

第四章 高级分析_排序与筛选

一、本章内容

1、排序

- ①手动排序
- ②计算排序

2、筛选器

- ①基本筛选器
- ②上下文筛选器
- ③通配符筛选器
- ④条件筛选器
- ⑤顶部筛选器

排序和筛选都是对数据的再次处理。

二、排序

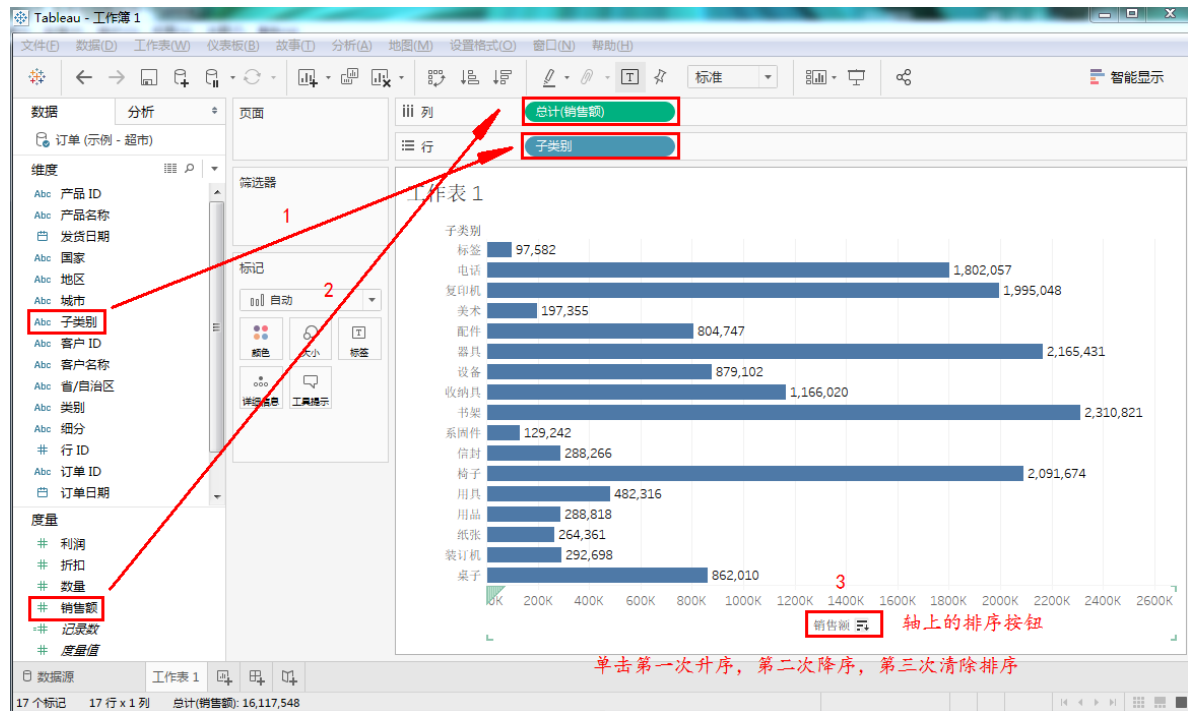
排序：指按照指定的度量对数据进行排序。

数据排序是数据分析的一个非常重要的特征。Tableau允许对属于维度的字段进行排序。

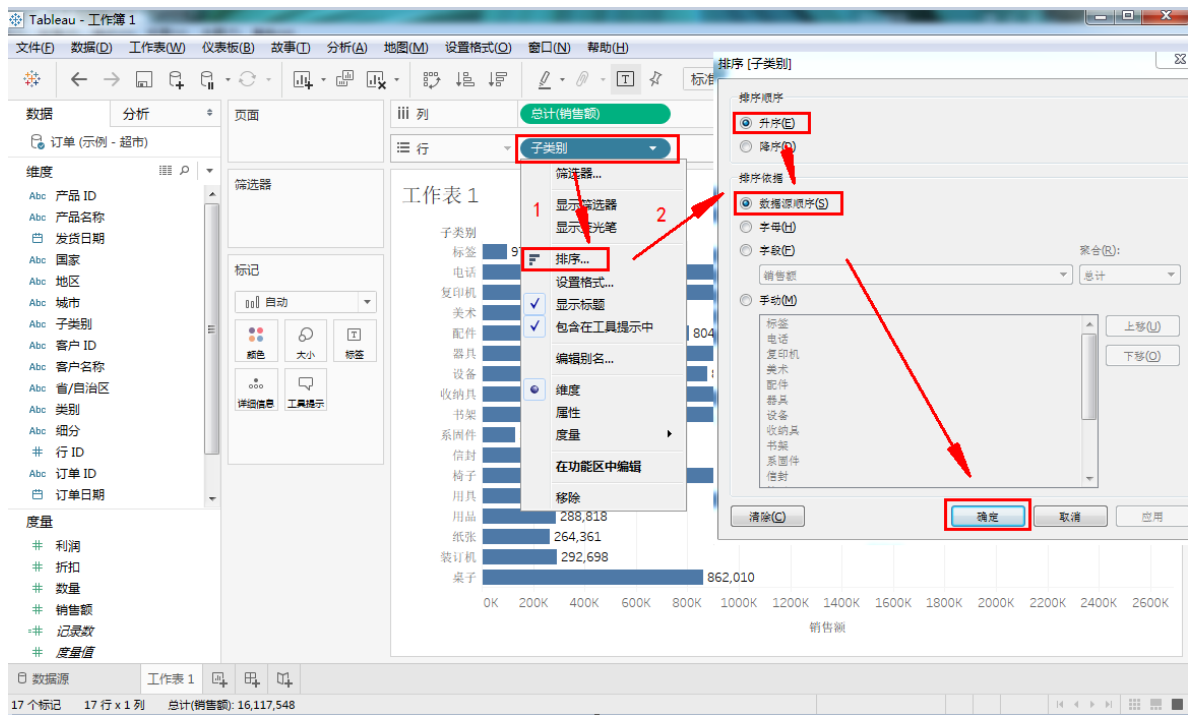
1、手动排序

手动排序：通过特定方式将他们彼此相邻拖动来重新排列维度字段的顺序。（系统默认的排序方式）

①通过轴上的排序按钮进行排序



②对可视化项中的特定字段进行排序



备注：

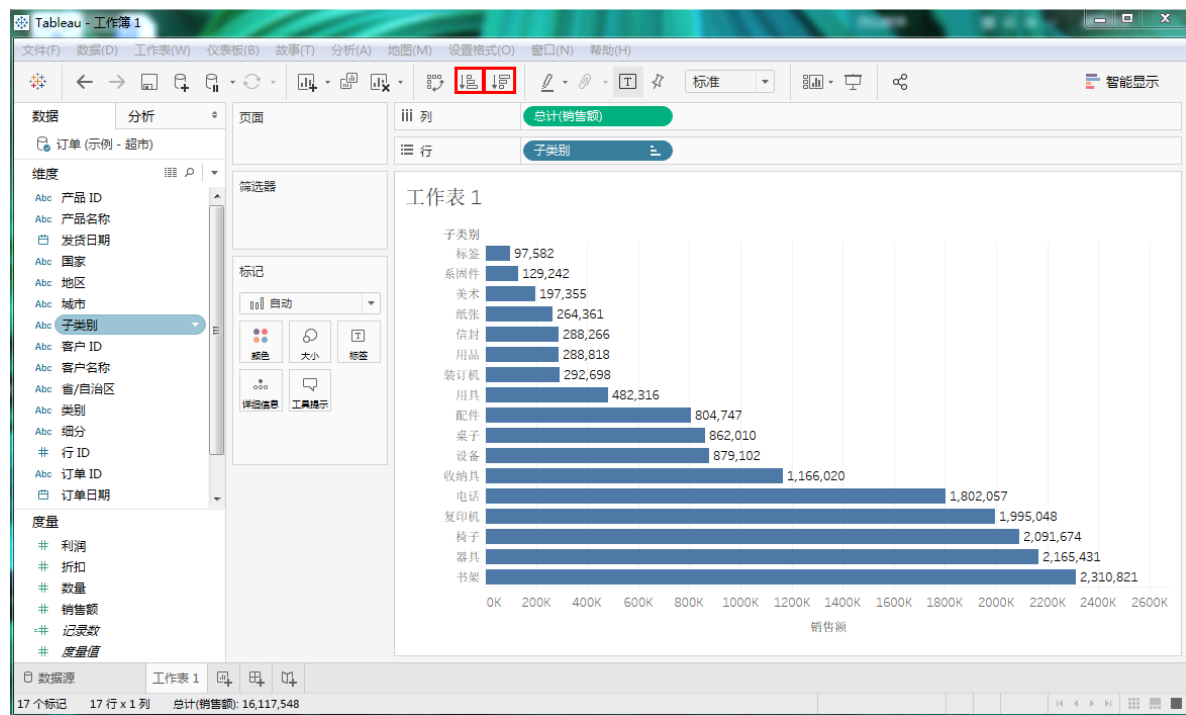
数据源顺序 — 数据源对数据进行自然排序的顺序。

字母 - 按字母顺序对数据进行排序。

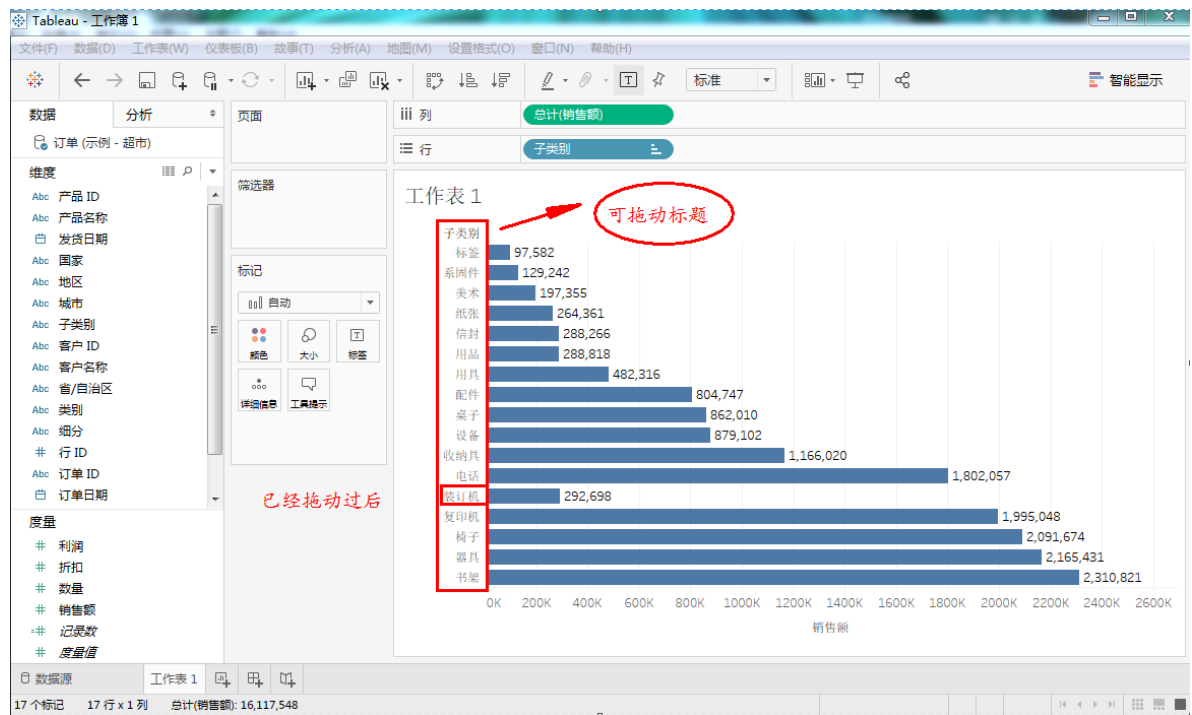
字段 — 基于另一个字段的值对数据进行排序。

手动 - 允许您按喜欢的顺序对数据进行排序。

③使用工具栏对数据进行排序

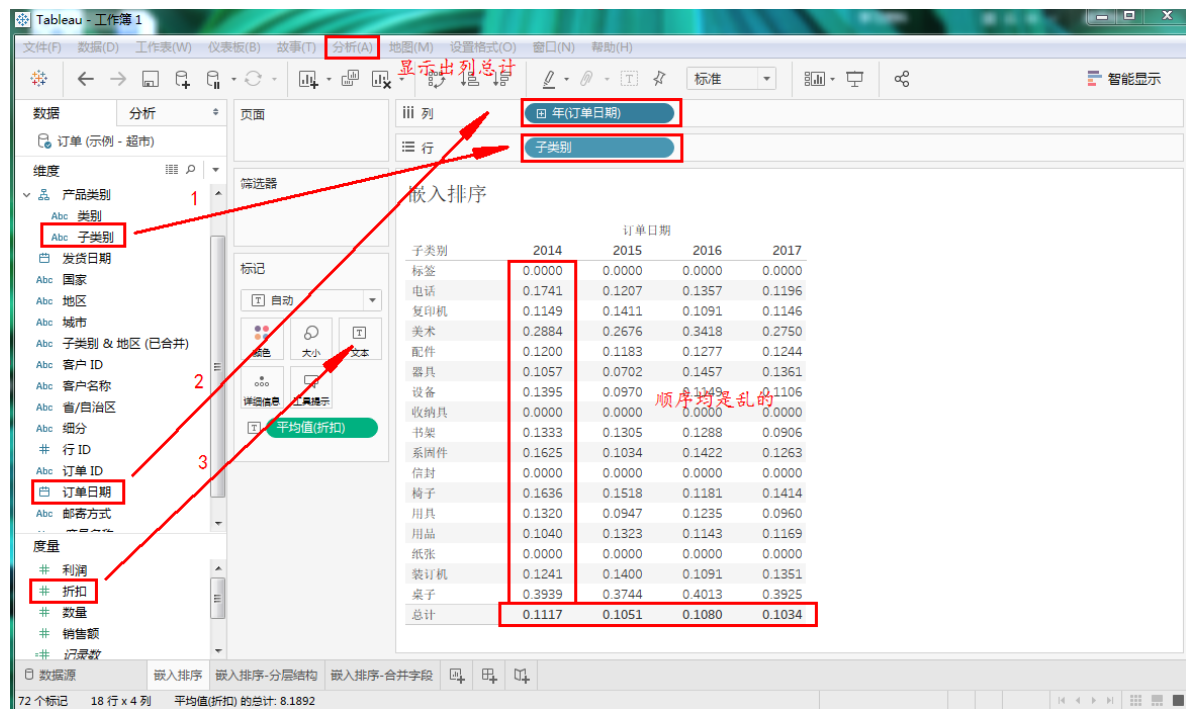


④使用标题对数据进行手动排序

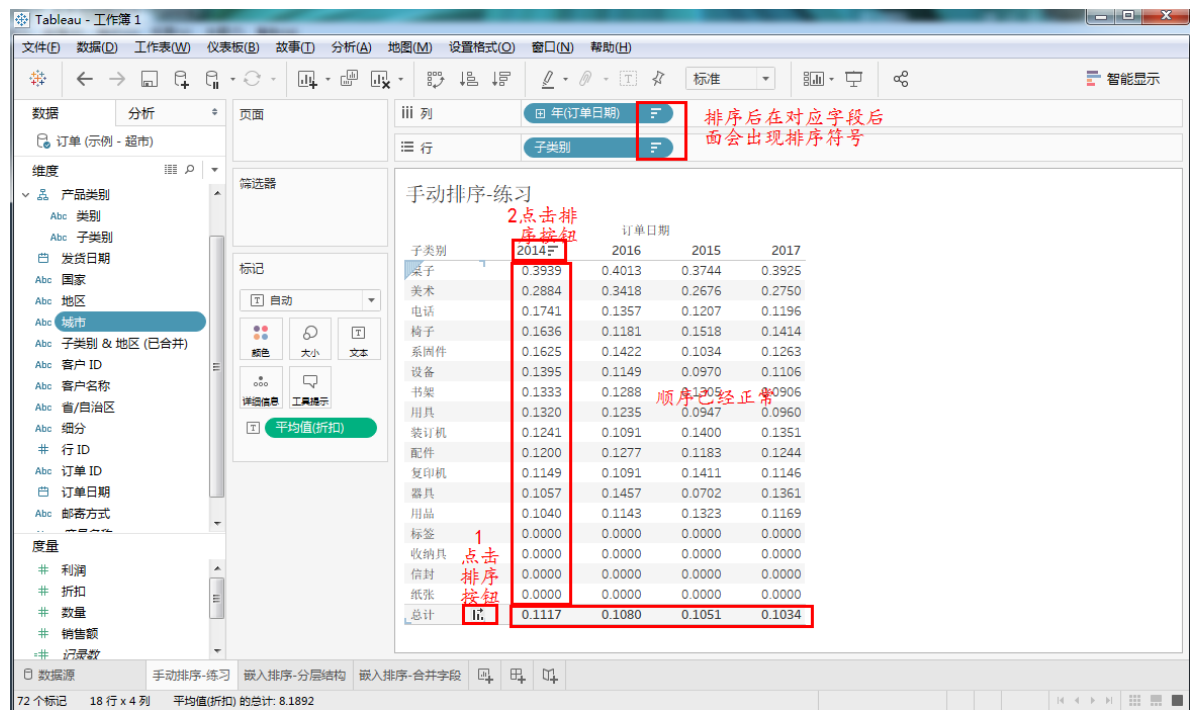


综合练习：对文本表的行和列进行排序，年份具有最高的平均折扣和某一年度哪些产品平均折扣最高（2014年为例）

①构建原始数据



②进行手工排序

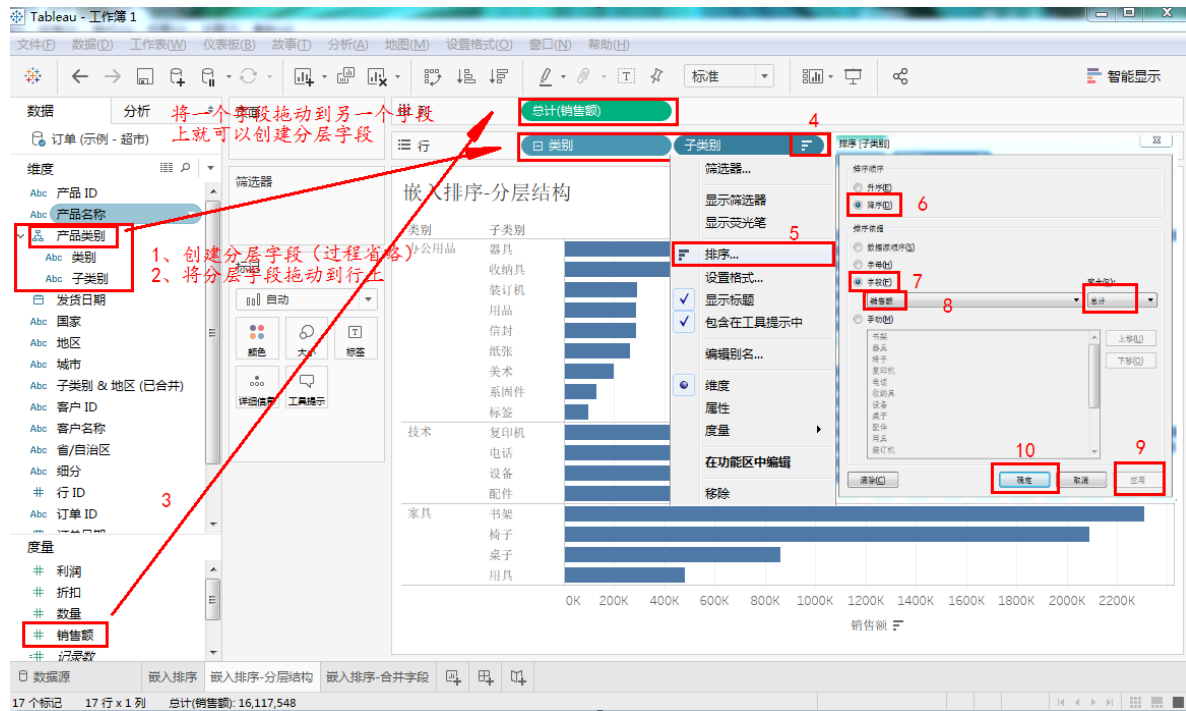


2、计算排序

计算排序：也称为嵌入排序，不使用系统定义好的排序规则，而使用自定义排序规则达到想要的效果。

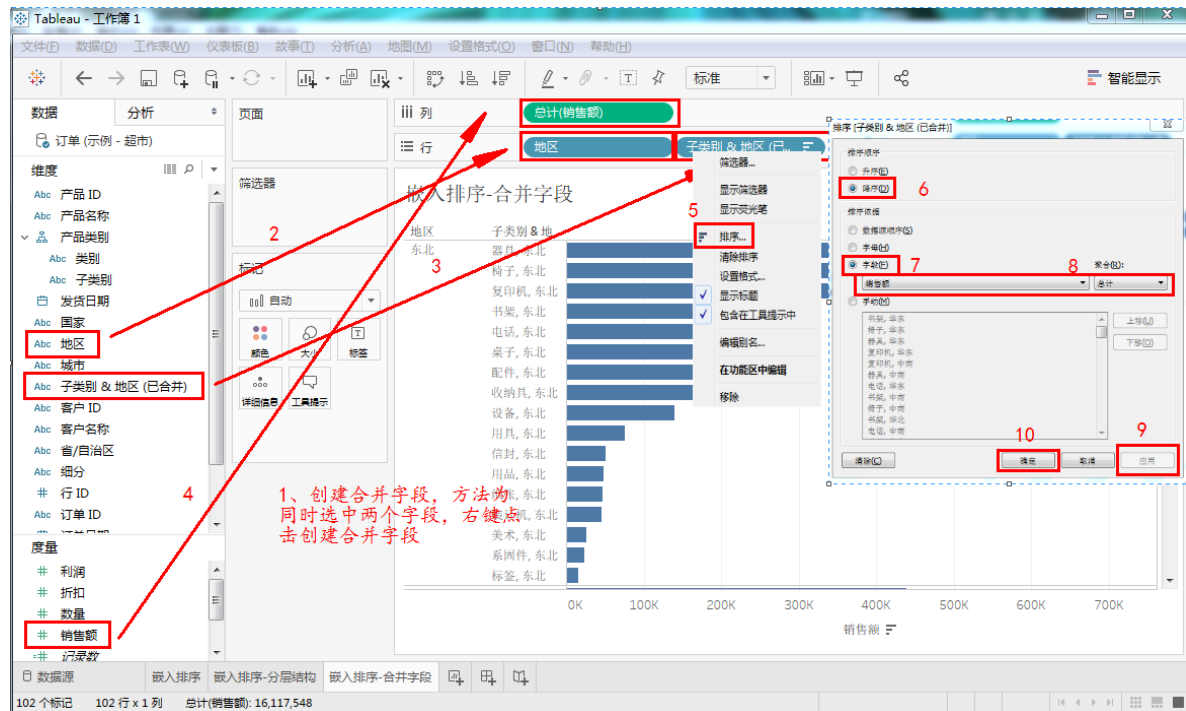
①对分层字段进行排序

例如：想要了解每种产品中哪种子类产品的销售额最高。



②对合并字段进行排序

例如：想要了解每个区域中哪个子类产品的销售额最高。



三、筛选器

要了解筛选器就要首先了解什么是筛选，顾名思义，筛选就是按条件查询。任何的数据分析和可视化工作都涉及到数据筛选。Tableau中有很多筛选器来满足这些需求。比如度量筛选器提供数字计算和比较，维度筛选器提供从列表中选择字符串或使用自定义值列表等等。筛选器通常也被称作为过滤器。

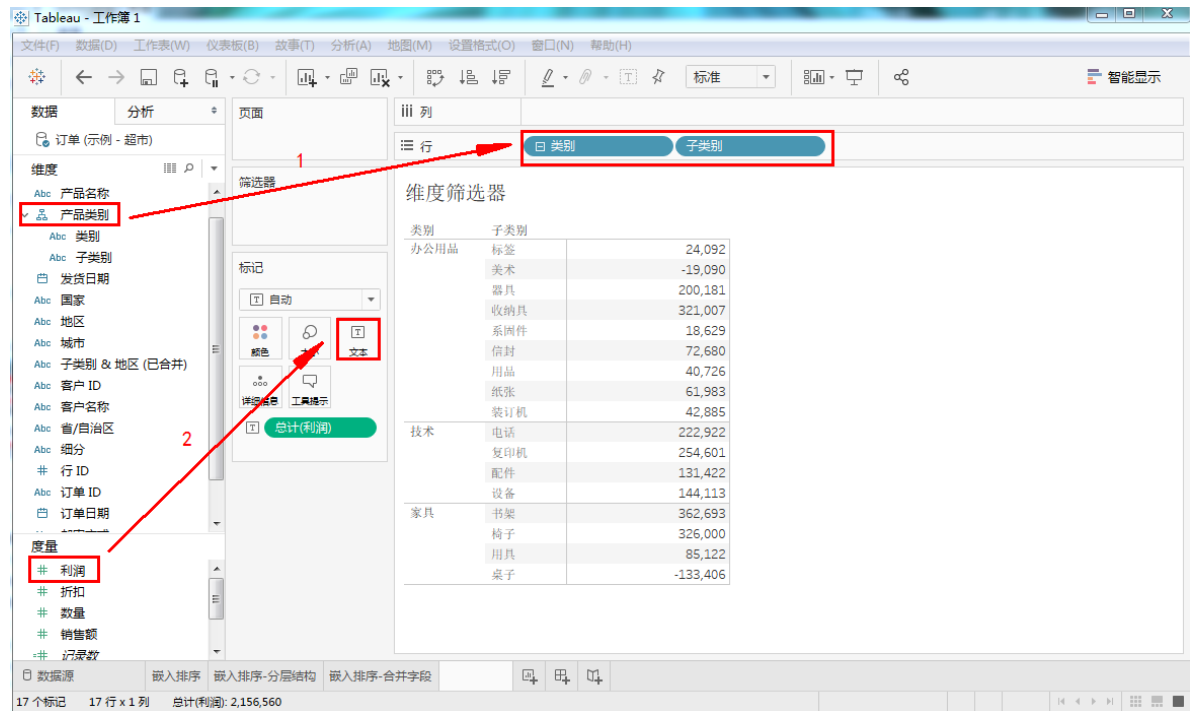
1、基本筛选器

A、维度筛选器

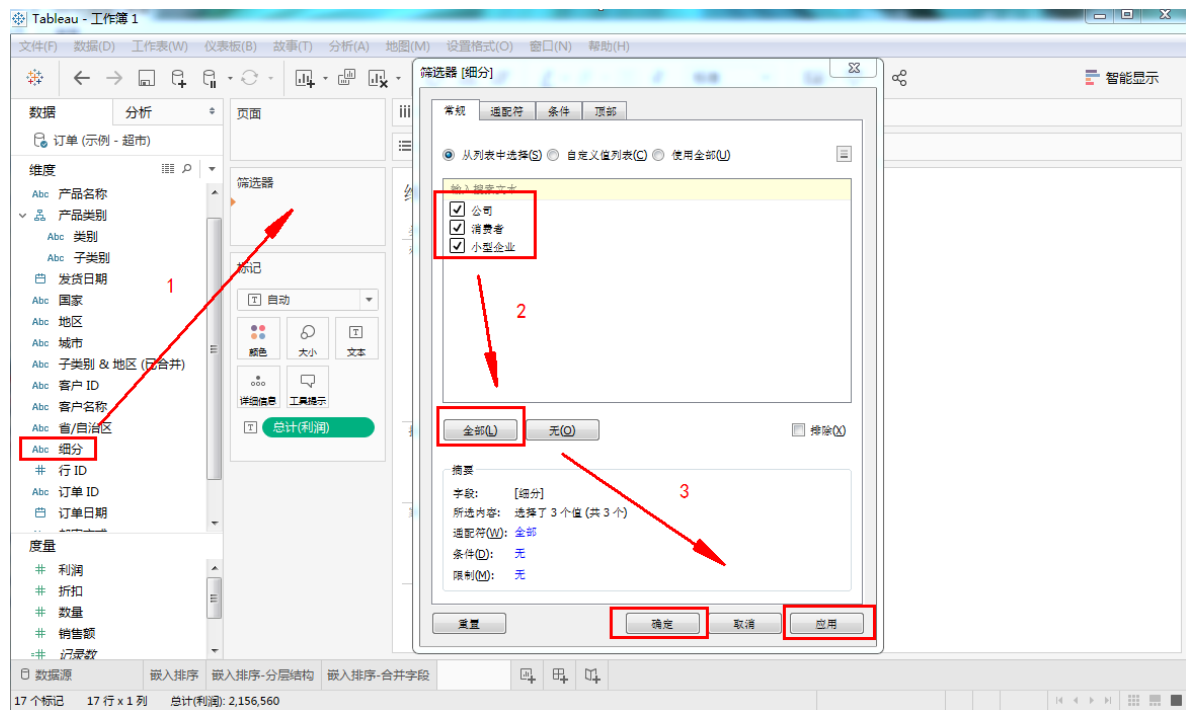
维度筛选器：应用于维度字段的筛选器。

例如：根据客户模式（细分字段---公司，消费者，小型企业）显示每个子类别产品的利润情况。

第一步：创建原始数据



第二步：创建筛选器



第三步：编辑筛选器和快速筛选

Tableau - 工作簿 1

文件(F) 数据(D) 工作表(W) 仪表板(B) 故事(T) 分析(A) 地图(M) 设置格式(O) 窗口(N) 帮助(H)

数据 分析

维度 产品类别

度量 利润

筛选器

维度筛选器

编辑筛选器...

显示筛选器

清除筛选器

添加到上下文

应用于工作表

创建集...

维度

属性

度量

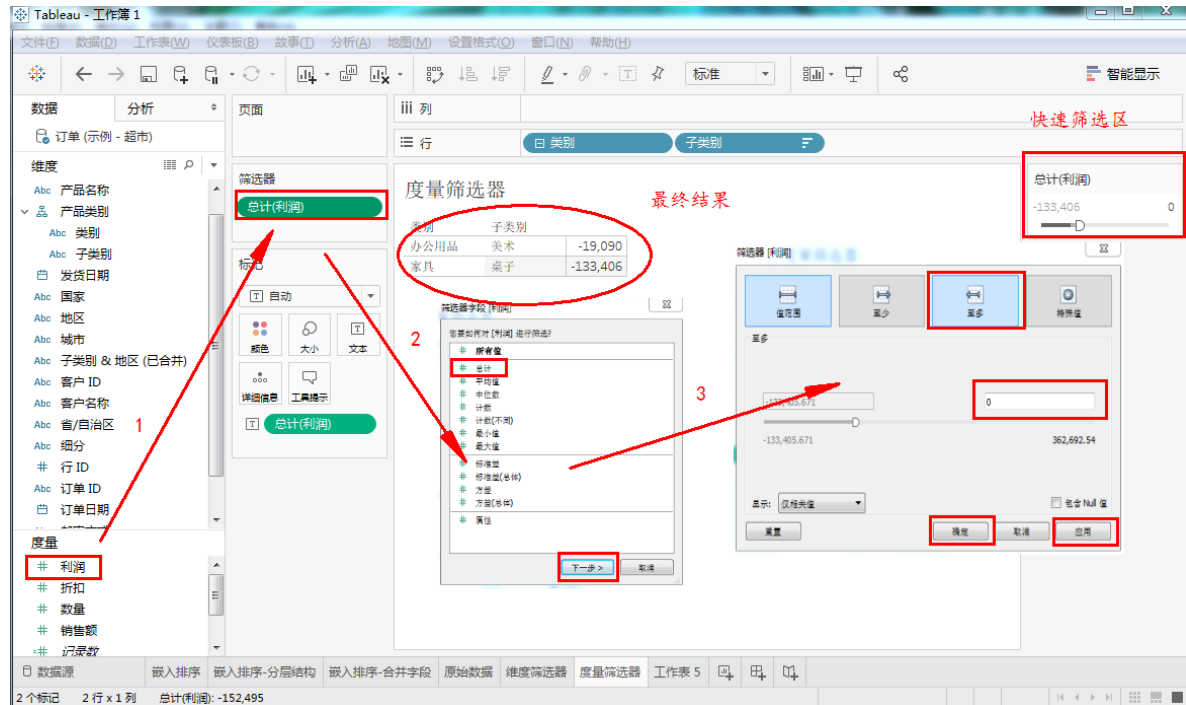
移除

进入编辑界面，与创建时相同

会在右侧生成一个快捷的筛选器

17 个标记 17 行 x 1 列 总计(利润): 1,096,606

第二步：创建筛选器



C、日期筛选器

日期筛选器：应用于日期字段的筛选器。

例如：统计出2016和2017年的产品子类别销售情况。

The screenshot shows the Tableau interface with the 'Date Filter' dialog box open. The dialog box is titled '筛选器 (订单日期)' and contains the following elements:

- 1**: The 'Date Filter' button in the 'Filters' shelf.
- 2**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 3**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 4**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 5**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 6**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 7**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 8**: The 'Date Filter' button in the 'Date Filter' dialog box.
- 9**: The 'Date Filter' button in the 'Date Filter' dialog box.

The resulting data table is shown below:

类别	子类别	2016	2017
办公用品	收纳具	86,575	113,641
	器具	29,917	56,832
	信封	17,459	24,203
	纸张	18,162	19,673
	装订机	11,352	12,776
	用品	10,416	12,941
	标签	6,490	7,663
	系固件	5,444	6,291
技术	美术	-10,288	-4,458
	电话	76,132	91,397
	复印机	100,423	63,988
	设备	40,070	50,352
家具	配件	37,939	35,638
	书架	97,664	135,175
	椅子	91,847	92,738
	用具	22,365	32,342
	桌子	-19,062	-66,200

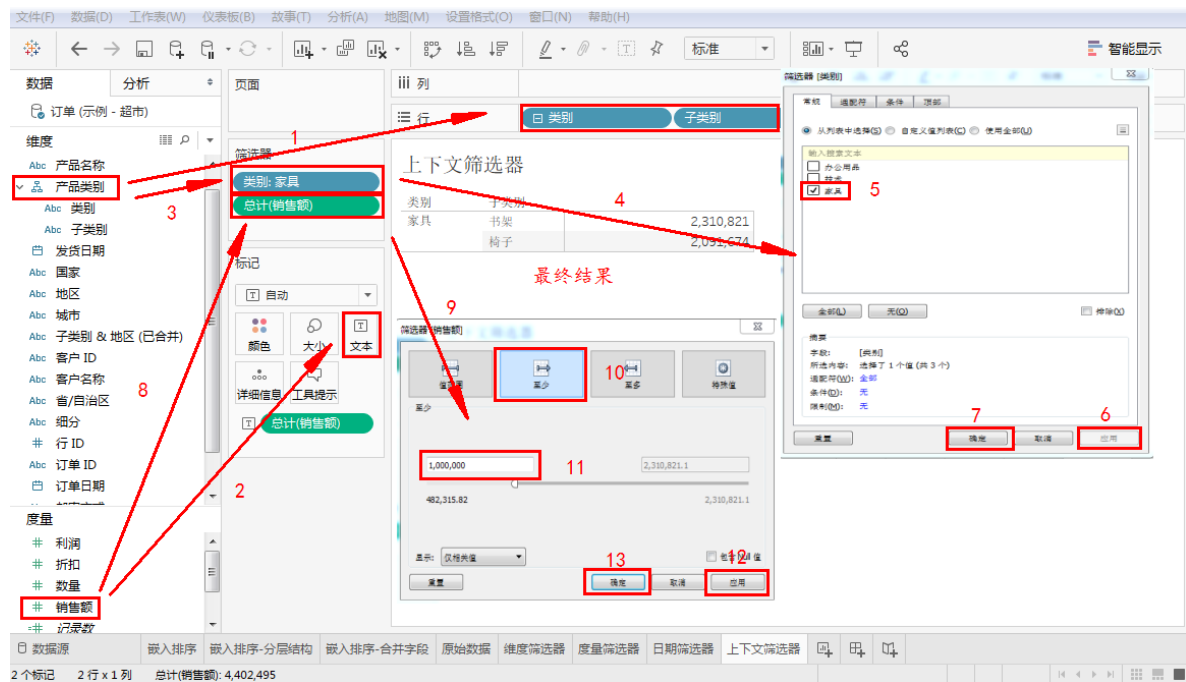
2、上下文筛选器

上下文筛选器：第二个筛选器只处理第一个筛选器返回的记录。通常第二个筛选器就称为依赖筛选器，因为它只能处理通过上下文筛选器的数据。

使用目的：1、提高性能 — 如果您设置了大量筛选器或具有一个大型数据源，则查询速度可能会很慢。可以设置一个或多个上下文筛选器来提高性能。

2、创建相关数字筛选器或“前 N 个”筛选器 — 可以设置一个上下文筛选器，以便仅包含相关数据，然后设置数字筛选器或“前 N 个”筛选器。

例如：查看家具类产品且销售额大于100万的产品有哪些（子类别）。



注意：将上下文筛选器视为独立筛选器；您设置的任何其他筛选器只处理通过上下文筛选器传递的数据。仅在上下文筛选器能够明显限制数据集大小时，才使用上下文筛选器。如果您要设置会显著减小数据集大小并将用于许多数据视图的筛选器，则应将这些筛选器设置为上下文筛选器。

3、通配符筛选器

通配符筛选器：将星号或者字符作为筛选（匹配）的值去匹配表中的所有值。

例如：查看产品名称以“罗技”打头的并且销售额大于5000的产品名称。

第一步：构建原始数据省略

第二步：筛选出所有产品名称以“罗技”打头的产品

The screenshot shows the Tableau interface with the following components and annotations:

- 1**: Points to the '产品名称' (Product Name) field in the '列' (Columns) shelf.
- 2**: Points to the '销售额' (Sales) field in the '度量' (Measures) shelf.
- 3**: Points to the '产品类别' (Product Category) field in the '维度' (Dimensions) shelf.
- 4**: Points to the '通配符筛选器' (Wildcard Filter) dialog box.
- 5**: Points to the '匹配模式' (Match Mode) section in the dialog box.
- 6**: Points to the '应用' (Apply) button in the dialog box.

The '通配符筛选器' dialog box shows the following configuration:

- 匹配模式** (Match Mode): ☒ 包含 (Contains)
- 匹配值** (Match Value): 罗技 (Logitech)
- 匹配范围** (Match Range): ☒ 为文本包括所有单词 (For text, include all words)

The main view displays a table of products with columns: 类别 (Category), 子类别 (Sub-category), 产品名称 (Product Name), and 销售额 (Sales). The table lists various Logitech products and their sales figures.

类别	子类别	产品名称	销售额
技术	配件	罗技记忆卡, 回收	2,172
		罗技记忆卡, 耐用	3,916
		罗技记忆卡, 实惠	5,511
		罗技记忆卡, 实惠	2,026
		罗技键盘, 回收	6,351
		罗技键盘, 可编程	3,451
		罗技键盘, 耐用	2,374
		罗技键盘, 实惠	2,975
		罗技键盘, 回收	2,265
		罗技键盘, 可编程	1,562
		罗技键盘, 耐用	3,691
		罗技键盘, 实惠	3,113
		罗技路由器, 回收	3,432
		罗技路由器, 可编程	7,896
		罗技路由器, 耐用	11,551
		罗技路由器, 实惠	2,539
		罗技闪存驱动器, 回收	1,269
		罗技闪存驱动器, 可编程	3,197
		罗技闪存驱动器, 耐用	516
		罗技闪存驱动器, 实惠	1,879

第三步：筛选出销售额大于5000的产品

The screenshot shows the Tableau interface with the following components:

- Left Pane (Data):** Dimensions (产品ID, 产品类别, 类别, 子类别, 产品名称, 发货日期, 国家, 地区, 城市, 子类 & 地区 (已合并), 客户ID, 客户名称, 省/自治区, 细分, 行ID, 订单ID, 订单日期) and Measures (利润, 折扣, 数量, 销售额, 订单数).
- Central View:** A table showing product details and sales amounts. The table has columns for Product Name, Sales Amount, and a calculated field for Sales Amount. The data is as follows:

产品名称	销售额	计算字段: 销售额
罗技 记忆卡, 耐用	5,511	5,511
罗技 键盘, 回收	6,361	6,361
罗技 路由器, 可编程	7,896	7,896
罗技 路由盒, 耐用	11,551	11,551

Right Pane (Filter): The 'Sales' field is selected. The aggregation is set to 'Sum'. The filter is applied to the 'Sales' field.

Annotations:

- 1 (6) 编辑筛选器... (Edit Filter...)
- 2 (7) 销售额 (Sales)
- 3 (8) OK

Bottom Status Bar: 4 个标记 4 行 x 1 列 总计(销售额): 31,319

4、条件筛选器

条件筛选器：对已存在的筛选器应用一些附加条件，使之达到我们预期的结果。

例如：查看家具类中销售额大于100万并且利润大于32.6万的子类产品。

第一步：构建原始数据省略

第二步：筛选出家具类

Tableau - 工作簿 1

数据 分析 页面

订单 (示例 - 超市)

维度

- 产品名称
- 产品类别
- 子类别
- 国家
- 地区
- 城市
- 子类别 & 地区 (已合并)
- 客户 ID
- 客户名称
- 省/自治区
- 细分
- 行 ID
- 订单 ID
- 订单日期

度量

- 利润
- 折扣
- 数量
- 销售额
- 记录数

条件筛选器

类别 子类别

类别	子类别	销售额
家具	书桌	2,314
家具	椅子	361
家具	椅子	2,091
家具	椅子	324
家具	椅子	481
家具	椅子	81
家具	椅子	861
家具	椅子	-131

最终结果

筛选器 (类别)

常规 适配器 条件 顶部

从列表中选择(S) 自定义值列表(C) 使用全部(U)

输入搜索文本

- ☐ 办公用品
- ☐ 技术
- ☒ 家具

全部(L) 无(O) 清除(O)

摘要

字段: [类别]

所选内容: 选择了 1 个值 (共 3 个)

适配器(U): 全部

条件(O): 无

限制(O): 无

重置 确定 取消 应用

4 个标记 4 行 x 1 列 总计(销售额): 5,746,822

第三步：筛选出销售额大于100万的子类产品（只是为了讲解单条件过滤）

1. 在维度列表中选中“子类别”

2. 在筛选器中选中“子类别”

3. 打开“筛选器 (子类别)”对话框

4. 选择聚合方式为“求和”

5. 选择聚合函数为“求和”

6. 选择筛选条件为“大于”

7. 输入筛选值 1,000,000

8. 点击“确定”

9. 点击“应用”

默认选中所有子类

此操作完成后即可获得销售额大于100万的家具子类

注意：此种方式只能做到单条件或者说是单范围的过滤

类别	子类别	销售额
家具	书架	2,310,821
		362,693
		2,091,674
	椅子	326,000

2 个标记 2 行 x 1 列 总计(销售额): 4,402,495

第四步：筛选出销售额大于100万并且利润大于32.6万的子类产品（满足多条件）

Tableau - 工作簿 1

文件(F) 数据(D) 工作表(W) 仪表板(B) 故事(T) 分析(A) 地图(M) 设置格式(O) 窗口(N) 帮助(H)

数据 分析

维度

- 产品类别
- 类别
- 子类别
- 发货日期
- 国家
- 地区
- 城市
- 子类别 & 地区 (已合并)
- 客户 ID
- 客户名称
- 省/自治区
- 细分
- 行 ID
- 订单 ID
- 订单日期

度量

- 利润
- 折扣
- 数量
- 销售额
- 行序号

筛选器

- 类别: 家具
- 子类别

条件筛选器

1 编辑筛选器...

2 产品 ID (计数)

3 值范围

4 最小值(0):

5 最大值(0):

6 筛选器(0)

7 筛选

8 取消

9 应用

得到最终结果

2,310,821

362,693

1 个标记 1 行 x 1 列 总计(销售额): 2,310,821

5、顶部筛选器

顶部筛选器：筛选器中的Top选项用于限制筛选器的结果集。通俗点讲就是找出Top多少的内容。

例如：查看家具类里面销售额最好的前三名子类产品。

第一步：构建数据省略

第二步：筛选出家具类省略

第三步：筛选出销售额前三名的子类产品

The screenshot shows the Tableau Desktop interface with the following components and annotations:

- Data Pane:** The '子类别' (Sub-category) field is selected under the '家具' (Furniture) category. An arrow labeled '1' points to this field.
- Filters Shelf:** The '子类别' field is added to the filter shelf. An arrow labeled '2' points to this field.
- Marks Card:** The '总计(销售额)' (Sum of Sales) measure is selected. An arrow labeled '3' points to this measure.
- Top Filter Dialog Box:** The 'Top' tab is selected. The 'Top of' dropdown is set to 'Sales' (销售额). The 'Number of Items' is set to 3. An arrow labeled '4' points to the 'Top of' dropdown. An arrow labeled '5' points to the 'Number of Items' dropdown. An arrow labeled '6' points to the 'Top of' dropdown. An arrow labeled '7' points to the 'Number of Items' dropdown. An arrow labeled '8' points to the 'OK' button.
- Table:** The table shows the top 3 sub-categories by sales for furniture: 家具 (2,310,821), 椅子 (2,091,674), and 桌子 (862,010). An arrow labeled '9' points to the '家具' row.

注意：如果结果不对，重新编辑筛选器，看看常规里是不是选中的“从列表中选择”，此选项未选中可能会导致结果不正确。

