# Junkyu Park

joonjunkyu.park@alum.utoronto.ca
https://www.linkedin.com/in/asdfzxcvjkl1/
https://joon3216.github.io/

## Education

**H.B.Sc.: statistics specialist and math minor** (GPA: 3.83) Jun 2019
University of Toronto

## Experience

**Methodologist** (MA-01) Jan 2020 - Present · 4 mos
Statistics Canada Ottawa, ON

- documented on various topics and ideas in statistics and computations
- involved in analysis and modeling process conducted by senior methodologists
- initiated a development of packages in R and Python

## Skills

- Data analysis
  - data manipulation: `dplyr`, `tidyr`, `pandas`
  - data visualization: `ggplot2`, `matplotlib.pyplot`, `seaborn`
  - regression analysis: R functions and packages, `statsmodels.api`
  - statistical computations: manually written functions, `numpy`, `scipy`
  - comprehension: able to find references and read documentations of external packages
- Technical
  - Markdown/R Markdown: able to form documentations
  - Git and git repositories:
    - able to use basic git commands
    - able to integrate dynamic badges supported by Travis CI and Codecov.io
  - Some knowledge of HTML and CSS: currently maintaining `joon3216.github.io`
  - SQL: able to join tables and write subqueries
  - Python and R
    - able to utilize external packages for modeling, computations, and visualiztions
    - able to develop a simple package and upload to the archives
- Languages: English, Korean

## Projects

**Package development** (work): `sprcr` package in R

- included the following functionalities to the package:
  - plotting kernel density estimates and empirical cdfs based on different colour and panel groups
  - smoothing out data points based on constraints and selected options
  - performing k-fold cross-validation in the smoothing process

**Modeling** (personal): credit card fraud detection

- built statistical models and provided analysis of data to detect fraudulent credit card use
- expected to achieve an accuracy of 97.105% and a true positive rate of 87.204% for every 6-hour worth of transaction data with the current model and the training set