**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 and Tableau. 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.

**Technical Skills**

**Expert:** Python, JavaScript, Tableau

**Advance:** SQL, MongoDB, Excel, VBA Script,

**Intermediate:** HTML/CSS, R, SAS, STATA, Git

**Areas of Expertise:** Data Visualization, ETL, Machine Learning, Data Mining, Data Modeling

**Projects**

**Prediction of Off-Shore Rig Utilization Rate** | **Github**: [bit.ly/2TCYxmW](file:///C:\Users\hanyang\Desktop\career\bit.ly\2TCYxmW) **|Tableau Public**: [tabsoft.co/366KwAm](file:///C:\Users\hanyang\Desktop\career\tabsoft.co\366KwAm)

* Implemented machine learning algorithms to predict global off-shore utilization rate.
* Ran XGBoost and neural network (LSTM) models in Python, tuned model hyperparameters and explored SHAP method to compute feature importance.
* Created Tableau dashboards to visualize analysis results.

**Analysis of Austin B-Cycle** | **Github**: [bit.ly/367Sh9g](file:///C:\Users\hanyang\Desktop\career\bit.ly\367Sh9g)| **Heroku**:[bit.ly/2ND1uA0](file:///C:\Users\hanyang\Desktop\career\bit.ly\2ND1uA0)

* Created a web dashboard to visualize the distribution of bike stations, the utilization of bikes and the composition of rider membership of Austin B-cycle program.
* Led a team to set up an ETL pipeline to process datasets of approximately 1 million trip records using Python and SQL.
* Used Javascript (D3.js, Leaflet.js and Plotly.js) to make interactive visualizations in an HTML webpage.

**Weather API** | **Github**:[bit.ly/2QhxCu8](http://bit.ly/2QhxCu8) **| Web Page**: [bit.ly/2xF5Pxs](http://bit.ly/2xF5Pxs)

* Constructed a visualization dashboard website to show the relationship of weather conditions of cities with respect to equator.
* Used Python to retrieve weather conditions and geo-coordinates of 500+ cities through the OpenWeatherMap API.
* Ran linear regression analysis and plotted results in Python to visually display the relationship between weather conditions and latitudes.

**Experience**

**Data Analyst/Lab Technician**  2017–Present

**SpectraCell Laboratories** Houston, TX

Perform lipoprotein profile test and collaborate with data analysis team on in-house projects.

* Wrote Python scripts to automate workflow which reduced processing time by 60%.
* Used SQL, Python and SAS to conduct a data-mining project to explore medical records of 10,000+ patients leading to the successful identification of 9 nutrient deficiencies associated with primary hypertension.
* Implemented a machine learning linear regression model in Python to optimize and validate micro-nutrient test results.

**Lab Technician** 2014 – 2016

**Auspicious Laboratory** Houston, TX

Led a team of 3 to perform clinical toxicology urine testing and assumed the role of working manager during the absence of lab manager.

**Teaching Assistant** 2011 – 2012

**University of Texas School of Public Health** Houston, TX

Lectured on hypothesis testing and statistical modeling and tutored students in the application of statistical software (STATA, R, SAS).

**Research Assistant** 2007 – 2010

**University of Missouri-Columbia** Houston, TX

Independently worked on a multi-disciplinary project of phytoremediation and solved research problems with minimal supervision.

* The 2nd Place of The Excellent Poster Presentation for the poster entitled “Plant species-specific phyto-degradation of TNT and RDX contaminants in soil” presented to the 2008 Research Symposium at Lincoln University of Missouri

**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