

Yufei Li (Kate)

- [linkedin.com/in/yufei-kate-l-773476a5/](https://www.linkedin.com/in/yufei-kate-l-773476a5/)
- <https://kate0528.github.io/>
- yufeili@uw.edu
- (206)480-9016
- Seattle, WA 98105

EDUCATION

- MS in **Information Management** | University of Washington, Seattle, WA | GPA: 3.79/4.0 Sep 2017 - Jun 2019
- Specialization: Data Science and Business Intelligence.
 - Related Coursework: Data Science (R); Machine Learning (Python); Interactive Data Visualization (D3.js)
- BS in **E-Commerce** | University of International Business and Economics, Beijing, China | GPA: 3.7/4.0 Sep 2013 - Jul 2017
- Related Coursework: Probabilities and Statistics; Linear Algebra; Data Structure

SUMMARY OF QUALIFICATION

- Data Analysis: Python, R
- Database: RDBMS (MS SQL Server), NoSQL (MongoDB)
- Interactive Data Visualization: Tableau, PowerBI, D3.js
- Other Techniques: A/B Testing, Spark, MapReduce, Microsoft Azure (Azure SQL, ML web app), AWS (EC2, RDS), Git, Putty, MS Office
- Web Development: HTML5, CSS, JSON, JavaScript, XML

RELEVANT PROJECTS

- Business Intelligence Solution Design on Sales Data** Mar 2018 - Jun 2018
- Built a dimensional data model ERD containing 10 tables in SSMS based on the source system and product analysis requirements.
 - Defined and implemented efficient ETL processes to move data through data pipeline for need of the business use cases in SSIS.
 - Developed five BI application dashboards with Tableau; provided advice to various regions/product categories and improved sales performance by 23%.
- Hot Spot Analysis on Seattle Real-time Emergencies Data** Oct 2018 - Nov 2018
- Developed an ETL solution for 760,000+ Seattle 911 raw records by using SQL Server Management Studio (SSMS).
 - Used K-means clustering and DBSCAN methods to perform hot spot analysis and identified top 20 hot spots for four different types of 911 emergencies in Seattle area.
 - Built geo-clustering visualizations with Python folium package to provide actionable insights for Seattle Police. Increased resource allocation ratio by 12%.
- Interactive Data Visualization Implementation** Jan 2018 - Mar 2018
- Used D3.js, HTML5, JSON and CSS to create four choropleth maps representing various kinds of violent crime rates by US states.
 - Applied zoom functions to show crime rates in a user-chosen state to help users to identify safe places.
- White Wine Quality Assessment Using Random Forest and XGBoost** Jan 2018 - Mar 2018
- Based on 17 white wine chemical properties, performed feature engineering to train a random forest (out-of-bag error 16.3%) and an XGBoost (accuracy rate 87.5%) classifier.
 - Identified the critical chemical properties (fixed acidity ratio, free sulfur dioxide ratio etc.) influencing white wine quality.

WORK EXPERIENCES

- University of Washington Business Intelligence & IT Department | Data Analyst Intern** | Seattle, WA Jul 2018 - Present
- Built seven space metrics dashboards in Tableau to model current space usage by using SQL Server Management Studio (SSMS).
 - Served as project consultant to Capital & Space Management team to construct logistic growth model to forecast enrollment number for the Provost's Space Report project.
 - Collated financial and enrollment records from the Enterprise Data Warehouse, a central repository of 5.2 billion records; extrapolated enrollment trends through next five years.
- Accenture | Research Intern** | Beijing, China Apr 2017 - Jun 2017
- Assisted to build knowledge graph demos for digital power plants; used MS SQL Server to build schema and encode equipment information (340+ devices) into the equipment knowledge base.
 - Collaborated with Accenture energy consulting team to improve electromechanical generators' KPIs in Pu'an Power Plants.