Mingming Pan

panxx408@umn.edu| C: 612-422-9299| LinkedIn |Minneapolis, MN 55414

EDUCATION

University of Minnesota, School of Public Health **Minneapolis, MN**
Master of Science in Biostatistics ***Sep 2020 – Present***

- **GPA:** 3.86/ 4.0 (current)
- Degree qualifies for a STEM designation; eligible for 36 months of OPT (12 months OPT+ 24 months of extension)

University of Minnesota, Twin Cities **Minneapolis, MN**
Bachelor of Science in Chemistry; Minor in Statistics ***Sep 2015 – May 2019***

Professional Skills

- Computer Skills: R Language, SAS, SQL, Microsoft Office (Excel, Word, PowerPoint, etc.)
- Analysis Skills: Data Wrangling and Visualization, Machine Learning Techniques, Deep Learning
- Biostatistical Skills: Clinical Trails, Correlated/Clustered Data Analysis, Survival Analysis, Casual Inference, Bayesian Analysis, Hypothesis test, Nonparametric Methods

EXPERIENCE

Global COVID-19 Trends |R | Data Wrangling |Data Visualization
Student Data Analyst ***Jan2022 - March 2022***

- Cleaned, transformed, and extracted relevant global COVID-19 data (Jan 2020 -Feb 2022) from Our World in Data (OWID).
- Analyzed statistics of total numbers and daily averages of new cases and deaths by continents and countries.
- Visualized trends of global new cases and death from Jan 2020 to Feb 2022 by global rolling average plot using R.
- Generated report of global COVID-19 trends with possible future analysis suggestions.

COVID-19 Dream Challenge |SQL |R | Machine Learning | EHR| Imbalanced Data
Student Data Analyst ***Jan 2020 - May 2020***

- Acquired, cleaned, and extracted 10 years of electronic health record (2010-2020) from ~9,500 patients using SQL.
- Employed undersampling to solve high data imbalance issue, and utilized LASSO for variable selection using R.
- Built classification models deploying Logistic Regression, KNN, Decision Tree and FNN to predict COVID-19 test results; delivered the best performing Logistic Regression based model for COVID test result. (96% accuracy)

ACADEMIC PROJECT

Adolescents Suicidal Ideation Project| Longitudinal Data Analysis | Clustered Data
Statistical Methods for Correlated Data, University of Minnesota ***Jan 2021 - May 2021***

- Designed data analyzing plan; created representative predictor, social support score, by merging and transforming relative survey questions.
- Selected, and cleaned longitudinal survey data of over 20,000 adolescents followed for five waves to date collected by Add Health; visualized correlations between demographic measured and suicide ideation using ggplot.
- Built GLMM and GEE models with various correlation structures to analyze how suicidal ideation changes by demographic measures and social support scores; tested model sensitivity by ANOVA and Likelihood Ration Test using R.
- Generated formatted 6-page scientific report with EDA, statistical analysis, and possible future investigation.

Treatment Effects of Diabetes Condition on Medical Insurance Cost| Explanatory Analysis
Methods for Causal Inference, University of Minnesota ***Sep 2021 - Dec 2021***

- Assessed insurance design rationality by estimating ATE (Average Treatment Effect) and ATT (Average Treatment Effects on the Treated) of diabetes condition on Yearly Medical Insurance Cost.
- Applied Regression adjustment, propensity score stratification, and inverse probability weighting separately to control confounding bias; visualized different estimated inferences of ATEs and ATTs separately by forest plot using R.
- Composed a 5-page project report with detailed EDA, data analysis and explanatory statements.

PROFESSIONAL EXPRIENCE

University of Minnesota, School of Public Health **Remote**
Graduate Teaching Assistant ***Sep 2020 – Present***

- Performed all assistant teaching duties for Biostatistics I & Biostatistics II, and SAS programming.
- Monitored labs and demonstrated associated R and SAS programming skills, including data cleaning, data visualization, and modeling.

OTHERS

Languages: Fluent in English and Mandarin (Chinese); Familiar with Japanese and Korean