

文章编号: 1005-8451 (2016) 09-0079-04

大数据时代两款敏捷商业智能工具对比分析

王 喆, 戚小玉, 东春昭, 杨涛存, 邹 丹

(中国铁道科学研究院 铁路大数据研究与应用创新中心, 北京 100081)

摘要: 大数据时代的来临使得对数据的高速捕获和实时分析变得越来越重要。相对于传统商业智能软件体量大、对业务变化响应时间长、灵活性差的缺点, 敏捷商业智能软件采用轻量建模方式迅速响应需求变更, 用户可以通过简单的拖拉拽实现对业务的分析和展示。本文从使用角度对比分析了两款主流敏捷性商业智能软件Tableau和Qlik Sense。

关键词: 商业智能; 大数据; 敏捷性BI; Tableau; Qlik Sense

中图分类号: U29-39 **文献标识码:** A

Comparative analysis of two agile business intelligence tools in big data era

WANG Zhe, QI Xiaoyu, DONG Chunzhao, YANG Taocun, ZOU Dan

(Research and Application Innovation Center for Big Data Technology in Railway, China Academy of Railway Sciences, Beijing 100081, China)

Abstract: Along with the advent of the era of big data, the high-speed data capture and real-time analysis are becoming more and more important. Compared with the traditional intelligent software of large size, long responding time for business changes and poor flexibility, lightweight modeling method was adopted by agile BI tools in order to adapt to changes in demand. Users could implement the analysis and display of business through a simple drag and drop. This article focused on the comparative analysis of two mainstream agile business intelligence software of Tableau and Qlik Sense.

Key words: business intelligence; big data; agile BI; Tableau; Qlik Sense

近年来, 随着移动互联网、物联网等技术和应用的兴起, 大数据在全球范围内备受关注, 对大数据的定义也有多种提法。麦肯锡全球研究院将大数据定义为“无法在一定时间内使用传统数据库软件工具对其内容进行获取、管理和处理的数据集合”^[1]。大数据可以认为是分析的另一种表述, 从数据中萃取知识, 并将其转化为商业优势的智能化活动^[2]。

商业智能(BI, Business Intelligence)软件通常被理解为是将企业中现有的数据转化为知识, 帮助企业做出明智的业务经营决策的工具。进入大数据时代以来, BI软件越来越注重快速建模、用户自主分析、数据可视化和多设备支持。数据可视化丰富的效果图简化了阅读者的分析使用过程, 从而大大提高了阅读者的决策和判断的能力、有利于企业提升信息资源和服务的水平^[3]。

从Gartner公司发布的2016年商业智能和分析

收稿日期: 2016-06-15

基金项目: 中国铁路总公司科技研究开发计划课题(2015X003-F); 中国铁道科学研究院基金重大项目(1551DZ8004)。

作者简介: 王 喆, 副研究员; 戚小玉, 助理研究员。

软件的魔力象限报告中不难看出(见图1), 在处于行业领导地位的象限中, Tableau公司和Qlik公司的商业智能产品在执行能力(考量产品的易用程度和价格、服务的完善程度和技术支持能力、管理团队的经验和能力等)评价中居于领先地位。



图1 Gartner公司2016年商业智能魔力象限

Qlik Sense是Qlik公司在成熟产品Qlik View基础上推出的一款轻量级BI产品, 其快速建模和自主式分析能力降低了BI工具的使用门槛。Qlik Sense

主要包括 Desktop 和 Server 两个产品, Desktop 针对单机版用户, Server 版针对企业用户, 二者在分析和可视化功能上并无差别。

Tableau 的核心产品是 Tableau Desktop 和 Tableau Server。其中, Tableau Desktop 是为企业数据工作者和数据分析人员设计的一个可视化数据分析工具。而 Tableau Server 是一个 Tableau 可视化内容的发布服务器, 它通过结合 Tableau Desktop 设计好的可视化仪表盘和分析图表, 让企业的业务用户, 可以通过使用 Web 浏览器或者移动设备访问查看可视化仪表盘的功能。

下文将主要从用户使用角度来对比敏捷性 BI 产品 Tableau Desktop&Server (简称: Tableau) 和 Qlik Sense Desktop&Server (简称: Qlik Sense)。

1 Qlik Sense产品

1.1 数据处理

1.1.1 数据源

在 Qlik Sense 中是按照应用程序→工作表的层级管理可视化报表, 报表的可视化效果如图 2 所示。

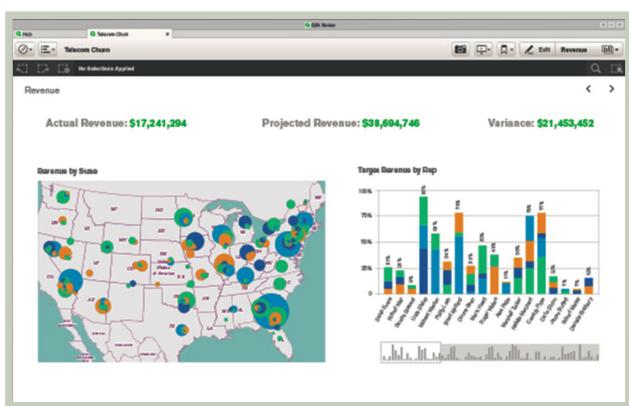


图2 Qlik Sense可视化效果图

对于同一个应用程序而言, 内部的各工作表共享同样的数据连接, 也就是连接到同一个或多个数据源。Qlik Sense 支持的数据源种类有很多, 包括: Excel 文件、ODBC 数据源、OLE DB 数据源、Web 文件、所有文件、逗号分隔文件几类。如果遇到了不在上述支持列表中的数据源, 还可以通过 Qlik 提供的编程接口 QVX SDK 来建立用户定制化连接。

1.1.2 数据加载

Qlik Sense 中数据加载方式有两种: 内存访问方

式和实时加载模式。内存访问方式是第 1 次装载数据的时候将数据全量加载到内存中, 其后的可视化分析运算都基于内存数据, 如果数据源有更新的话可以在 QMC (Qlik Management Console) 中通过运行加载任务的方式进行增量加载。内存访问方式的特点是速度快且对数据实时性要求不高。实时加载模式是实时地从数据源读取数据并刷新可视化视图, 特点是数据实时性较高但是对资源消耗比较大, 比较考验数据源的性能。数据实时加载模式需要在数据加载编辑器的脚本中添加 Direct 关键字来实现。

1.1.3 数据模型

数据源加载完成后, Qlik Sense 自动将不同数据表中字段名称相同的字段或者数据一致性程度较高的两个字段识别为关联字段 (也可以手动建立或者取消关联)。通过表与表之间的关联字段建立支持整个应用程序的数据模型, 如图 3 所示。

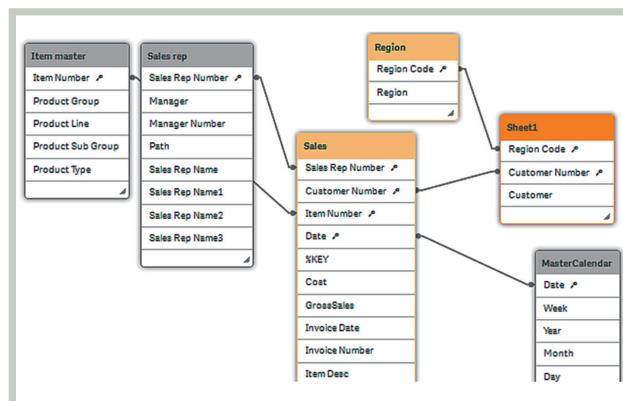


图3 在 Qlik Sense Enterprise 模型查看器中浏览模型

数据模型是创建应用程序后进行数据分析和可视化的基础。在创建数据模型的过程中, Qlik Sense 还提供了较多的数据表转换工具, 例如: 交叉表转换、数据行变字段、按层级组织数据等, 分别通过在数据加载脚本中添加 Crosstable、Generic、Hierarchy 关键字来实现。这些转换相当于在原始数据源和分析数据源之间做抽取、转换、加载 (ETL) 工作, 方便用户使用。

1.2 可视化方式

1.2.1 基本控件

Qlik Sense 为用户提供了最基本的可视化控件, 包括: 饼图、地图、散点图、筛选器、树形图、条形图、透视表、KPI、仪表、折线图、组合图等, 不同的视

图展现方式针对不同的场景。Qlik Sense 的每张工作表上都可以添加多个可视化控件，通过优化组合的方式针对同一主题从不同的角度分析和展现数据。

1.2.2 扩展控件

除了系统提供的最基础的可视化控件之外，如果用户还需要定制个性化的可视化效果，可以通过扩展控件的方式实现。Qlik Sense 提供了扩展控件应用程序接口（API），用户可以使用 Javascript 调用扩展 API 编写出满足自己需求的控件，丰富 Qlik Sense 的展示效果。另外，在 Qlik 的官方社区中，爱好者们公布了各式各样的个性化控件，任何人都可以下载并集成到自己的 Qlik Sense 中使用。

1.2.3 同企业业务系统集成

如果要将 Qlik Sense 的可视化报表集成到企业现有的业务系统中，必须使用 Qlik Sense Enterprise。通过 Qlik Sense Enterprise 创建的应用程序和工作表都自动保存在服务器端，每个用户只能访问到自己目录下的应用程序。可以在 QMC 中发布应用程序从而将应用程序共享给多人浏览，应用程序一经发布只可以浏览不可以更改。

每个工作表创建完成后无论是否发布，Qlik Sense 都会为其分配一个统一资源定位符（URL）来唯一标识。在企业应用中，用户可以通过在企业 Portal 中访问该 URL 的方式将 Qlik 的报表视图集成到企业应用中，便于和现有业务的整合。

1.3 地图扩展

企业业务分析系统越来越注重企业数据同地理信息数据的集成。Qlik Sense 把地图作为一个控件显示在工作表中，地理底图只能从 Qlik 默认的地图 URL 中选择，底图上面可以叠加用户地图和数据。用户使用其他工具制作的地图或者图层数据可以通过导出 KML 文件的方式关联到 Qlik Sense 的数据源中，然后叠加显示在地图控件上，这样就完成了用户数据同地图数据的集成。由于不支持自定义底图，可视化分析时必须保证服务器能够连接到互联网才能正确显示底图。

1.4 API

Qlik Sense 提供了较为丰富的 API，用户几乎可以从各个层面根据自己的需求对 Qlik Sense 进行定制

化开发，表 1 列出了 Qlik Sense 几种 API 类型和功能。

表1 Qlik Sense 提供的 API 种类和应用场景说明

API类型	使用场景
Client side APIs	定制化控件、mashup，系统集成
Server side APIs	操作用户、管理认证和 session 等
QVX SDK reference	定制化数据连接
.Net SDK reference	使用 .Net 在 Windows 应用程序中集成应用

2 Tableau 产品

2.1 数据处理

2.1.1 数据源

Tableau 支持的数据源种类繁多，多达上百种。其中，除了常见的 Oracle、DB2、SQLServer 等数据库，还包括 Hortonworks Hadoop Hive、Cloudera Hadoop Hive&Impala、Teradata、Google Analytics、Web 数据连接器等大数据类型的数据源。另外，所有兼容 ODBC 3.0 的数据库和应用程序都可以直接作为 Tableau 的数据源。Tableau 软件的可视化效果如图 4 所示。

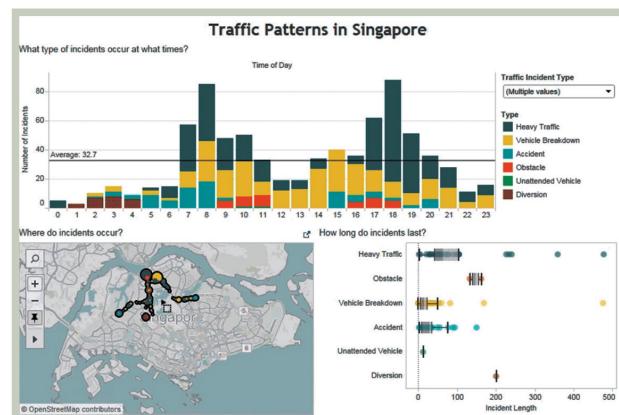


图4 Tableau的可视化效果图

2.1.2 数据加载

无论是 Tableau Desktop 还是 Tableau Server，其数据加载方式都有两种：实时方式和数据提取方式。这两种方式与 Qlik 中的类似，实时数据加载表示 Tableau 直接连接到数据源，加载性能由数据源的性能决定；数据提取方式是将数据源中数据加载到 Tableau 快速数据引擎中，后面对数据的操作都基于数据引擎中的数据进行。与 Qlik 不同的是，Tableau 软件在这两种数据加载方式之间切换时可以非常方便地在界面上直接选择，无需修改数据加载脚本。

2.2 可视化方式

在 Tableau 中, 对数据源中各数据表的关联关系建立好之后, 系统会自动根据数据表中的字段类型、特点划分成维度和度量两类。维度是分析问题的角度, 比如: 年、月、日、城市、单位等; 一切可以测量的值都可以被称作度量。Tableau 根据字段特点自动创建划分的维度和度量, 还可以通过人工调整的方式修改, 或者添加新的计算字段用于补充现有维度和度量。在确定了维度和度量之后, 在 Tableau 中就可以通过鼠标拖拉拽的方式实现自主式分析。

Tableau 提供的可视化控件和 Qlik 类似, 包括: 饼图、柱状图、树形图、折线图等。所不同的是, Tableau 会根据用户鼠标选中的度量或维度, 分析该字段的数据特征并智能地对推荐使用的显示方式加以高亮显示, 辅助用户决策。

2.3 地图扩展

Tableau 提供了较为丰富的地图支持方式, 比如: 脱机地图、Tableau 地图服务、Web 地图服务 (WMS) 等。另外, Tableau 还可以自定义地图的地理编码、图层、图例等信息。

2.4 API

Tableau 提供的 API 大体上可以分为以下 3 种类型, 如表 2 所示。

表2 Tableau提供的API及其功能

API类型	使用场景
ExtractAPI	在第三方软件中创建TDE文件; 定制化数据连接
REST API	通过HTTP来管理和更改 Tableau Server 资源
Javascript API	可以将Tableau可视化项整合到用户自己的Web应用程序中

2.5 与R的集成

R 是用于统计计算的免费开源编程语言和环境。它是发展最快的统计编程语言之一, 并正在成为大多数数据科学家首选的工具。R 语言的优势是统计和分析功能强大、函数众多, 但是缺点是对大数据的管理较差, 人机交互功能较弱。R 与 Tableau 的集成, 可以使用 R 来处理统计分析的“重头戏”, 而 Tableau 让用户在简单的拖放环境中研究并可视化表示结果, 将统计发现转变为漂亮的仪表板。

3 结束语

Qlik Sense 和 Tableau 软件都很好地兼具了 BI

分析工具的数据分析能力以及简单易操作的特点, 符合敏捷型 BI 工具的要求。二者强大的数据可视化能力在数据分析专家和抽象数据之间架起了一座桥梁, 能够使人们以最直观的方式从数据中获取知识和经验。同时, 两款软件在一些细节功能上各有特点, 细节功能对比参见表 3。

表3 Qlik Sense和Tableau软件细节功能对比

	多数据源	数据加载方式	可视化控件	地图功能	API及扩展性	易用程度
Qlik Sense	支持	支持内存访问和实时加载	提供基础控件; 还提供用户自定义控件编程	不支持离线地图和WMS服务	提供丰富API	基础应用简单, 高级应用可以通过编写脚本实现
Tableau	支持	支持实时和数据提取	提供基础控件; 尚未提供用户自定义控件	支持离线地图和WMS服务	除了API之外还提供与R的集成	非常简单, 无需编写脚本

通过上述内容不难发现, Tableau 软件易用性好、用户上手快, 比较适合个人用户以及没有经过可视化分析培训的企业人员使用, 使用者能够以最快的时间将自己所掌握的数据以可视化图表的形式展现出来, 并通过 Tableau Server 共享给其他用户; Qlik Sense 以其丰富的二次开发接口和可扩展性, 提供给用户充分的定制能力, 使用户可以通过数据加载脚本和 Qlik API 来定制数据连接和可视化报表的展现方式, 比较适合较大规模的企业级应用。

参考文献:

- [1] Manyika J, Chui M, Brown J, et al. Big Data: The Next Frontier for Innovation, Competition and Productivity [R]. McKinsey Global Institute, 2011: 11.
- [2] McAfee A, Brynjolfsson E. Big Data: The Management Revolution [J]. Harvard Business Review, 2012 (October): 61-68.
- [3] 段建, 翟慧敏, 程杰. 大数据背景下可视化设计趋势应用研究 [J]. 电脑知识与技术, 2015 (11): 3-4.
- [4] Tableau Software[EB/OL]. <http://www.tableau.com>, 2016-6-12.
- [5] Qlik[EB/OL]. <http://www.qlik.com>, 2016-6-12.

责任编辑 王浩