

# Lei Wang

Newark, DE, 19711 | 919-813-8868 | lwang8406@gmail.com

<https://www.linkedin.com/in/lei-wang-5106a5112/>

## **EDUCATION**

University of Delaware, Newark, DE **M.S. Statistics** Expected May 2018

- GPA: 3.8/4.0
- Courses: Statistical inference, Database management, Statistical learning, Machine learning (Coursera), Marketing Analytics (Udemy)

Beijing Forestry University, Beijing, China **B.S. / M.S. Ecology** 2010

## **TECHNICAL SKILLS**

Programming language: SQL, R, Python, SAS (Advanced Certified)

Tools: Excel, SPSS, JMP, Tableau, Google analytics

## **EXPERIENCE**

### **Data Scientist Intern**

YouRok – San Francisco Bay area, CA

Mar. 2018 – present

- Analyzed and tracked app user data and campaign performance, interpreted user behavior
- Applied text analytics to discover the popular topics from users to improve app function
- Managed data from various sources (Google analytics, Free app analytics, etc.), created dashboard and identify key metrics and delivered analysis reports

### **Statistical Analyst**

University of Delaware – Newark, DE

Sep. 2017 – present

- Analyzed survey data, performed feature selection (PCA) on 120 variables to reduce dimension reduction, and applied linear regression to interpret relationship between attributes and target
- Provided instruction and suggestions on experiment design and result analysis (ANOVA, multiple testing) and software development assistance for R, SPSS, SAS, JMP
- Designed case study applied in the class, managed StatLab clients and activities

### **Data Analyst Intern**

Fang88 - Palo Alto, CA

Jun. 2017-Aug. 2017

- Developed neighborhood analysis metrics and determined key indicators by collecting demographic data to optimize website and property report contents
- Manipulated 800M data, conducted exploratory data analysis, built linear regression and neural network model under Tensorflow in Python (pandas, numpy, scikit-learn), reduced 2% error
- Delivered house price trend and single property neighborhood analysis articles for social media

## **PROJECTS**

### **Survival analysis on customer retention**

- Applied survival analysis on customer tenure and churn to detect the time when customers tend to churn, grasped the retention trend over time, identified the important factors related to retention

### **AudioTex marketing analysis**

- Evaluated the influence of personal attributes and interaction with AudioTex on customer purchases by establishing multivariate linear regression model to identify key factors

### **Human resources analytics**

- Identified variables that affect loss of employees by implementing classification models (Logistic regression, Decision tree, Random forest, SVM, neural network) and applied clustering analysis (hierarchical analysis, PAM) to segment employees

### **Credit risk modeling**

- Prescreened variables by calculating WOE and IV and implemented binning for some variables, applied logistic regression, random forest and Xgboost models to predict default