

工具创新，让数据分析回归本源

# 三个数据分析师的日常

# 数据分析

业务理解 + 统计知识 + 技术工具

# 数据分析（理想）

业务理解 + 统计知识 + 技术工具

# 数据分析（现实）

业务理解 + 统计知识 + 技术工具

数据分析师

需要从繁重的操作实现中解放出来！

需要有更多精力去解决更高层次的问题！

我们需要更好的工具！

Exited: Acquisition or IPO

## Applications

The figure is a large, colorful grid of logos for various data science and technology companies, categorized by industry and function. The grid is organized into rows and columns, with each cell containing a cluster of related logos. The categories include:

- NoSQL Databases:** Includes logos for Redis, MongoDB, Couchbase, Aerospike, HypeTable, Sqrrl, Amazon, and Neo4j.
- NewsSQL Databases:** Includes logos for MarkLogic, RainStar, deepdb, Citusdata, Clustrix, VoltDB, and others.
- MPP Databases:** Includes logos for Teradata, InfiniDB, Vertica, Netezza, and others.
- Graph Databases:** Includes logos for Neo4j, Asterdata, and InfiniteGraph.
- Crowd-sourcing:** Includes logos for Servio, Microtask, and others.
- Cross Infrastructure:** Includes logos for SAP, SAS, IBM, Google, Microsoft, VMware, Amazon, and others.
- Analytics Platforms:** Includes logos for Cloudera, Amazon, Microsoft, Pivotal, and others.
- BI Platforms:** Includes logos for Tableau, Qlik, and others.
- Machine Learning:** Includes logos for IBM Watson, Palantir, and others.
- Location/People/Events:** Includes logos for Radius, Fliptop, and others.
- Big Data Search:** Includes logos for HP, LucidWorks, and others.
- Crowd-Sourced:** Includes logos for Kaggle, DataKind, and others.
- Ad Optimization:** Includes logos for Rocketfuel, Tapad, and others.
- Finance:** Includes logos for Lenddo, Cignifi, and others.
- Government/Regulator:** Includes logos for Mark43, Enigma, and others.
- Health:** Includes logos for Recombine, Ginger.io, and others.
- Industries:** Includes logos for Tubular, OPOWER, and others.

100

---

# 哪些新特性即将成为分析工具的主流

交互式可视化

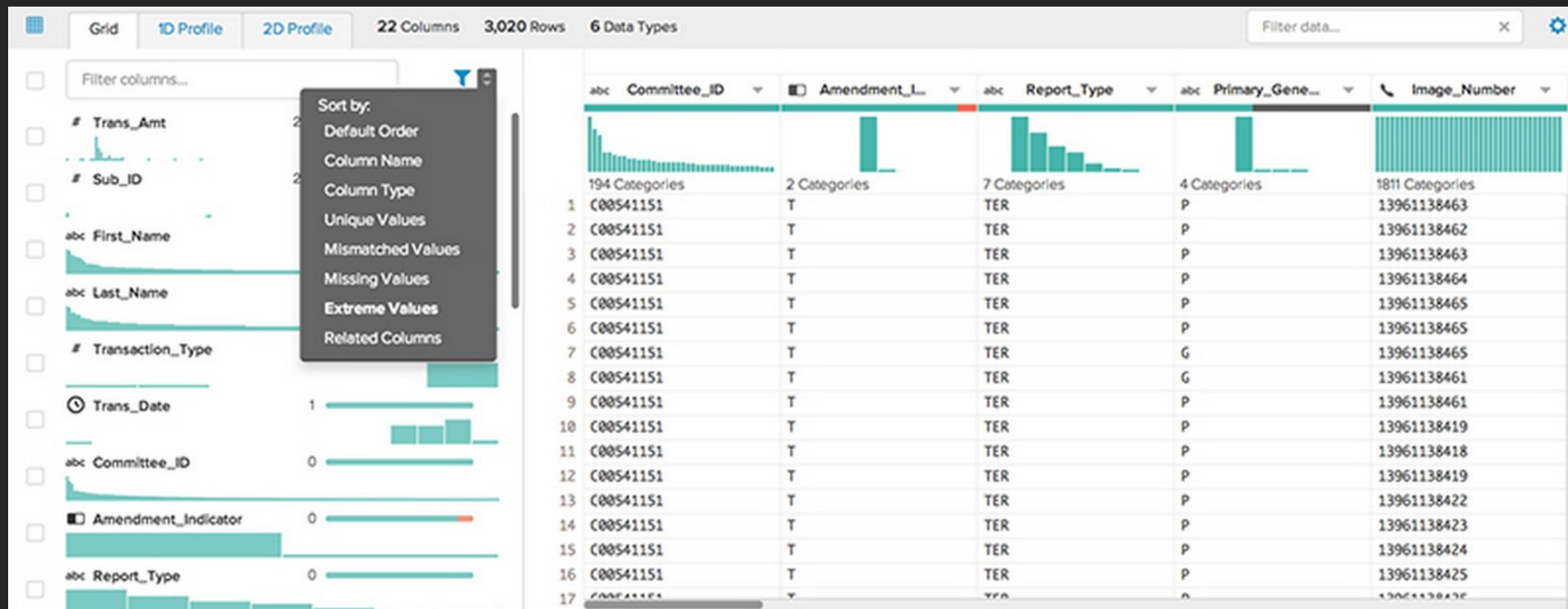
高性能云计算

一体化

数据一键接入

低获取成本

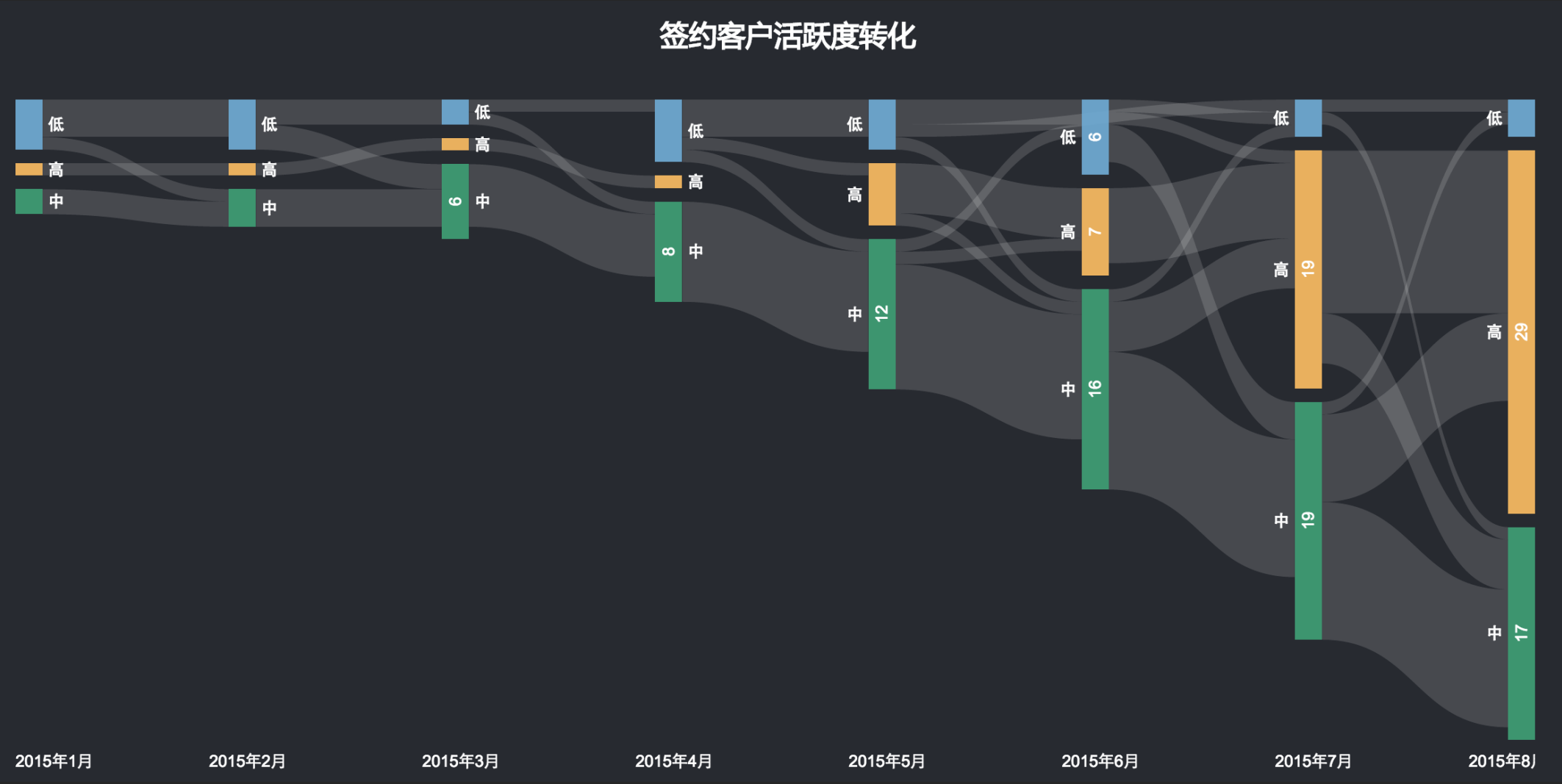
# 交互式可视化



# 交互式可视化

	Year	extract	Property_crime_rate
0	Reported crime in Alabama	Alabama	
1	2004		4029.3
2	2005		3900
3	2006		3937
4	2007		3974.9
5	2008		4081.9
6	Reported crime in Alaska	Alaska	
7	2004		3370.9
8	2005		3615
9	2006		3582
10	2007		3373.9
11	2008		2928.3
12	Reported crime in Arizona	Arizona	
13	2004		5073.3
14	2005		4827
15	2006		4741.6
16	2007		4502.6
17	2008		4087.3
18	Reported crime in Arkansas	Arkansas	
19	2004		4033.1
20	2005		4068
21	2006		4021.6
22	2007		3945.5
23	2008		3843.7
24	Reported crime in California	California	
25	2004		3423.9
26	2005		3321
27	2006		3175.2
28	2007		3032.6

# 交互式可视化



# 高性能云计算

计算性能很重要！

计算性能很重要！

计算性能很重要！

# 计算性能很重要

- 计算性能的大幅提升能带来完全不一样的灵活性和分析效率
  - 超高计算性能是交互式可视化的前提条件
  - 数据量和数据的使用量都在不断增大

# 云计算是答案



# 一体化



# 数据一键接入



## 低获取成本

- 云端产品，即开即用
  - 价格低廉

工具创新，正在让数据分析回归本源

国外的工具创新正如火如荼

中国呢？

BDP

成员来自百度、微软、IBM、微策略等数据领域知名公司

2年多的研发



现在，我们希望开放我们的能力，  
让更多数据分析师受益



分析师个人版

免费

2015 Q4

微信发起群聊 -> 面对面建群 -> 输入 0911



敬请期待