

深入理解旋转透视图

2016-02-29 poon BI实验室

看了不关注↑都是耍流氓 

[TOC]目录

- 什么是旋转透视图 (pivot table)
- 理解旋转透视图的几个要点
- 旋转透视图制作过程分析
- 旋转透视图和交叉表
- 旋转透视图与obiee 的度量标签
- 总结

关键字：旋转透视图，交叉表，列联表，互分类表，crosstabs，Contingency table，聚合，cross tabulation

什么是旋转透视图 (pivot table)

1. 从数据存储上来说，不管是普通表还是旋转透视图，都是对应同一份数据，同一个表，所以旋转透视图跟数据存储无关，而跟展现有关。
2. 从数据运算来说，比如excel，有完善完善的旋转透视功能，excel上的大部分计算，其实都是矩阵的计算，而矩阵可以用数组/列表来表达，所以，旋转透视图其实可以看作是数组（列表）运算。
3. 从数据粒度来说，一般旋转透视图最大的功能就是数据聚合，并且根据需要，可以切换不同的维度来看聚合数据的变化。
4. 从展示效果来说，旋转透视图最大的特点就是行列转换，人机交互。也就是说，数据展示可以根据人的需要动态变更，实现人机交互。
5. 从维度数量来说，旋转透视图要求数据表至少有两个或以上维度。如果数据表连

2个维度都不具备，那么没必要使用旋转透视图。

6. 从应用需求来说，最常使用的是行列转换和数据聚合功能。也就是把行转成列，制作成交叉表／情形分析表等。

所以，可以这样理解旋转透视图：

旋转透视图是一种基于列表的、支持多维度数据聚合的、可以人机交互的数据展现工具。

理解旋转透视图的几个要点

1. 核心功能：行列转换、数据汇总。
2. 扩展功能：数据排序,聚合方式扩展,数据筛选
3. 应用特点：人机交互（根据用户要求切换数据视角。不是改变数据，而是改变数据的展示方式和计算规则）

我们来看看最原始的选择透视图特征：

A tool that could help the user recognize these patterns would help to build advanced data models quickly. With Improv, users could define and store sets of categories, then change views by dragging category names with the mouse. This core functionality would provide the model for pivot tables.

-- 来自 wikipedia 的 pivot table 词条

Improv 是最早的旋转透视图。注意这几个关键词：user，build，data models，define，change views，by .. mouse。

显然，这段话着重强调了*人*的参与。即人机交互。

旋转透视图制作过程分析

无论我们用excel、biew或者其他工具制作旋转透视图,一般都会经历筛选维度,筛选度量,选择聚合方式这几个步骤。经历这几个步骤之后,我们得到一张新的表。然后我们切换维度或度量,变更聚合方式,又得到一张新的表。

其实这个过程,跟我们创建一个视图(create view),或者手工从一张 **交易明细表** 汇总出一张 **地区 - 销量表** 的过程没有什么两样。只是对于旋转透视图来说,不用我们自己写代码,由程序自动完成了而已。

通常发生在旋转透视图的计算,都是基于一个清单级别的数据集,计算出一个粒度较粗的数据集合,并且在内存中直接完成。因此,我们在excel中使用旋转透视图的时候,常常会有电脑变慢的情况,因为excel在做旋转透视的时候,需要较多的内存去做计算和数据展现。

旋转透视图和交叉表

交叉表,又叫列联表,又叫情形分析表, cross tabulation , crosstabs , Contingency table ... 越叫越复杂,其实都是指同一个东西。其实就是对一张数据表,取出两个维度,一个按列展示维度值,一个按行展示维度值,度量数据放中间。

由于旋转透视图的行列转换功能,导致旋转透视图非常适合做交叉表。一度让我有一段时间把交叉表当作旋转透视图了。其实旋转透视图不仅限于交叉表的制作啦。

旋转透视图与obiew 的度量标签

biew 的旋转透视图里,有一个度量标签。其作用是分配度量列的标头的展示的位置。有点拗口? 简单来说,度量标签放哪里,度量列对应的表头就展示在什么位置。

比如有表(dim1,dim2,dim3,mea1,mea2,mea3),我们把 dim1,dim2 放在行(rows)区域,把dim3放在列(columns)区域,把mea1,mea2,mea3 放在度量 measures区域,那么就形成了一张经典的透视图。当我们把度量标签放到 columns 区域的时候,度量列的表头(字段名称) mea1,mea2,mea3 将展示在列区域中。

当我们把度量标签放到 rows 区域的时候,度量列的表头(字段名称) mea1,mea2,mea3 将展示在rows区域中,并把mea1,mea2,mea3 按行列出。

参考:

```
http://obiee101.blogspot.com/2008/05/obiee-pivot-table-measure-label.html
```

(请自觉翻墙访问)

总结

旋转透视图是为了提高客户感知而开发出来的。其亮点是数据展示和人和数据展示的交互。

通过旋转透视图，一份底表数据（明细数据）可以通过不同维度的组合，按不同视角生成另外一份汇总数据。



小知识：题图威尔士柯基犬。虽然属于小型犬，但性格非常稳健，完全没有一般小型犬的神经质，是非常适合小孩的守护犬。它们的胆子很大，也相当机警，能高度警惕地守护家园，是最受欢迎小型护卫犬之一。有着小狗的外表，却拥有大狗的灵魂。——摘自百度百科。

[Read more](#)
