Sylebra Internet Data Analyst Case Study

Search keyword coverage project

* We would like to conduct a project that is able to identify advertisers, coverage, and pricing from 2 search engines, Baidu and Qihoo
* Search keywords are ready
* Need a program to automatically search keywords and grab advertisers information
* Also need to identify advertisers on the right hand side and in the algorithm results
* Need to identify advertisers in brand zone
* IP
* BIDU keyword pricing trend
* Use API to track down the pricing of keywords by categories
* API username yuysun, password Youth@!@!2121
* Website <http://cas.baidu.com/?tpl=www2&fromu=http%3A%2F%2Fwww2.baidu.com%2F>

* For the results, should be able to show
* # of advertisers under each category
* Net out repeat advertisers
* Keywords coverage comparison Baidu vs. Qihoo
* Keywords pricing trend
* # of brand zone advertisers
* Resources
  + Computer
  + Internet
  + Printer (Brother Printer A)

hi, 在么

我现在把抓取的程序做好了，也抓了一部分数据，关于最后的output我还有一些问题想要为问问，

1. 关于第二点和第四点，你们想要的是不是每个keyword在奇虎和百度不同位置的广告主的url. 以及数量

2. 关于第一点不是很明白of advertisers under each category， Net out repeat advertisers

3. pricing那个东西据我所知做不到，因为没有这种api