

Kris Hong

krishong.com

Birmingham, UK

07737 148818

kris.hong@hotmail.com

Education

- 2012–2018 **B.Sc. Statistics – Computer Science Track**
University of British Columbia, Vancouver, Canada
2:1 – Departmental Rick White Award for Excellence in Applied Statistics

Skills

- Highlight** 6 years of writing concise, efficient, and reproducible code to clean, link, explore, analyze, visualize, and extract insights from very large and messy datasets
- Computing** R (tidyverse), Python (pandas, keras), SQL Queries, RegEx, Git, Command Line
Building data pipelines, web scraping, spatiotemporal/audiovisual data competency
- Storytelling** Interactive Shiny apps, ggplot2, Matplotlib
R Markdown, Jupyter Notebook, Microsoft Office, G Suite, LaTeX

Experience

- 2018–2020 **Data Analyst, McGill University, Montreal, Canada**
- Lead analyst on five projects, first author on three of four publications
 - Built data pipelines, trained models, optimized code, wrote data protocols, organized fieldwork, managed an RStudio/Jupyter server with several users
 - Invited to talk at an International Conference on Machine Learning 2019 workshop
- 2019–2020 **Research Intern, Element AI, Montreal, Canada**
- Trained models in parallel on remote computing clusters using the command line
- 2016–2018 **Data Analyst, BC Centre for Disease Control, Vancouver, Canada**
- Lead analyst on four projects, first author on three of four publications
 - Conducted literature reviews, filed data requests, cleaned data, trained models, created visualizations and wrote reports for government stakeholders
- Jan–Jun 2018 **Teaching Assistant, University of British Columbia, Vancouver, Canada**
- Developed/taught workshops for data cleaning, exploration, and visualization in R
- May–Aug 2013 **Summer Intern, PROOF Centre, Vancouver, Canada**
- Developed an algorithm for improving the accuracy of microarray data

Selected Projects

Predicting short-term air pollution from videos of the street

- Built an efficient data pipeline in R for automatically linking nine data sources
- Trained a ConvNet to predict air pollution from street images and spectrograms

Visualizing spatial variations in summer temperature in Montreal

- Scraped temperature data from thousands of weather underground web pages
- Linked thousands of csv files from multiple sources into one dataset
- Built an interactive Shiny app to visualize millions of points across space and time

Publications

1. **Hong KY**, Pinheiro PO, Weichenthal S. Predicting spatiotemporal variations in outdoor ultrafine particle concentrations and size using street-level images and audio data. *Environ Int*, 2020.
2. Weichenthal S, Dons E, **Hong KY**, Pinheiro PO, Meysman FJR. Combining citizen science and deep learning for large-scale estimation of outdoor nitrogen dioxide concentrations. *Environ Res*, 2020.
3. Dill-McFarland KA, Koenig S, Mazel F, Oliver D, McEwan L, **Hong KY**, Hallam SJ. An integrated, modular approach to data science education in the life sciences. *Under Review*.
4. Gan WQ, Henderson SB, McKee G, Yuchi W, McLean KE, **Hong KY**, Auger N, Kosatsky T. Snowfall, temperature, and the risk of death from myocardial infarction: a case-crossover study. *Am J Epidemiol*, 2020.
5. **Hong KY**, Pinheiro PO, Minet L, Hatzopoulou M, Weichenthal S. Extending the spatial scale of land use regression models for ambient ultrafine particles using satellite images and deep convolutional neural networks. *Environ Res*, 2019.
6. **Hong KY**, Pinheiro PO, Weichenthal S. Learning global variations in outdoor PM_{2.5} concentrations with satellite images. *ICML AI4SG Workshop*, 2019.
7. **Hong KY**, Tsin PK, van den Bosch M, Brauer M, Henderson SB. Urban greenness extracted from pedestrian video and its relationship with surrounding air temperatures. *Urban For Urban Gree*, 2019.
8. **Hong KY**, King GH, Saraswat A, Henderson SB. Seasonal ambient particulate matter and population health outcomes among communities impacted by road dust in British Columbia, Canada. *JAPCA J Air Waste MA*, 2017.
9. **Hong KY**, Weichenthal S, Saraswat A, King GH, Brauer M, Henderson SB. Systematic identification and prioritization of communities impacted by residential woodsmoke in British Columbia, Canada. *Env Pollut*, 2017.

Other

Languages	English, Native speaker Mandarin Chinese, Spoken fluency
Right to work	Canadian Citizen Tier 5 Unrestricted Work Authorization Visa