

# Yaning (Sasha) Shi

647-674-6816 | [yaning.shi@hotmail.com](mailto:yaning.shi@hotmail.com) | [linkedin.com/in/yaning-shi/](https://www.linkedin.com/in/yaning-shi/) | [github.com/sashashi](https://github.com/sashashi)

## SUMMARY OF SKILLS

- 5 years of quantitative and qualitative data analysis and visualization experience, comfortable with unconventional data types
- Knowledgeable in research methodologies and familiar with data in multiple fields including health, life sciences, social sciences, and business
- Hands-on problem-solving experience with application of various statistical methods and machine learning models; Agile values advocate
- Technical Skills: Python, R, SQL, Tableau, Power BI, Apache Spark, IBM SPSS, Jupyter, Microsoft Excel, git, AWS, HTML; self-starter

## EDUCATION

**University of Toronto - Honours Bachelor of Science**  
Applied Statistics Specialist with Focus in Health Studies

Sep-Jun 2022

## EXPERIENCE

**Research Analyst** - University of Toronto Faculty of Law & Dept of English Oct-Apr 2022

- Conducted data analysis on real-time survey and historical textual data independently
- Evaluated research objectives, progress, and limitations in a collaborative environment
- Presented counter-intuitive statistical insights and their in-context implications to clients with written EDA, statistical analysis, and final reports

***The Birth of Modern Detective Stories: Popular Writing Styles***

Feb-Apr 2022

- Concluded unchanging overall sentiment and dominant cognitive mode as common characteristics in popular detective stories. Implemented Natural Language Processing, Textual Analysis, and Topic Modelling using Semisupervised Latent Dirichlet Allocation and Logistic Regression

***How Safe is Your Bicycle in Toronto?***

Feb-Mar 2021

- Composed a data-driven story that illustrated current situation and future trends of bicycle thefts in Toronto

***COVID-19 Cases and PM 2.5 Concentration in the UK: Demographic Statistics***

Nov-Dec 2020

- Discovered proportion of ethnic minorities as a confounding variable between ambient air pollution and COVID-19 Cases. Performed Spatial and Survival Analysis on Censored Data, with Bayesian Disease Mapping Models

***COVID-19 Mortality in Quebec***

Oct-Nov 2020

- Estimated COVID-related deaths in Quebec with pre-COVID Predictive Analytics using Bayesian Inference and Semi-parametric Time Model

***Tobacco Use in American Youth***

Sep-Oct 2020

- Determined demographic characteristics of youth tobacco in youth users with Generalized Linear Mixed Effects Models

***The Social Foundation of World Happiness***

Mar-Apr 2020

- Fitted a predictive model for happiness score using Principal Component Analysis

***Property Sale Price in Greater Toronto Area***

Nov-Dec 2019

- Selected predictive property specifications for sale price using Information Criteria