Kevin Dang

(647) 909-9810 | dang.kevin@outlook.com | dang-kevin | dang-kevin | dang-kevin

EDUCATION

UNIVERSITY OF TORONTO

Honours Bachelor of Science

Applied Statistics Specialist
Mathematics Minor

2016 - Present

COURSES

- Computer Programming
- Design & Analysis of Experiments
- Linear Algebra
- Machine Learning
- Methods of Data Analysis
- Multivariate Calculus
- Partial Differential Equations
- Statistical Practice
- Statistical Theory

MOOCs

- MIT: Computer Science using Python
- Stanford: Machine Learning
- UofT: Learn to Program

SKILLS

PROGRAMMING

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

OTHER

• 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
• Board games • Fishing • Investing
• Piano • Soccer • Table Tennis

EXPERIENCE

ROTMAN SCHOOL OF MANAGEMENT | *Research Assistant*

May 2018 - Present | Toronto, ON

- Worked under the supervision of **Dr. Christopher 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

MOSAIC SALES SOLUTIONS | *Brand Ambassador*

Oct 2015 – Sep 2017 | Toronto, ON

- Promoted different types of brands for numerous companies and **consistently increased product sales** by more than the daily target of **25%**
- Wrote reports containing information regarding customer interaction, sales made, products purchased, demo issues and conflict resolution

PROJECTS

PREDICTING CREDIT CARD APPROVALS | *Python*

- 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** using **scikit-learn** 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*

- Collected data on gravitational strength using a gravimeter, manipulated data with **NumPy** and fit **linear regression** models to the data using **SciPy**
- Plotted models using **matplotlib.pyplot**, 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 (Jupyter Notebook)*

- Cleaned data using **dplyr** 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

DUNGENESS CRAB GROWTH | *R Markdown*

- Summarized the data in a table and plotted a **boxplot**, then used a **t-test** to compare the mean shell sizes of the two groups of crabs
- Plotted a **normal quantile plot** and **histogram** using **ggplot2** to check the normality condition and performed an F-test to compare two variances to check the constant variance assumption

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

- 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