Fei Liu

1600 South Eads Street
Apartment 316N
Arlington, Virginia 22202

Phone: +1 313-320-4960
liufei8643@gmail.com
https://www.linkedin.com/in/liufei15

Summary

- 5 years professional experience in quantitative analysis, data analytics, and statistical inference
- 5 years experience in statistical programming experience using SAS
- 15 data reports used by New York Times, Washington Post, U.S. News, State of the Union Address, and Department of Education

Education

2008 – 2010	**George Washington University**, *M.S. in Statistics*, GPA 3.91/4.0.
2004 – 2008	**China Agricultural University**, *B.S. in Mathematics*, GPA 3.30/4.0.

Publications

Framework Content Coverage Variation in NAEP: Grade 8 Mathematics and U.S. History. Young Yee Kim, Fei Liu, Sharyn Rosenberg. In *The American Educational Research Association Annual Meeting, 2014.*

Technical Skills

Statistics: multiple linear regression, Ridge/Lasso regression, logistic regression, generalized linear models, hierarchical liner model (HLM), experiment design, hypothesis test, multiple comparison (Bonferoni and FDR), ANOVA, categorical data analysis, factor analysis, model-assisted sampling survey, variance estimation (Taylor series, Jackknife, and BRR), cross-validation, bootstrap, classification-tree, k-nearest neighbors, principle component nalysis

Programming Languages: proficient in SAS (BASE, STAT, GRAPH.Macro, PROC SQL); prior experience in R, Python, STATA, SPSS, and SQL

Operating Systems: Mac OS X and Windows

Databases: MySQL

Certifications

- SAS Certified Base Programmer
- SAS Certified Advanced Programmer
- Coursera Computing for Data Analysis

Experience

1/12 – Present	**Statistical Analyst/SAS Programmer**, *American Institute for Research/MacroSys LLC*, Washington DC.

National Assessment of Education Progress (NAEP) project: NAEP's long-term trend data is the largest nationally representative and continuing assessment of America's student academic progress in various subject areas.

- Utilized factor analysis and multiple linear regression with fixed and random effect to predict difficulty of Grade 4 and 8 reading items with Coh-Metrics factors.
- Implemented bootstrap re-sampling method to create 10,000 hypothetical assessment samples to determine Mathematics and History framework content coverage variability; validated results with the most recent 5 assessments.
- Analyzed complex multi-stage survey data, using techniques such as variance estimation using Jackknife, bias analysis using Chi-square test, t-test, multiple comparison, and HLM.

11/10 – Present	**Statistical Analyst/SAS Programmer**, *American Institute for Research/MacroSys LLC*, Washington DC.

Common Core of Data (CCD) survey project: CCD is the national statistical database of all public elementary and secondary schools and school districts, containing data designed to be comparable across all states.

- Implemented review tests to ensure data integrity, accuracy and consistency; developed SAS programs with SAS/Base, SAS/Stat, Macro, PROC SQL, and ODS for data documentation and reporting.
- Summarized up to 27 years of CCD historical data with descriptive statistics; presented their central tendency and dispersion in concise figures and tables.
- Conducted multicollinearity diagnostics and identified correlations between CCD items to reduce reporting burden.

1/10 – 9/10	**Research Data Analyst Intern**, *Education Data and Policy Center, Academy for Educational Development*, Washington DC.

- Extracted education indicators from household surveys using STATA and SPSS.
- Queried large data sets using SQL: e.g., joining tables, aggregating data with GROUP BY, and deleting duplicates, using RANK function with OVER, sub-setting data with WHERE and HAVING, and creating user-defined variables.

11/07 – 1/08	**Statistical Programmer**, *China Agricultural University*, Beijing, China.

- Conducted univariate analysis, t-test, and ANOVA on data from factorial experimental design and randomized block design.
- Imputed missing data.

Awards

2013	Excellent Employee of the Year, MacroSys LLC
2006, 2007	3rd Prize Academic Scholarship, China Agriculture University

Graduate Courses

Statistical Inference, Statistical Computing in R, Data Analysis, Linear Regression, Categorical Data Analysis, Practical Machine Learning, Model-Assisted Sampling Survey, and Multivariate Data Analysis.