**Selina WONG**

+65 8751 8702

Email: selinawj@msn.com | GitHub: <http://github.com/selinawj> |

Portfolio: <http://selina-wong.github.io/> | LinkedIn: <http://linkedin.com/in/selinawj>

**Work Experience**

**SRX (Streetsine Technology Group)**

**Junior Data Analyst,** Singapore

June 2019 - Present

* Deep learning-based Face detection using the YOLOv3 algorithm (Keras and Tensorflow) on filtering out poor quality listing photos
* Deployed a predictive model API for estimating rental property values with median percentage absolute error of an average of 5% with Python Hug & Postman
* Automated graphics in monthly reports with Python (Matplotlib, Seaborn & Pandas) and SQL, reducing production time by 90%
* Developed Tableau map visualization dashboard for listings around schools

**RentHop**

**Data Scientist Intern,** New York City

August 2017 - July 2018

* Led creation of a statistical multinomial logistic regression model for classifying scammers with 90% accuracy (R)
* Reported an alternative decision tree model with 88.7% accuracy
* Initiated and managed 2 viral data analytics projects picked up by over 160 global and nationwide news including [**New York Post**](https://nypost.com/2018/06/12/new-york-citys-rat-problem-is-getting-worse/)**, Gothamist, TimeOut, UK DailyMail,** [**Wall Street Journal**](https://www.wsj.com/articles/new-yorkers-are-complaining-more-about-rats-1528978480)**, Xin Hua &** [**Chicago Tribune**](http://www.chicagotribune.com/news/local/breaking/ct-met-chicago-rat-capital-20180719-story.html) etc (sparked an equivalence of ~80% growth in SEO efforts)
* Applied statistics and Natural Language Processing (NLP) on unigrams and bigrams for Naive Bayes text classification and feature prediction
* Automated geospatial distance calculations between linear streets and pairs of coordinates, and reverse-geocoding of geospatial coordinates to their respective neighborhoods
* Spearheaded the production of over 50 customizable and interactive geospatial data visualization, maps and infographics with open-sourced data

**Technical Projects**

**[Twitter News Tweets Analysis](https://github.com/selinawj/Twitter_Extractor_Scraper)**

* Automatically looped Json format files and parsed 27,500 lines of Twitter records into CSV (Python, Tweepy)
* Matched and retrieved Twitter users followers from a 26GB .Net database file of 41.7 million user profiles & 1.47 billion social relations into individual CSV records

[**New York City**](https://www.renthop.com/studies/nyc/rats-continue-to-thrive-in-new-york-city-especially-during-summer) **&** [**Chicago**](https://www.renthop.com/studies/chicago/chicago-rat-complaints-continue-to-grow)**’s Rat Complaints on the Rise**

* Developed data studies from open-sourced data, and interviewed with news reporters to provide further insights

**The Coca Cola Company Business Analytics Project**

* Scraped Twitter for tweets regarding Coca Cola in R
* Carried out sentimental analysis to evaluate effectiveness of marketing strategies

**Education**

**National University of Singapore**

Bachelor of Science (Honors) in Business Analytics

August 2015 – June 2019

* Freshmen Social Camp 2016 Programs and Logistics Organizing Committee
* Represented the School of Computing in Rag and Flag 2015 as a dancer

**New York University**

Entrepreneurship

August 2017 – December 2017

**Skills**

R, Python, SQL, HTML5/CSS, SAS, Tableau, Javascript, Java, MongoDB, Matlab, Julia, SEO, Google Analytics, Microsoft Excel, Photoshop, Git, Pandas, GeoJson, NLTK, Power BI

**Competitions & Achievements**

HackPrinceton Spring 2018 - [**DocentAR**](https://devpost.com/software/docentar)

HackMIT 2017 – [**Put a Face to It**](https://devpost.com/software/put-a-face-to-it)

NEC Laboratories Singapore Industrial Innovation Competition 2016 Merit Award – Advanced Queue System (AQS)

NEA 3 R Photography Competition 2012 Winner