

27、数据可视化

1、什么是数据可视化

数据可视化是将数据以图形图像等形式表示，直接呈现数据中蕴含信息的处理过程。

有许多软件或者工具支持数据可视化。在前面学过的内容中，Excel等电子表格软件可以通过插入图表操作对表格数据进行可视化，Python中的matplotlib库也可以对图表进行可视化，wouldcloud 库则可以根据文本内容生成标签云，从而实现对文本数据的可视化。

数据可视化的作用包括：快捷观察与追踪数据、实时分析数据、增强数据的解释力与吸引力等。

2、可视化的基本方法

有关时间趋势的可视化——线形图、柱形图

有关比例的可视化——饼图、环形图

有关关系的可视化——散点图、气泡图

有关差异的可视化——雷达图

有关空间关系的可视化——地图上进行标注

3、可视化工具

软件与平台：大数据魔镜、Gephi、Tableau

编程语言：JS——D3.js、Highcharts、Google Charts

Python——matplotlib

JS平台上的基本都支持H5，用于网页应用程序

4、可视化探索——百度指数

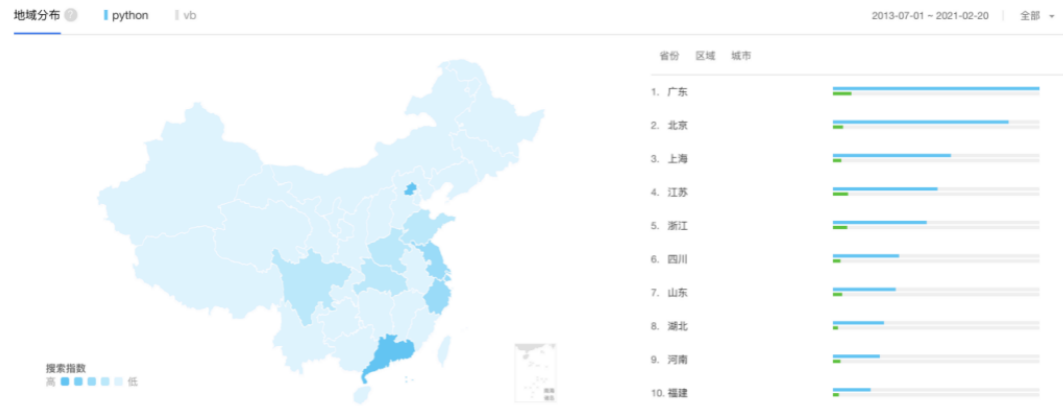
百度指数（Baidu Index）是以百度海量网民行为数据为基础的数据分析平台，是当前互联网乃至整个数据时代最重要的统计分析平台之一，自发布之日便成为众多企业营销决策的重要依据。

“世界很复杂，百度更懂你”，百度指数能够告诉用户：某个关键词在百度的搜索规模有多大，一段时间内的涨跌态势以及相关的新闻舆论变化，关注这些词的网民是什么样的，分布在哪里，同时还搜了哪些相关的词，帮助用户优化数字营销活动方案。

使用搜索引擎搜索“百度指数”或者直接通过域名<http://index.baidu.com/>即可访问使用百度指数。尝试在百度指数搜索框中输入关键词“python，vb”，即可以查看python和vb两种编程语言这些年以来的热度变化趋势，如下图：



切换到“人群画像”模式，可以进一步看到我国搜索python和vb的人群地域分布，如下图：



请尝试访问百度指数网站，搜索自己感兴趣的关键词，并尝试发现一些现象、得到一些结论。