Matthew J. Delhey

2035 Sheridan St
Houston, TX 77030

Phone: (770) 312-3475
Email: matt.delhey@rice.edu
Homepage: mattdelhey.com

Education

Rice University B.A. Statistics and Philosophy, Mathematics minor; *Expected May, 2015.*

Coursework GPA: 3.54 Statistics GPA: 3.8

- Data Mining & Statistical Learning
- Mathematical Statistics
- Database Systems
- Data Analysis and Visualization
- Linear Regression & Statistical Computing
- SAS Statistical Programming

Skills & Activities

Programming Languages R, SQL/Hive, Unix, C, C++, Python, SAS, Matlab.

KTRU (Rice Radio Station) Technical Director & Station DJ

Experience

TripAdvisor *Data Scientist Intern* May 2013 - August 2013

Evaluated the impact of television ads on website traffic, creating a robust model for traffic prediction in R, automated for individual ad assessment. The results served as the primary consideration in TripAdvisor's expansion to a national television advertising campaign.

Created "Best of" list for honeymoon hotels & cities across several countries, identifying relevant reviews through text analysis and regular expressions in Hive and creating the ranking algorithm in R.

Created NLP backend in R for Hackathon project for topic detection of restaurant reviews, resulting in a new sorting option for users.

Projects

(Full reports and code available at mattdelhey.com/projects.)

Independent Research on Interactive Visualization Analytics

Developed Shiny (R library) web application, flyvis.com, in order to explore improving data analysis through the use of interactive web visualization.

Application allows users to analyze airport usage in America. In our paper we demonstrate using our application to find subsets of the dataset whose flight delays can be successfully modeled.

Presented our research which outlines both the application and the design take-aways for future interactive visualizations, receiving an A in the course.

Participation in Kaggle competitions

Stat640: Movie Recommender – Problem similar to Netflix prize; implemented ensemble of various SVD algorithms and KNN using R and LensKit Java framework (placed 18/41).

Titanic – Ensembled & cross-validated multiple models using R and waffles (C++ framework) (placed 253/1181).

Obama for America 2012: An Exploratory Sentiment Analysis

Responsible for scraping tweets and analysis of their sentiment and in R; created timeline visualization demonstrating temportal relationship between Obama's healthcare policy and message sentiment.