📷 dang-kevin

🌐 dang-kevin

🏛️ EDUCATION

UNIVERSITY OF TORONTO
Honours Bachelor of Science
Applied Statistics Specialist
Mathematics Minor
2016 - Apr. 2020 (expected)

COURSES

- Computer Programming
- Design & Analysis of Experiments
- Geographic Information & Mapping
- Machine Learning
- Methods for Multivariate Data
- Methods of Applied Statistics
- Methods of Data Analysis
- Statistical Computation
- Statistical Consultation & Collaboration
- Time Series Analysis
- Theory of Statistical Practice

⚙️ SKILLS

PROGRAMMING

- Python • R • SQL • Stata
- Matlab • HTML • CSS

OTHER

- ArcGIS • LaTeX • MS Access
- MS Excel • MS Office

🏆 AWARDS

UofT Entrance Scholarship
• 92%+ average

AP National Scholar
• 98th Percentile

Mathematics Award
• Highest overall average across all senior math courses

★ INTERESTS

- Volunteering
- Eco-Team Executive
 - Student Council Representative
 - Statistics Study Group Leader
 - Tennis Canada (Fundraising)

- Hobbies
- Fishing • Piano • Table Tennis
 - Trading • Video games • Weightlifting

📁 EXPERIENCE

DATA ANALYST | *University of Toronto*
Sep 2019 - Present | Toronto, ON

- Spearheaded a new data analysis project about agriculture and worker rights with Dr. Greg Distelhorst and a team of researchers
- Performed **data cleaning** and **data wrangling** on millions of rows of data using the **dplyr**, **tidyr** and **lubridate** libraries
- Visualized relationships between variables and time trends with **ggplot2** with specific details geared towards non-technical audiences

STATISTICAL CONSULTANT | *University of Toronto*
Sep 2019 - Present | Toronto, ON

- Currently working on a **data analysis project** for a management consulting firm to answer their business problems
- Previously completed a statistical analysis project to determine the effect of **auditory distraction** on cognitive flexibility for university students

RESEARCH ASSISTANT | *Rotman School of Management*
May 2018 - Aug 2019 | Toronto, ON

- Worked under the supervision of Dr. Chris Liu with a team of graduate students on projects about scientific publications and careers
- Queried scientific databases using Python-based API-Wrappers, worked with **dataframes** using **Pandas** and exported data into csv files
- Merged and manipulated large datasets with **Stata**, extracted desired information, cleaned data and generated new variables
- Used **BeautifulSoup** for web scraping and exported data into Excel to improve efficiency in creating new datasets

📁 PROJECTS

PREDICTING CREDIT CARD APPROVALS | *Python (NumPy, Pandas, Scikit-learn)*

- Cleaned data by filling in missing values with **mean imputation** or most frequent observations, used **label encoding** to convert non-numeric data to numeric format and split data into train and test sets
- Scaled features, fit a **logistic regression classifier** with **84% accuracy** and performed a **grid search** of the model parameters to improve the model's ability to predict credit card approvals

RADIUS OF THE EARTH | *Python (NumPy, SciPy, Matplotlib)*

- Collected data on gravitational strength using a gravimeter, manipulated data with **NumPy** and fit **linear regression** models to the data using **SciPy**
- Plotted models and performed **chi-squared analysis** on the models to check for goodness of fit and estimated the Earth's Radius to **within 30 kilometres**

DEGREES THAT PAY YOU BACK | *R (dplyr, ggplot2)*

- Cleaned data and used **elbow**, **silhouette**, and **gap statistic** methods to determine the **optimal number of clusters** to be used in applying the **k-means algorithm** to the data
- Visualized the starting and median salaries with **ggplot2**, plotted each cluster individually to look for patterns in career growth for certain majors

SUPER BOWL, T.V. & HALFTIME SHOWS | *PostgreSQL*

- Investigated tables containing Super Bowl, television, and halftime show data by **writing queries** containing various **filter and join clauses**
- Explored questions involving game outcomes, T.V. viewership, ad costs and musician performances