Yueqi Liu

yueqi.liu@uconn.edu | 860-931-9870 | www.linkedin.com/in/yueqiliu

1812 Stearns Hill Road, Waltham, MA 02451

Summary:

Experienced data analyst with solid skills in statistical modeling, data manipulation and data visualization. Proficient in supervised and unsupervised machine learning methods.

Education:

UNIVERSITY OF CONNECTICUT, STORRS, CT *08/2016-05/2018*

Master of Science, Statistics GPA: 3.8/4.0

***Relevant Coursework*:** Applied Statistics, Survival Analysis, SAS and R data management, Applied Time Series, Mathematical Statistics, Statistical Consulting, Designed and Analysis of Experiments, Supervised/ Unsupervised Learning in R (Data Camp).

TAIYUAN UNIVERSITY OF TECHNOLOGY, TAIYUAN, CHINA *09/2012-07/2016*

Bachelor of Science, Statistics GPA: 3.5/4.0

Skills and Certification:

Computer: SAS, R, Power BI, Python, SQL, A/B Testing, Linux, Hadoop, Excel **Certification:** SAS Advance Programming

Professional Experience:

Statistician

LLX Solutions LLC, Waltham, MA *07/2018 - Current*

- Doing efficacy analysis in Oncology studies, using Kaplan-Meier Survival Analysis method to analyze tumor response.
- Generating and validating analytical datasets (ADAM) based on statistical analysis plan along with data Specification.
- Creating and validating Study Data Tabulation Model (CDISC SDTM) domain datasets from existing clinical trial data.
- Producing data reports and data visualization including listings, tables, figures and analysis results using R, SAS and Power BI.
- Lead 5-people team to do data visualization and new-drug analysis for different clients and developed great presentation skills, excellent analytical and problem solving skills with the ability to work simultaneously in multiple tasks and teams.

2017 Travelers Case Competition (1st place in Storrs)

The Travelers Company, Hartford, CT *10/2017-01/2018*

- Conducted predictive modeling for identifying earlier cancel policies using R.
- Fitted GBM, random forest, logistic regression and neural network mode to accomplish predictive tasks.
- Implemented cross validation to prevent overfitting and evaluate model performance.
- Achieved a 0.74 AUC and presented business advice that helps the company to save $15M on unprofitable business.

Analysis of Energy Consumption of Buildings at the Storrs campus

Statistical Consulting Service, UConn, STORRS, CT *01/2017-05/2017*

- Hierarchically clustered buildings and fitted mixed regression model based on time series to forecast future consumption.
- Forecasted the energy consumption of different buildings of UCONN in following years and assisted clients to solve problems.
- Communicated with clients weekly on their requests, assisted the project leader to wrote and transmit reports to clients.

Academic Projects:

Survival Analysis of Heart Failure Patients: Cox Regression

Department of Statistics, UConn, STORRS, CT *10/2017-12/2017*

- Developed a Cox regression model to estimate death rates due to heart failure and to investigate its link with major risk factors.
- Visualized the regression by Kaplan Meier plot to study the survival at different levels of factors: Ejection fraction and gender.
- Developed time-dependent ROC Curves to validate that the discrimination ability of the model is higher at longer follow up time.
- Obtained Martingale residuals to access the functional form of variables and did log-transformation on non-linear variables.

Demographic Choropleth Map using R shiny

Department of Statistics, UConn, STORRS, CT *10/2017-12/2017*

- Retrieved, cleaned and merged shapefile datasets and zip-code datasets in R to create a final dataset to be analyzed.
- Created an interactive choropleth map with the R package 'leaflet' to visualize the demographic information initially.
- Visualized USA zip codes on the Choropleth map and a shiny dropdown to dynamically switch States using R shiny.

Inadequate Sleep as a Risk Factor for Obesity (NHANES): Data Visualization and Logistic Regression

Department of Statistics, UConn, STORRS, CT *11/2016-12/2016*

- Applied logistic regression to identify and determine the association between sleep duration and obesity in SAS.
- Visualized the logistic regression results to better present the results.