PETER TEA

(+1) 613 501 1108 ◇ ptea@sfu.ca ◇ petertea.netlify.app ◇ Vancouver, Canada

TECHNICAL STRENGTHS

Statistical Tools: R, Python, SQL, SAS, SPSS
Other Tools: Git, BASH, LaTex, Markdown.

EDUCATION

Simon Fraser University Sep 2019 - Current
Master of Science in Statistics
Awards: Natural Sciences and Engineering Research Council of Canada Graduate Scholarship (NSERC CGS-M), BC Graduate Scholarship, Dean's Graduate Fellowship

University of Ottawa Sep 2014 - Apr 2019
Honours Bachelor of Science - Biostatistics (*Summa Cum Laude*)
Award: NSERC Undergraduate Student Research Award (USRA)

STATISTICS WORK EXPERIENCE

Data Science Co-op - Aquatic Informatics May 2020 - Aug 2020

- Generated tools with *Python* and *Boto3* to wrangle customer data from an AWS bucket. Reported user effort and organizational costs with an interactive data visualization tool developed in *plotly-dash*—allowing managers to make informed data-driven decisions.
- Applied *anomaly detection* algorithms on customer time series data, and evaluated its performances. Recommended improvements based on observed user behaviours and customer qualitative feedback.

R Consultant - Research Commons Library Sep 2019 - Apr 2020

- Provided individual and group consultations to graduate level researchers on data wrangling, visualization and statistical analysis to advance ongoing research project deliverables.
- Lead *R* and *Git* workshops and assisted on *Python* and *BASH* workshops to effectively communicate programming concepts for beginners.

Data Science Intern - Canada Revenue Agency May 2019 - Aug 2019

- Generated SAS scripts to decipher patterns in tax-evasion behaviour in Canadian sub-populations.
- Organized large datasets into a consistent and analyzable format.
- Collaborated with international data scientists on Machine Learning and Econometric Model applications to complex data. Presented these findings to management to advise on future strategic plans.

Statistical Genetics Research Assistant - University of Ottawa May 2018 - Apr 2019

- Implemented novel data dimension reduction approaches in R. Scripts were then applied to analyze high volume Crohn's disease datasets to identify genetic risk factors.
- Coded BASH scripts to automate large scale simulations of genetic data to illustrate power of data dimension reduction techniques using a computer cluster.
- Presented a poster to a non-technical audience at an academic Health research conference attended by medical students.

Analyst - Transport Canada Jul 2018 - Apr 2019

- Improved decision-making by verifying concerns of stakeholders and upper management by supporting and cross-referencing unknown claims with added context and evidence.
- Maintained data integrity by meticulously fixing data entry errors.

DATA PROJECTS

MIT Sloan Sports Analytics Conference Hackathon Mar 2020

- Create visual representations of college basketball player tracking data with ggplot and gganimate.
- Analysed the relationship between player shot release angles and release velocities on the success of shot attempts.

A Dynamic Approach to modelling career All-NBA selection counts Oct 2019

- Applied Regression models to predict a player's All-NBA selection count at any stage of their career.
- Scraped and engineered meaningful features from raw NBA box-score data from the past 30+ years into a tidy, usable dataset.