Department of Statistics & Data Science Colloquium Series
Fall 2019

Speaker: Dr. Andriy Koval
Title: Visualizing Logistic Regression
Date: Friday, November 8th, 2019
Time: 11:00AM – 12:00PM
Location: Technology Commons II, Room 222

Abstract:

Visualising results of statistical modeling is a key component of data science workflow. Statistical graphs often is the best means to explain and promote research findings. However, in order to find that one graph that tells the story worth sharing, we sometimes have to try out and sift through many data visualizations. How should we approach such a task? What can we do to make it easier from both production and evaluation perspectives? This talk will demonstrate a reproducible graphing system designed for the IPDLN-2018 hackathon. The system evaluates synthetic socioeconomic and mortality data with logistic regression. The data was prepared for the hackathon by Statistic Canada and represents Canadian population. First, I will introduce a visualisation technique that uses color to create a meaningful expectations from the results of a logistic regression. Then I will discuss the workflow of the project that implements this graphing system. I will conclude by building the case to prefer reproducible workflows with version control over computational notebooks (e.g. Jupyter, R Notebook).

Speaker Bio:

Andriy Koval, Ph.D. is a data scientist with a background in quantitative methods and interests in data-driven models of human aging. He was a Health System Impact Fellow with Observatory for Population and Public Health (Centre for Disease Control of British Columbia, UBC) and a research fellow with the Vancouver Island Health Authority. Andriy's works centers around developing tools for reproducible research with R and GitHub as key components. Presently, Andriy's work focuses on developing statistical methods for analysing transactional data extracted from the electronic health records (EHR). His current interests include design of information displays with R, literate programming, statistical modelling in general, and longitudinal analysis in particular.

Notice to students: Please sign in at the event!