

Shijia Xie

601 1st Street, Apt 7
Hoboken, NJ 07030

(607)319-1188
sx85@cornell.edu

EDUCATION

Cornell University , Ithaca, NY	May 2015
M.P.S in Applied Statistics, GPA: 3.72/4.0	
SAS Base Programming for SAS9: 95/100	November 2014
SAS Advanced Programming for SAS9: 90/100	December 2014
University of Michigan , Ann Arbor, MI	May 2013
B.S in Economics, Minor Mathematics	

RELEVANT COURSES

Applied Statistical Computation with SAS	Probability Models and Inference	Linear Models with Matrices
Bayesian Data Analysis: Principles and Practice	Statistical Data Mining	Survival Analysis
Database and SAS High Performing Computing		

SKILLS

SAS, R, SQL, Python, Linux, Unix, Github, Shell, Bash, Stata, SPSS, Spreadsheet, Excel, PowerPoint

PROJECT

Path Analysis of Online Activity, ComScore, Inc December 2014-May 2015
Analyze 1 million+ panelists' online activities based on the panel provided by ComScore, Inc to find the differences between the behavioral patterns of online shoppers, general visitors, and non-visitors.

- Build a predictive model of buyers /visitors / non-visitors of one of the largest online retailers using the comScore panel (2000 buyers and 30000 visitors in the dataset).
- Use Logistic Regression, Step-wise Regression, LASSO Regression, Random Forest and Boosting for dimension reduction and parameter selection.
- Visualize the paths of integrated activities using Tableau, Excel and Power Point.

Analyze Honeybee Genes Sequence using Python, Oracle SQL and SAS April 2015

- Wrote Python files to calculate the relative frequency of the elements that compose bee genes. (107 types of bee genes with the largest sequence composed of 14 thousands elements)
- Connected Python with Oracle SQL, stored the processed data in and Oracle database table, extracted data from the database table when required, and ran queries to find the gene with the highest relative frequencies.
- Designed different web pages to allow users to upload local data files, acknowledge the successful uploading process, and displayed the results gathered from Oracle database.
- Integrated SAS with Oracle SQL to conduct cluster analysis using Proc SQL procedures.

Oracle Database and SAS High Performing Computing December 2014

- Created 5 relational tables from one large table (13588 observations, 52 variables) by setting constraints based on business rules.
- Queried relational database in RDBMS using subqueries and aggregate functions to solve business problems.
- Integrated Oracle database with SAS and SQL developer and did Correlation and Regression Analysis in SAS.
- Visualized the relational database by using ERD Diagram, Microsoft Visio and Composite Usage Map.

RELEVANT EXPERIENCE

Citibank (China) Co., Ltd, Shanghai Branch September 2013-April 2014
Securities & Fund Services Department

- Conducted research on economics and capital markets for advising foreign corporate clients on improving performance of investments in China Capital market.
 - Used Excel to calculate and analyze data from foreign corporate clients' Investment Reports.
- Summarized the results in tables and graphs to complete clients' Investment Plans and Quota Applications.

LEADERSHIP EXPERIENCE

Dream Corps International University of Michigan Chapter

- Community Service Department Founder & Chair September 2012-May 2013
- Events & Fundraising Director September 2011-May 2012