CHEN, Ziyang

400W 113TH ST, 1120
czyang0357@gmail.com · (413)-406-9382

EDUCATION BACKGROUND

College of Natural Science, University of Massachusetts Amherst, MA 9/2014 – 5/2018
B.S. in Physics, Mathematics and Statistics Overall GPA: 3.45 / 4; Major GPA: 3.96 / 4

Mailman School of Public Health, Columbia University, NY 9/2018 - Now
MPH in Biostatistics, Advanced Epidemiology Certificate Overall GPA: 3.56 / 4

Coursework (On campus): Multivariate Calculus, Advanced Calculus, Linear Algebra, Fundamental Concepts of Math, ODEs, Probability, Statistics, Regression Analysis/Variance, Statistical Computing, Linear Regression Model, Epidemiology, Mechanics, Electricity & Magnetism, Computational Physics (Matlab), Theoretical Physics, Modern Physics, Wave Thermodynamics, Advanced Mechanics, Statistical Physics, Quantum Mechanics, Cosmology & General Relativity, Relational database and SQL programming, Statistical computing with SAS.

Coursework (Online): Real Analysis, Time Series Analysis, Inferential Statistics, Non-parametric Statistics, Object-Oriented Programming (C++), Numerical Analysis, Stochastic Processes, Machine Learning, Data Mining.

INTERNSHIP EXPERIENCE

NetEase, Inc. Fall, 2018
Data Analyst in NetEase Youdao Advertising Department
- Screened user information and select effective advertising targets.
- Produced data cleaning and data processing tools by using R and Excel.
- Recorded the activity data (flow rate, click rate) of the advertising merchants and generated the flow statement.
- Visualized the daily consumption of merchants in specific NetEase and Youdao products.

YY, Inc.
Data Scientist in YY Data Operations Technology Department Fall, 2019
- Work focused on short video app developed by YY called Biugo.
- Observed and recorded the daily use indicators of Biugo such as PV(page view), UV(unique visitor), retention and click rate etc. through Haido database.
- Employed HiveQL along with Spark engine to analytically tidy and process the indicator data.
- Created design-specific visual reports for different data analysis strategies.
- Created weekly business report by coding on HiveQL.

RESEARCH PROJECTS

Research Assistant, Study on Spatial Distribution of Pistachio Production using Ising Model Spring, 2017
- Employed the Ising Model, a statistical physics model concerning spatial patterning of ecological system, to statistically explain the spatial distribution of pistachio production; employed Matlab for programmatic implementation of algorithms.
- Conducted visualization on spatial distribution of pistachio production.
- Analyzed spatial correlation of pistachio production; investigated its consistency with Ising model for 1D and 3D cases.
- Statistically established that the current distribution fails to honor the Ising model, indicating the inappropriateness of using a single model to address the complexity of the ecological process.
- Significantly strengthened technical ability concerning data visualization, statistical analysis, Matlab programming, and mathematical modeling.

Research Assistant, Study on Redlining in Chicago Fall, 2017
- Employed R to investigate the presence of redlining (a practice of insurance companies to deny services due to racial discrimination) in the metropolitan Chicago area.
- Conducted data visualization to reveal patterns.
- Established linear regression models; carried out model selection based on a suite of criteria; determined influential point and outliers based on Cook's distance, DIFFITS, and DFBETAS.
- Provided data-driven support to the research team.
- Rigorously disproved the existence of redlining in Chicago; revealed an array of factors affecting the denial of providing voluntary insurance.
- Significantly bolstered interpersonal skill and technical leadership.

Research Assistant, Epidemiological Study Concerning Impact of Climate on Dengue Fever in Thailand Fall, 2017
- Proposed a theoretical hypothesis relating the distribution of Dengue fever to a multitude of weather-related factors.
- Conducted a thorough pre-processing of data in R, e.g., treatment of outliers, irrelevant data, and missing data.
- Collectively used linear regression models and data visualization techniques to determine the existence of interaction term between variables.
- Established statistical correlation between the number of patients with dengue fever and temperature and precipitation.

SKILLS

Programming Languages / Software: C/C++, Java, Matlab, Python, R, SAS, Excel, SQL, Access, HiveQL, Spark