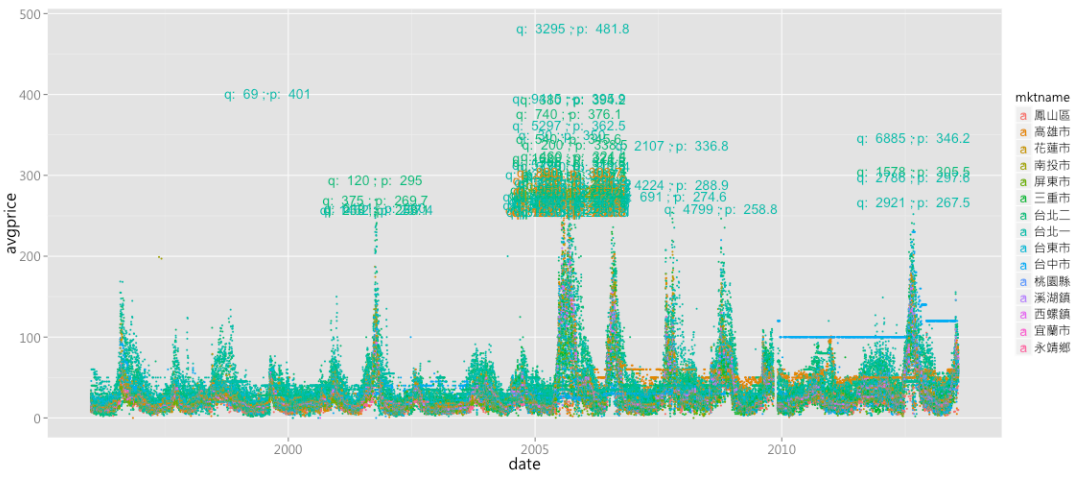
Real-time visualization with Python and d3.js

PyCon APAC 2014 Lighting talk proposal

video license: cc 3.0 BY
type: 25 min Talk
language: English
target audience: the technical part will benefit programmers that are interested to Data Visualization; the underlying message of the talk will benefit the general audience on why collaboration between human and machine is needed

Pyhton has great power in scrapping and analyzing data, and D3.js is a great tool for building visual interface. In the first part of the talk, I will demonstrate how to set up D3.js as an interactive layer on top of Python. In the second part of the talk, I will show what it can achieve, by using the "Taiwan Vegetable Auction" dataset (past 10 years transaction data of 127 kinds of vegetable, 1GB).

The dataset is too large for human to see through, a machine learning algorithm will be able to fit a regression model on the dataset, but it can't make sense of it. For example, in the following graph:



You can see the average price of green onion in different markets for the past 20 years, the fluctuation between different markets are similar, yet since 2010, there is a perfectly horizontal blue line: while price in the other markets have changes, price in "Taitung" has remained exactly the same, suggesting a possible case of monopoly. One can then asks the system to detect other cases of monopoly. This process of exploratory data analysis can only be possible with both machine and human.

From a technical perspective, this talk can benefit front-end developer/data scientist to set up such a system. Yet a more profound value of the talk will be to explore how machine and human can work together.

Muyueh Lee is a programmer focus on Data Visualization, he is also the administrator of the d3js.tw group, and hosts "Visualization Lighting Talk", a gathering for programmers doing data visualization work in Taiwan.