Kim Eng Ky

kykimeng@gmail.com • +1 (612) 295-5937 • github.com/kykimeng • kykimeng.com

EDUCATION	**MS in Statistics**, University of Minnesota Twin Cities, Minneapolis, Minnesota	Sep 2019 – Present
	BA in Applied Math & Statistics, Macalester College, Saint Paul, Minnesota	Sep 2012 – May 2016

COMPUTER SKILLS

Proficient: R, R Shiny, R Markdown, Hive, SQL, Git, Microsoft Excel
Familiar: Python, LaTex, SAS

WORK EXPERIENCE

Federal Reserve Bank of Minneapolis, Minneapolis, MN
Data Scientist (Project Director) — Mar 2020 – Present
- Provide data science expertise for projects in areas such as affordable housing, access to credit and early childhood development

UnitedHealth Group, Minnetonka, MN
Principal Data Scientist — Dec 2019 – Mar 2020
- Used data analytics to prove or disprove hypotheses related to operations of UHC products, e.g. members who call about dental or vision benefits call more often in a fixed time period
- Developed a method using time series forecasting and probability to set 2020 goals for claim and call related metrics for customer service advocates

Senior Data Scientist — Aug 2018 – Dec 2019
- Predicted MRI procedures in the next two months based on historical diagnoses and procedures using long short-term memory algorithm to help direct members to a more affordable site of care
- Predicted likelihood of members leaving UHC's Medicare products and understand key drivers using Regularized Logistic Regression, Random Forest and XGBoost to assist in retention effort
- Scored healthcare providers based on their patients' answers to questionnaire about their experience using Bayesian Hierarchical Model
- Improved rate of contacting members via outbound calls by recommending optimal time to outreach, which could lead to an increased enrollment in clinical programs and improved health outcome
- Developed and maintained an internal R package including ggplot2 theme and other common functions

Metro Transit, Minneapolis, MN
Data Scientist — Feb 2017 – Aug 2018
- Received employee recognition award at the Transportation Committee meeting for quality work
- Contributed to an internal R package including designing R Shiny application template, Shiny gadget for filtering spatial data, and functions to pull data from relational databases
- Designed, developed and maintained R Shiny applications to visualize spatial data, time series, regression analysis output, and survey data
- Performed monthly transit ridership forecasts at route-level and mode-level using Auto-Regressive Integrated Moving Average Model, Exponential Smoothing State Space Model, and Seasonal Decomposition of Time Series using Loess
- Developed and maintained internal weekly performance reports for bus routes affected by the construction on I-35W highway
- Analyzed survey data (e.g. Customer Satisfaction survey and Employee Engagement survey) using Partition Around Medoid clustering method and Bayesian logistic regression
- Estimated bus time budget (in-motion time, passenger dwell, and other delays) and reliability metrics (e.g. speed and travel time) using automatic vehicle location records

The Brattle Group, Washington, DC
Research Analyst — Jul 2016 – Feb 2017
- Designed and ran Monte Carlo simulations on historical stock and bond returns to estimate expected returns for portfolios with different stock-to-bond ratios
- Estimated quantity demanded based on price and price elasticity of demand derived from literature
- Audited spreadsheets, regression and data manipulation in R, and economic expert reports

Metro Transit, Minneapolis, MN
Research Analyst Intern May 2014 – May 2016

- Designed R Shiny apps to analyze transit ridership trends and forecasting which are used by analysts and schedulers across Metro Transit
- Presented the R Shiny apps to agency-wide audience
- Conducted intensive literature review and statistical analysis on bus accidents, on-time performance, and cross-price elasticities of transit ridership with respect to gasoline prices
- Developed algorithms and visualization to explore bus speed at every hundredth of a mile

VOLUNTEER EXPERIENCE

Women in Machine Learning and Data Science Twin Cities chapter
Founder and co-organizer Jul 2019 – Present

MinneAnalytics 2020 Women in Analytics and Data Science Conference
Co-founder and co-chair Nov 2019 – Present

MinneAnalytics 2020 Data Tech Conference
Co-chair Jan 2020 – Present

noRth 2020 Conference
Organizaing Committee Jan 2020 – Present

UHG Women in Analytics and Data Sceince Group
Co-founder and co-leader Sep 2019 – Mar 2020

PUBLICATIONS

CONFERENCES

[1] Huting, J., Reid, J., Nwoke, U., Bacarella, E., and Ky, K.E. (2016). Identifying Factors That Increase Bus Accident Risk by Using Random Forests and Trip-Level Data. *Transportation Research Record: Journal of the Transportation Research Board*, (2539), 149-158.

PRESENTATIONS

POSTER PRESENTATIONS

[1] Huting, J., Ky, K. E., Lind, E., Freese, R., and Pansch, J. (2018). Understanding Public Transit Rider Satisfaction Using Clustering and Bayesian Regression Methods. *Transportation Research Board 97th Annual Meeting*, (No. 18-05209).

INVITED TALKS

[1] *UHG Analytics Conference*: Calculate Provider Rating Score using Bayesian Hierarchical Model Sep 2019

[2] *UHG Analytics Conference*: Predict Medicare member lapse during Annual Enrollment Period Sep 2019

[3] *NoRth Conference*: Introduction to *data.table* Aug 2019

[4] *UMN Machine Learning Camp for high school students*: Career in Data Science Jun 2019

[5] *MinneAnalytics Big Data Tech*: Build your own R package May 2019

[6] *MinneFRAMA*: Understanding public transit rider satisfaction using clustering and Bayesian methods Dec 2018

[7] *Twin Cities R User Group*: Visualizing Transit Behavior Inventory Onboard Survey with R Shiny Oct 2018