YILIU CAO

+ 1 (249) 872-1796 | yiliu.cao@mail.utoronto.ca | https://github.com/yiliuc/Projects-Papers.git

Education

University of Toronto **Toronto, ON**
Double Major in Economics & Statistics, with a Mathematics Minor *Sep.2020 - Present*
- GPA: 3.79/4
- Awards: Dean's List Scholar *(2021 Summer, 2022 Winter, 2023 Winter)*
- Core Courses: Methods of Data Analysis, Design and Analysis of Experiments, Methods of Multivariate Data, Data Visualization, Time Series

University of Oxford **Oxford, UK**
Summer Abroad Program *Aug.2023-Sept.2023*
- GPA: 94/100
- Core Course: Big Data Tools and Applied Machine Learning for Economists

Professional Experiences

United Nations Global Platform for Big Data China Hub(UNBDC) **Hangzhou, China**
Big Data Research Intern *Jun.2023-Aug.2023*
- Used R to process and analyze large volumes of industrial data to deliver specialized insights, such as examining patterns of industry inflows and outflows in Zhejiang Province, identifying reasons for industry relocation and influencing policy.
- Participated in State Grid data collection training as well as learning the practical application of linear regression analysis of expected total energy consumption by type in 2030
- Engaged in three United Nations Statistics Division 'Beyond GDP Sprint' meetings to better understand global statistics frameworks and their implications for socioeconomic policies.

Project Experiences

ECO250: Did COVID-19 Defeat Donald Trump? An Electoral Analysis for the 2020 US Election
- Employed Python to examines how COVID and economic factors relate to Donald Trump's 2020 election loss
- Used big real-world data which combines the election data in 2016 and 2020, and ACS 5-year estimates in 2016 and 2020 (DP03 and DP05). The total number of observation is 20000.
- Performed the data cleaning and merging, OLS regression models and Machine Learning techniques such as Regression Trees and Random Forest.
- Concluded that the COVID heavily impact the voting pattern for Trump on 2020, and high-income counties are less likely vote for Trump. Counties voting for Trump at 2016 are also less likely to vote for Trump.

STA304: Analysis of popular vote for the Liberal and the Conservative Party in the 2025 Canadian Federal Election
- Utilized R in this individual project to analyze data and create linear regression models to better understand voting patterns.
- Used post-stratification to divide the population into cells based on age, sex and university completion.
- Predicted that the Conservative Party will obtain a bigger popular vote than the Liberal Party, and received a perfect score of 95/100.

STA313: Advocating for Sustainable Development within Canada
- Solved difficult issues in the Tableau portion as a leader in this group project, and delivered a presentation.
- Tableau was applied to create four interactive dashboards that portrayed increasing energy usage, possible affections on people's health, and predictions using linear regression.
- Achieved 14.5/15 in the Tableau section and 28.5/30 in the final grades. Find the project here

Extracurricular Activities

Henan Normal University Statistics Tutor *Jun.2023 - Jul.2023*
- Developed and delivered statistics tutor sessions to a group of Masters and PhD Physics students.
- Addressed gaps in knowledge of statistical principles to students primarily schooled in a different academic area.
- Systemically explained a wide range of statistical subjects from basic concepts (like distribution) to advanced theories (like linear regression).

University of Toronto Volunteer Notetaker Program *Jan.2023 - Apr.2023*
- Provided academic support to students with disabilities by taking detailed notes during lectures for two courses (STA305 and STA347)

Computer Skills

Technical skills: R, Tableau, LaTeX, Python, SQL, Linux, Fortran, C++
Packages: **Python:** pandas, numpy, matplotlib, sklearn (sklearn.ensemble, sklearn.metrics), BeautifulSoup, geopandas, seaborn, request, linearmodels **R:** tidyverse, ggplot2, Shiny, openintro, knitr, kable, gridExtra, lsmens