Langyi Tian

Phone: +1 347 406 1727   Email: langyi.tian@columbia.edu   GitHub: github.com/lt2710

EDUCATION

M.A. Quantitative Methods in the Social Sciences (Data Science Focus)   **12/2019 (expected)**

Columbia University (New York, NY)

B.A. Global Business Studies and Finance   **05/2018**

The Chinese University of Hong Kong (Hong Kong)

SKILLS

R (web app with Shiny, reporting with Xaringan, machine learning, clustering, visualization, GIS, statistical inference),
Python (NumPy, Pandas, seaborn, scikit-learn, Keras), **AWS** (EC2, S3), **Git**, **SQL**, **Tableau**, **Qlik Sense**, **Spark** (in R)

PROFESSIONAL EXPERIENCE (SELECTED)

Data science consultant, people analytics [Code] [Deliverable]   **05/2019 to 10/2019**

The Agora Companies (Baltimore, MD) Took ownership of data-driven end-user products for attrition and revenue management

- Worked with client executives and middle managers to define business problem and analytical solution
- Built auto-populated visualization reports decomposing KPIs by business units and teams for client executives
- Deployed a Shiny interactive dashboard with revenue projection and metrics monitoring for middle managers
- Developed a logistic regression model to predicting employee attrition for training and HR decision making
- Formulated recommendations on labor force and performance management implemented company-wide

Predictive modeling intern [Code] [Deliverable]   **02/2019 to 05/2019**

First Street Foundation (New York, NY) Predicted market values of 3 million properties in South Florida from administrative data

- Performed EDA, feature selection and imputation with high dimensional administrative data (>200 variables)
- Overcame a unique missing data issue by constructing process to build 85 individual models for subsamples
- Performed modeling selection with regularization Ridge, LASSO, random forest and gradient boosting in AWS
- Built functionalities to measure the sensitivity of performance to data quality and tune data filtering parameters

Market research consultant [Deliverable]   **11/2018 to 03/2019**

Porteñas (New York, NY) Explored marketing solutions of overseas expansion for yerba mate (an Argentina drink) portfolio

- Connected a US-based team and a China-based team to explore partnership opportunities in a new market
- Conducted 4 focus groups, 3 virtual ethnographies to gather feedback on brand image and user experience

ACADEMIC RESEARCH (SELECTED)

The impact of parental wealth on offspring living standards [Code] [Deliverable]   **10/2018 to 07/2019**

- Aggregated data from more than 20 relational data sets in the data lake of a large-scale panel social survey
- Modeled education, income, and wealth with diagnostic models (OLS, quantile, cumulative link, and logit)
- Received school research grant to present research at New York University (Aug 2019)

Social stratification: exploration with unsupervised learning [Code] [Deliverable]   **05/2017 to 09/2018**

- Applied Multiple Correspondence Analysis to visualize consumer dimensions and reduce dimensionality
- Developed a new algorithm for segmentation with bootstrapped hierarchical and K-medoids clustering in R
- Visualize the geographical distribution of 4 social group clusters with GIS packages
- Received travel grant and matching fund from school to present research at Princeton University (Aug 2019)