**Han Yang**

Phone: 573-864-1753 | Email: [hyangkenny@gmail.com](mailto:hyangkenny@gmail.com) | Houston TX

LinkedIn: [bit.ly/33hcPfv](http://bit.ly/33hcPfv) |[Github](https://github.com/hanyang2019): [bit.ly/38knsQ5](http://bit.ly/38knsQ5) | Personal Page: [bit.ly/2WdsrPM](http://bit.ly/2WdsrPM)

**Summary**

Data analyst equipped to run analysis and visualization with skills in Python, JavaScript, SQL, Tableau and Machine Learning. Enjoys challenges, providing detailed analysis of large datasets and proposing insights to solve related problems. Analytical skills, strength in collaborating with people of diverse backgrounds, and effective communication abilities with team members make a valuable addition to any team.

**Skills**

**Technical skills**: Python, JavaScript, Tableau, SQL, MongoDB, Excel VBA, R, SAS, STATA

**Machine Learning**: Random Forest, GBM, XGBoost, SVM, Neural Network

**Experience**

**Data Analyst**SpectraCell Laboratories2014–Present

* Increase daily productivity by creating Python scripts to automate workflow.
* Identify 9 hypertension-associated deficiencies by data-mining 10000+ medical records using SAS.
* Optimize test results and shorten TAT by implementing Machine Learning algorithms using Python.

**Data Analytics Trainee** Rice University2019

* **Prediction of Off-Shore Rig Utilization Rate** | **Github**: [bit.ly/2TCYxmW](https://bit.ly/2TCYxmW) **| Tableau**: [tabsoft.co/366KwAm](https://tabsoft.co/366KwAm)
* Train XGBoost and neural network (LSTM) models in Python for prediction.
* Explore SHAP method to compute feature importance.
* Construct Tableau dashboards to present analysis results.
* **Analysis of Austin B-Cycle** | **Github**: [bit.ly/367Sh9g](https://bit.ly/367Sh9g)| **Heroku**:[bit.ly/2ND1uA0](https://bit.ly/2ND1uA0)
* Manage an ETL pipeline to process 1 million+ trip records using Python and SQL.
* Design web-based interactive visualizations using Javascript to uncover hidden trends.
* **NYC Citi Bike Analytics** | **Github**:[bit.ly/2WAHqTX](https://bit.ly/2WAHqTX) **| Tableau**: [tabsoft.co/2UzKops](https://tabsoft.co/2UzKops)
* Extract and clean/organize 7 million+ trip history data using Python.
* Build Tableau dashboards to reveal bike share trends.

**Teaching Assistant** University of Texas School of Public Health2011 – 2012

* Lecture on hypothesis testing and statistical modeling.
* Tutor students in the application of statistical software (STATA, R, SAS).

**Research Assistant** University of Missouri-Columbia2007 – 2010

* Execute a multi-disciplinary phytoremediation project to restore explosive-contaminated soil.
* Analyze experiment results in SAS.

**Education**

**Data Analytics Certificate: Rice University**, Houston, TX

**Biostatistics M.S.: University of Texas Health Science Center at Houston**, Houston, TX

**Environmental Science M.S.: University of Columbia-Missouri**, Columbia, MO

**Biotechnology B.S.: Southwest University**, Chongqing, China