

夏强

- Github: <https://github.com/ShawnXiha>
 - Email: huakaijianwo2009@163.com
 - Phone: 18679188050
 - Location: 下城区, 杭州, 浙江
-

• Summary

I am a graduate student of food science who find his true passion in data science. As a member of my mentor's research team, I have handled tons of data. I found it very tedious to analysis data with SPSS, so I want to find a better way to do this. After I understand what python and R can do with data, I said this is what I am going to do.

I start teaching myself programming and data manipulating. Now, I finally graduated from Udacity's Data Analyst Nanodegree. I can expertly clean manipulate data with python, R and SQL, find Pattern in data with machine learning, and visualizing data with d3.js.

• Technical Skills:

- Languages: python,R,Javascript,HTML5,CSS3,
 - Frameworks/Libraries: NumPy,Pandas,Matplotlib,scikit-learn,d3.js,dimple.js, ggplot2
 - Other: git/github,SQL,MongoDB
-

Projects

https://github.com/ShawnXiha/Udacity_Data_Analiyst_Nanodegree

1. Investigate a Dataset

- Investigate, analysis and find problems in a dataset using NumPy and Pandas
- Plot data with Matplotlib and find relationship between features
- Analysis significance scipy

2. Explore and Summarize Data

- Use R and apply exploratory data analysis techniques to explore data set for distributions, outliers, and anomalies.

3. Identify Fraud from Enron Email

- Analyze, clean and get data ready for machine learning, combine features with PCA (principal component analysis)
- Tune machine learning algorithm to get best performance. My final classifier have scores: Precision: 0.38639 Recall: 0.59600

4. **Wrangle OpenStreetMap Data**

- Parse map data of Shanghai from OpenStreetMap, find problems in map data
- Fix Chinese mixed English and Pinyin street name, and shape data into MongoDB
- Analysis data with MongoDB

5. **Make Effective Data Visualization**

- Data visualization with d3.js and dimple.js
- Get feedback, improve visualization according to feedback

6. **A/B testing**

- Analyze the results of an A/B test that was run by Udacity and recommend whether or not to launch the change.
 - Design a A/B test on reducing the number of student cancel after free trial
-

Education

Udacity Data Analyst Nanodegree 2016.6~2016.9

Nanchang University MSc of Food Science 2013.9~2016.6

Other Skills:

Experiment design and analysis

PPT, Speech, Pomodoro