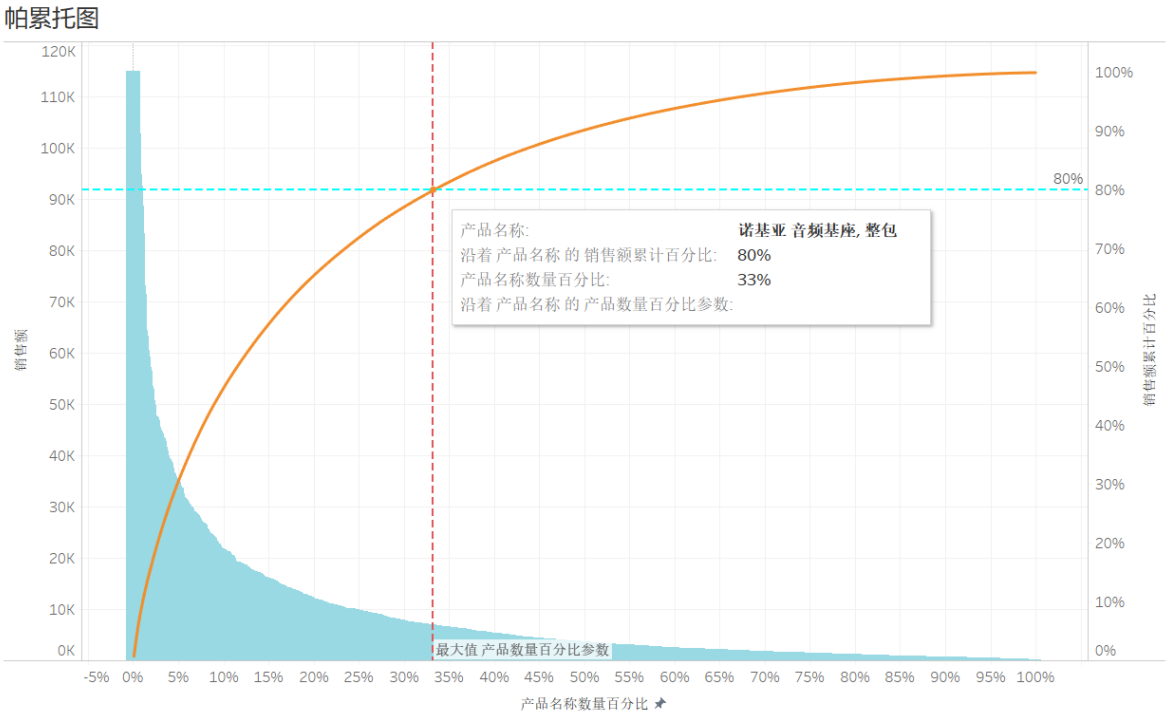


第五章：高阶可视化，满足你一切想象力

一、帕累托图

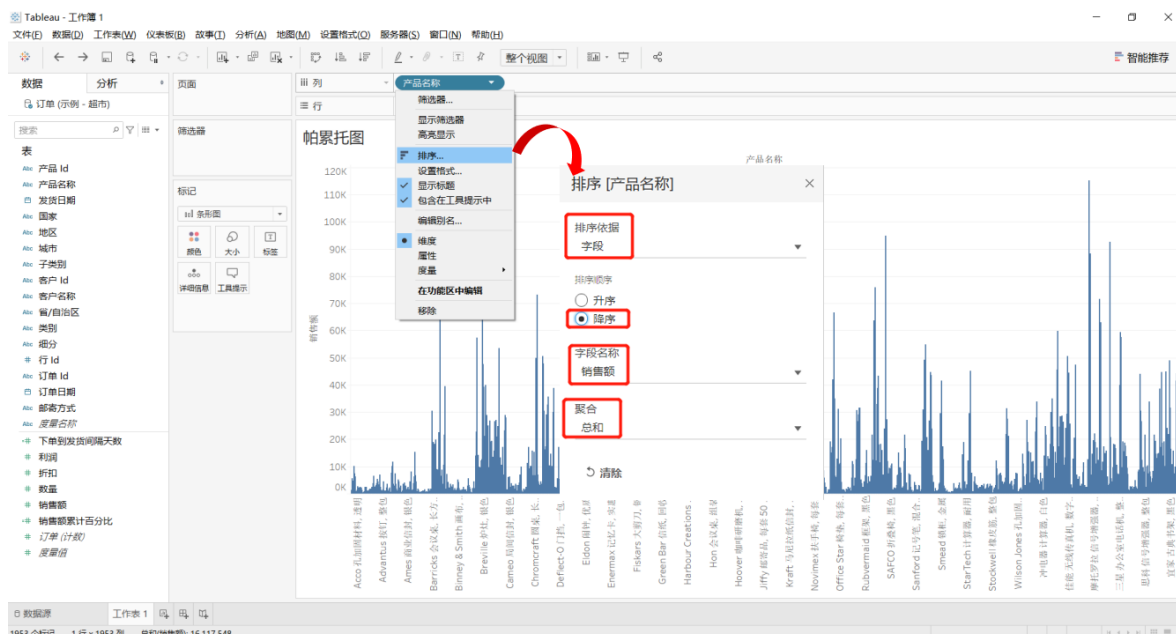
帕累托图是按照一定的类别根据数据计算出其分类所占的比例，用从高到低的顺序排列成矩形，同时展示比例累积和的图形，主要用于分析导致结果的主要因素。帕累托图与帕累托法则（又称为“**二八原理**”，即80%的结果是20%的原因造成的）一脉相承，通过图形体现两点重要信息：“至关重要的极少数”和“微不足道的大多数”。

接下来我们就通过帕累托图来挖掘《示例-超市》的订单数据，带大家迅速发现隐藏在数据中的重要信息。通过“二八原理”，分析贡献80%收入的产品比例，来评估该企业的销售的健康程度。

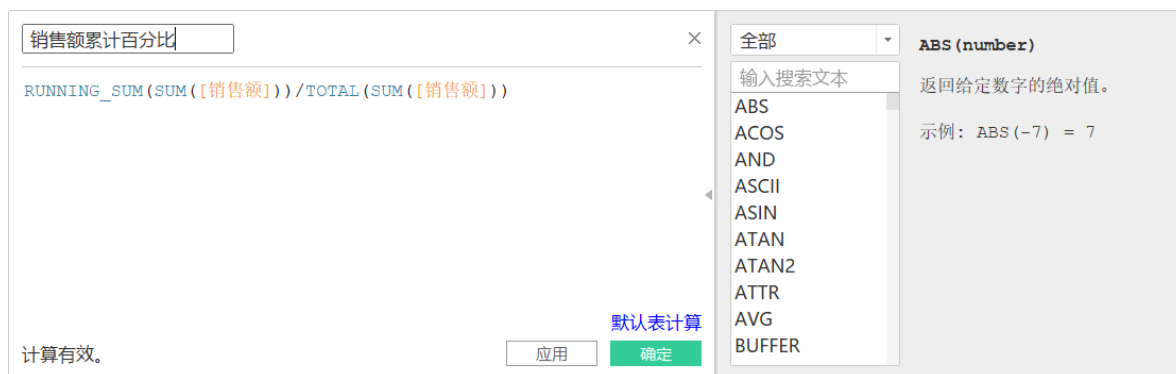


1.1 帕累托图应用场景

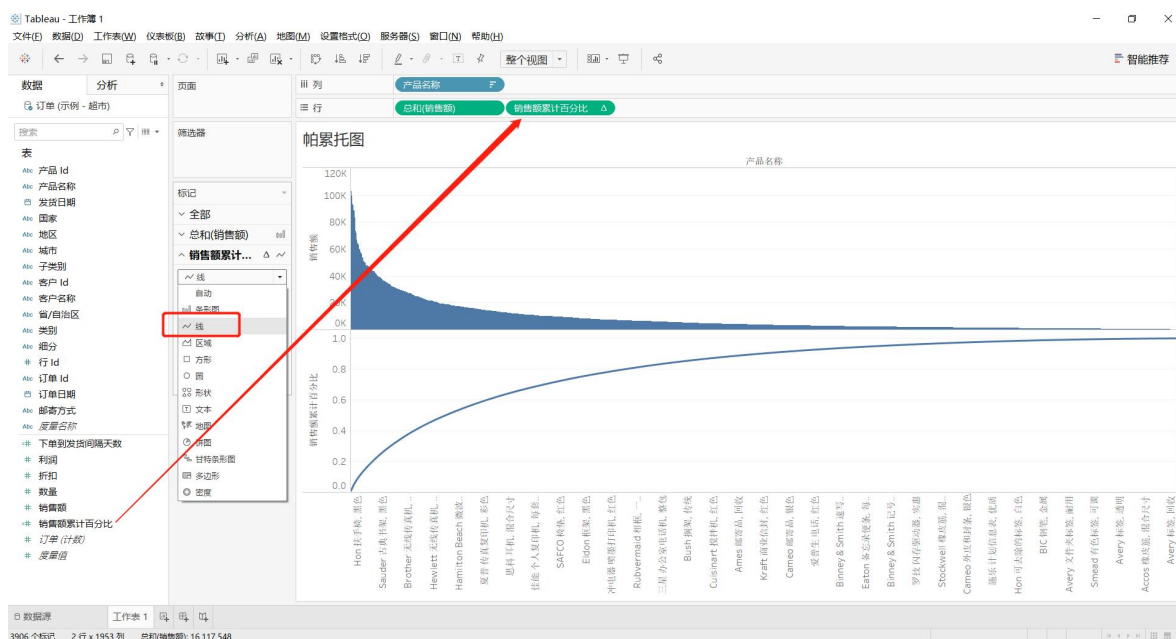
- **Step1:** 将“销售额”拖到功能区，产品名称拖到“列功能区”，点击下拉菜单，选择排序，依据“销售额”进行降序排序



- Step2: 创建计算字段“销售额累计百分比”

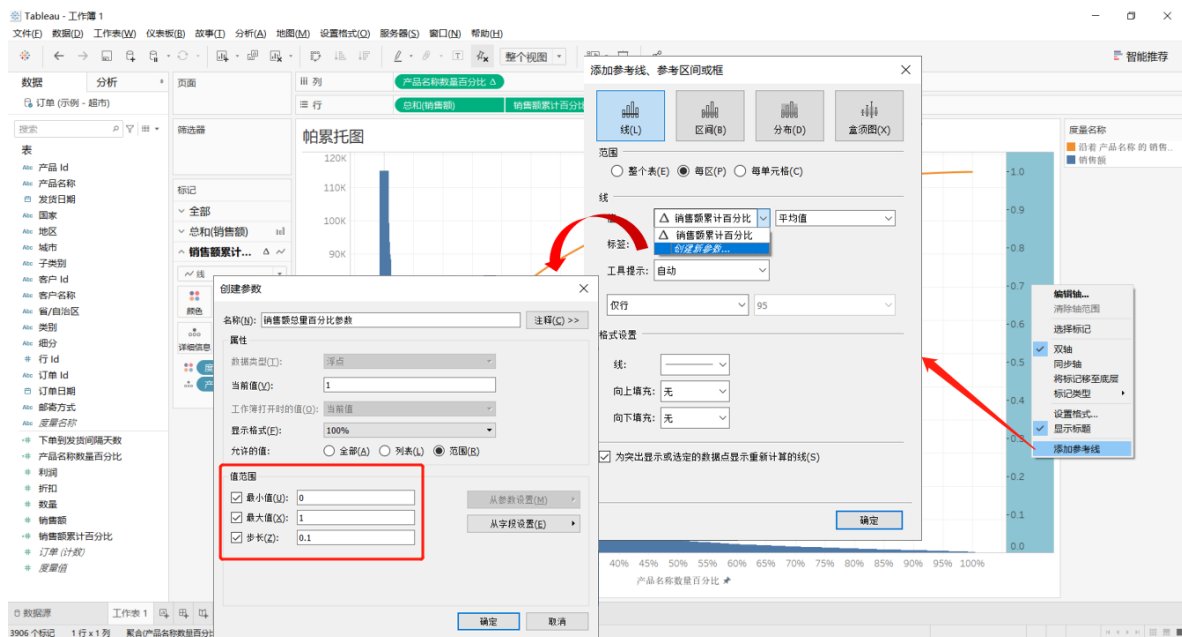


- Step3: 将“销售额累计百分比”拖到行功能区，点击下拉菜单，计算依据“产品名称”，形状选择线



- Step4: 点击“销售额累计百分比”下拉菜单，选择“双轴”

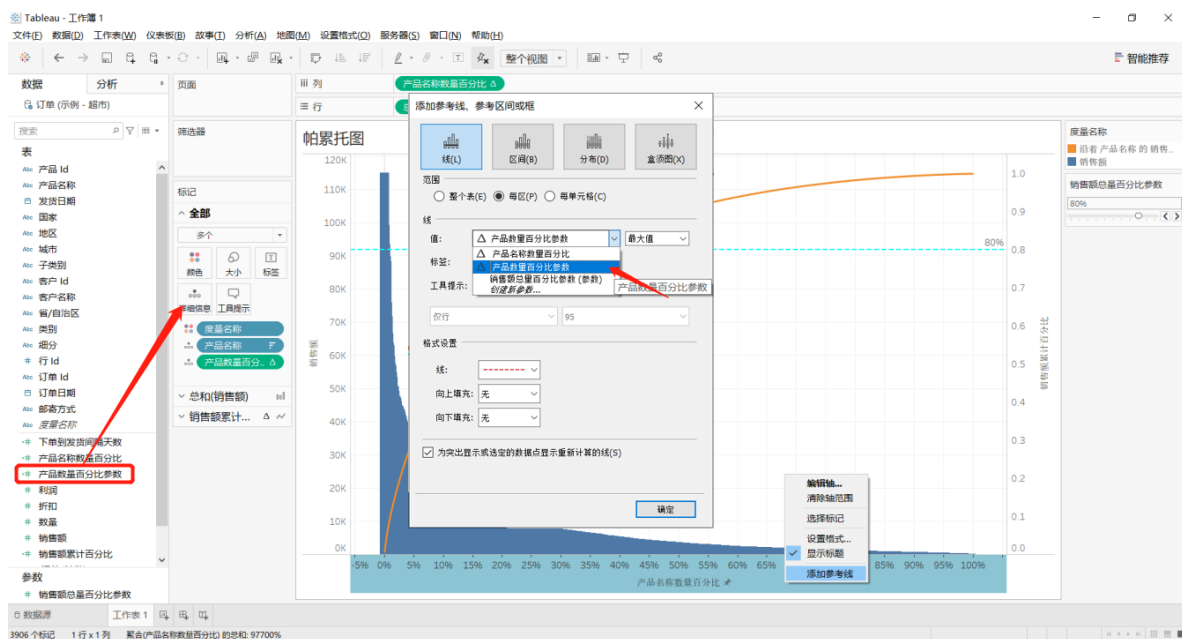
• Step7: 给纵坐标“销售额累计百分比”添加辅助参考线

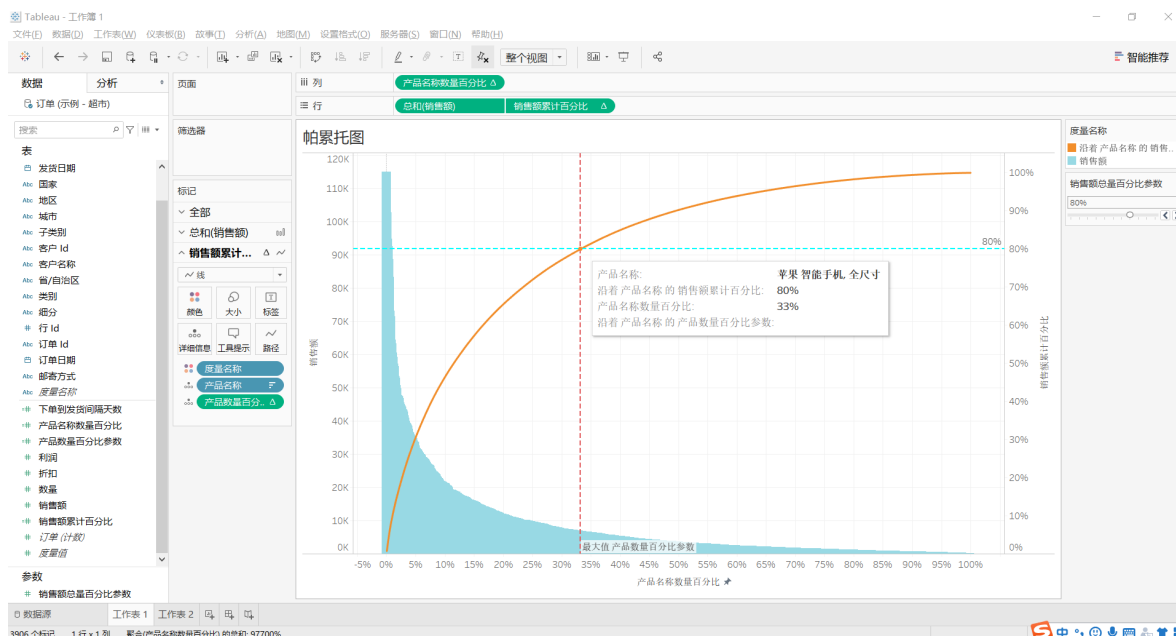


• Step8: 创建计算字段“产品名称数量百分比”



• Step9: 将创建的字段“产品名称数量百分比”拖进“标记卡 - 详细信息”中，然后给横坐标“产品名称数量百分比”添加辅助参考线

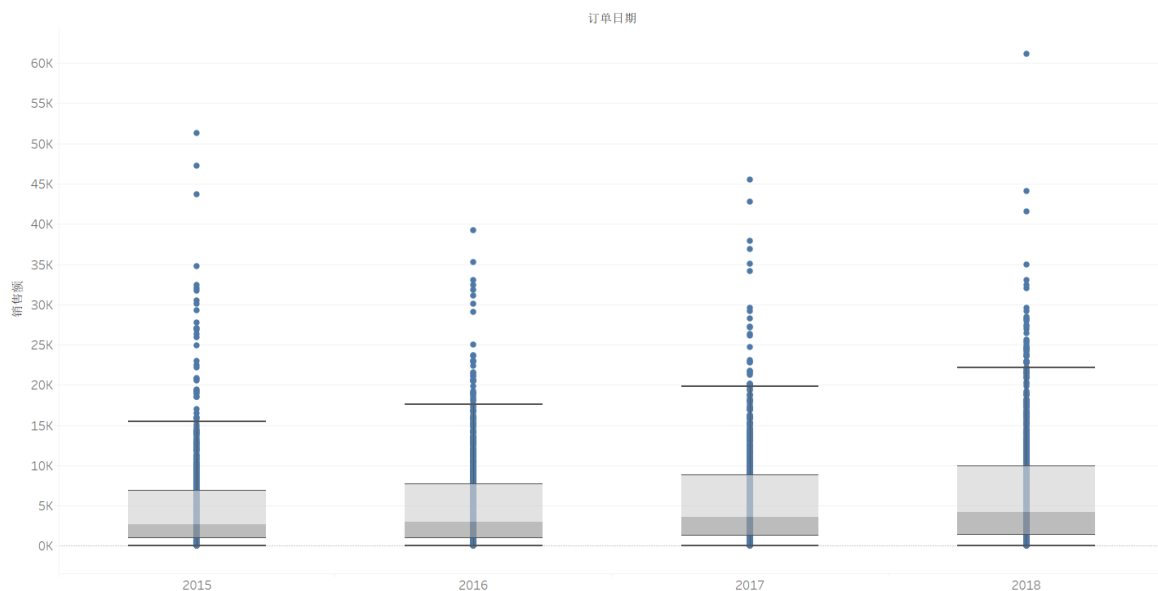




二、盒须图

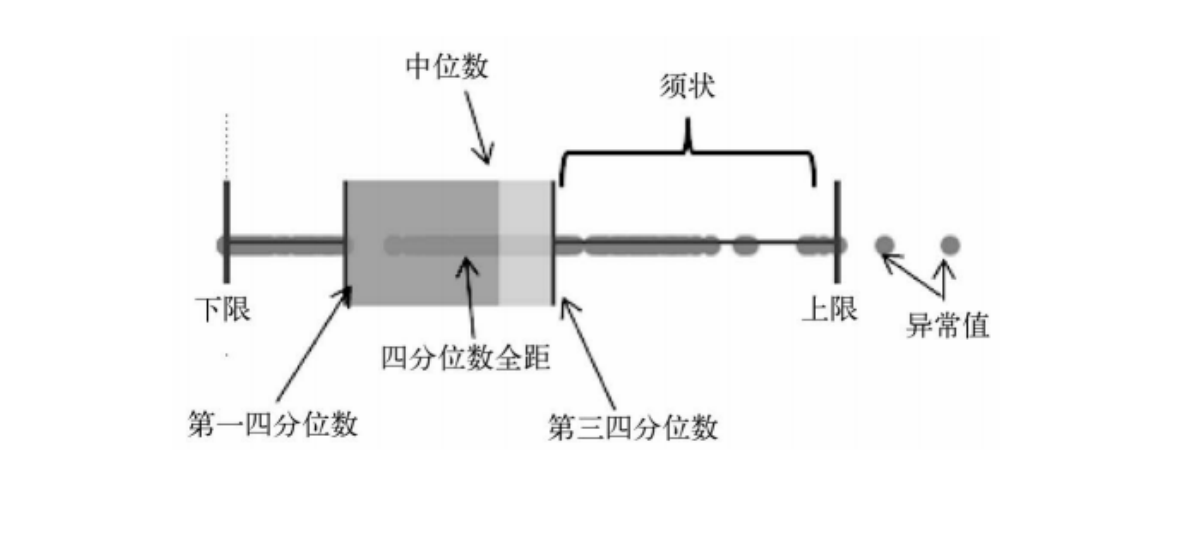
盒须图又叫箱线图，是一种常用的统计图形，用以显示数据的位置、分散程度、异常值等。箱线图主要包括6个统计量：**下限**、**第一四分位数**、**中位数**、**第三四分位数**、**上限**和**异常值**。通过绘制盒须图，观测数据在同类群体中的位置，可以知道哪些表现好，哪些表现差；比较四分位全距及线段的长短，可以看出哪些群体分散，哪些群体更集中。

盒须图



- **第一四分位数**：数据按照大小顺序排列，处于总观测数25%位置的数据
- **中位数**：数据按照大小顺序排列，处于中间位置，即总观测数50%的数据。
- **第三四分位数**：数据按照大小顺序排列，处于总观测数75%位置的数据为第三分位数
- **下限**：第一四分位数 - 1.5 * IQR
- **上限**：第三四分位数 + 1.5 * IQR
- **异常值**：在上限和下限之外的数据

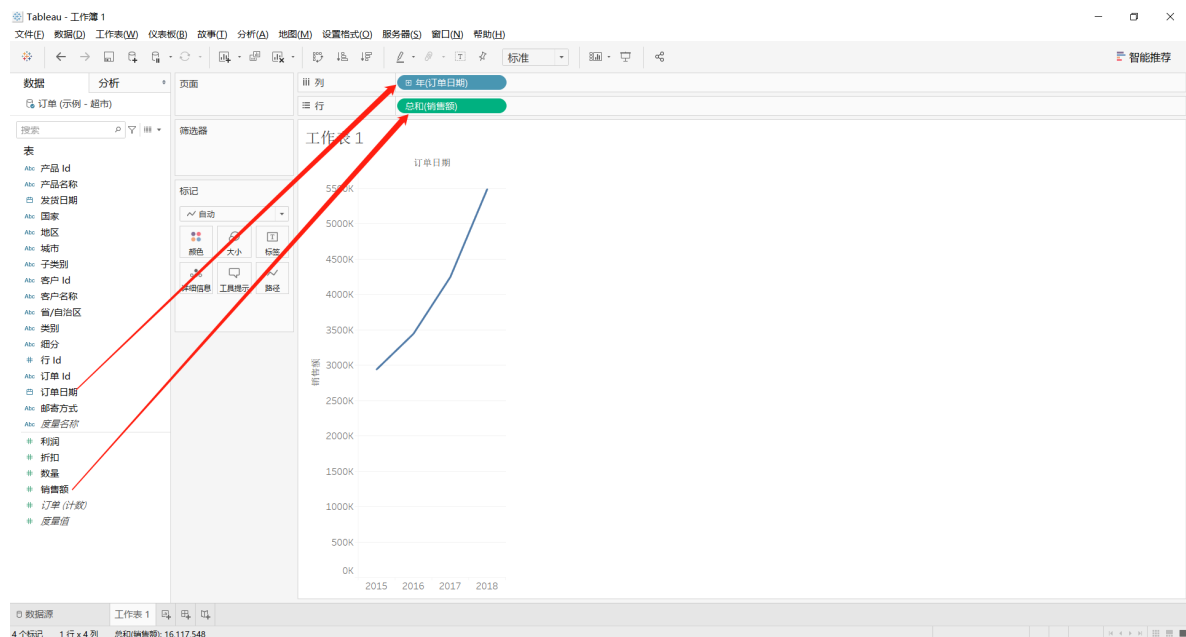
- **IQR**：四分位全距，即第三四分位数与第一四分位数之差



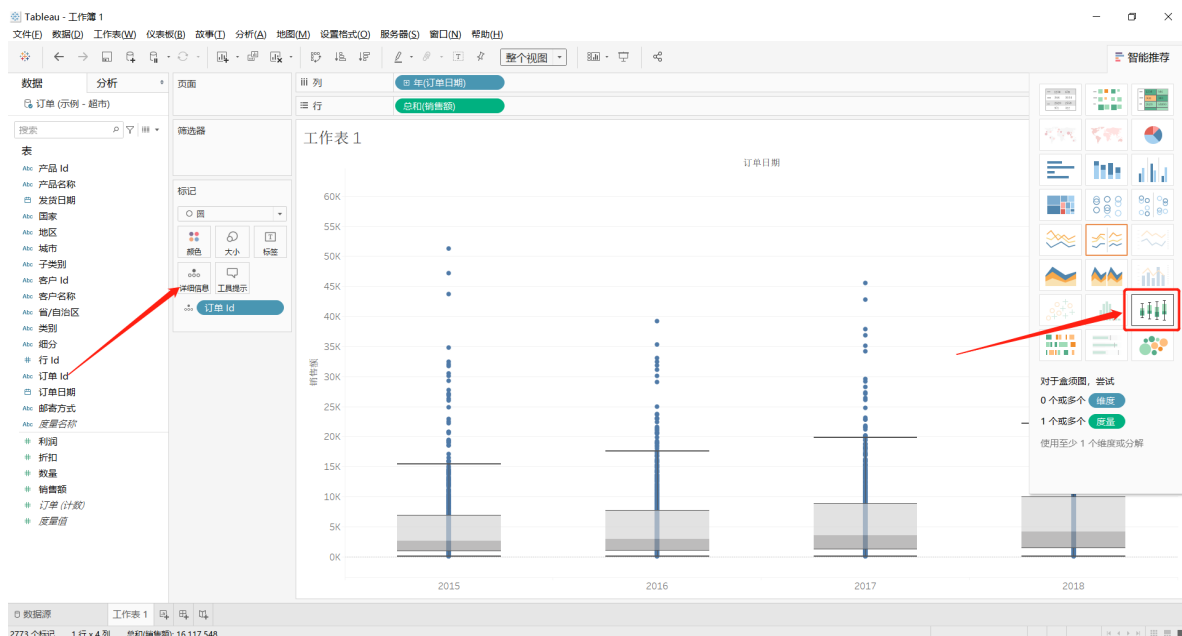
2.1 盒须图应用场景

如果我们想要对《示例-超市》的订单销售额数据进行深入分析，就可以通过构建盒须图分别对比销售额的分位值、上下限值在2015-2018年的变化趋势，从而能够很直观地发现销售额的变化规律。

- **Step1**：将“订单日期”拖进行列功能区，“销售额”拖进行功能区



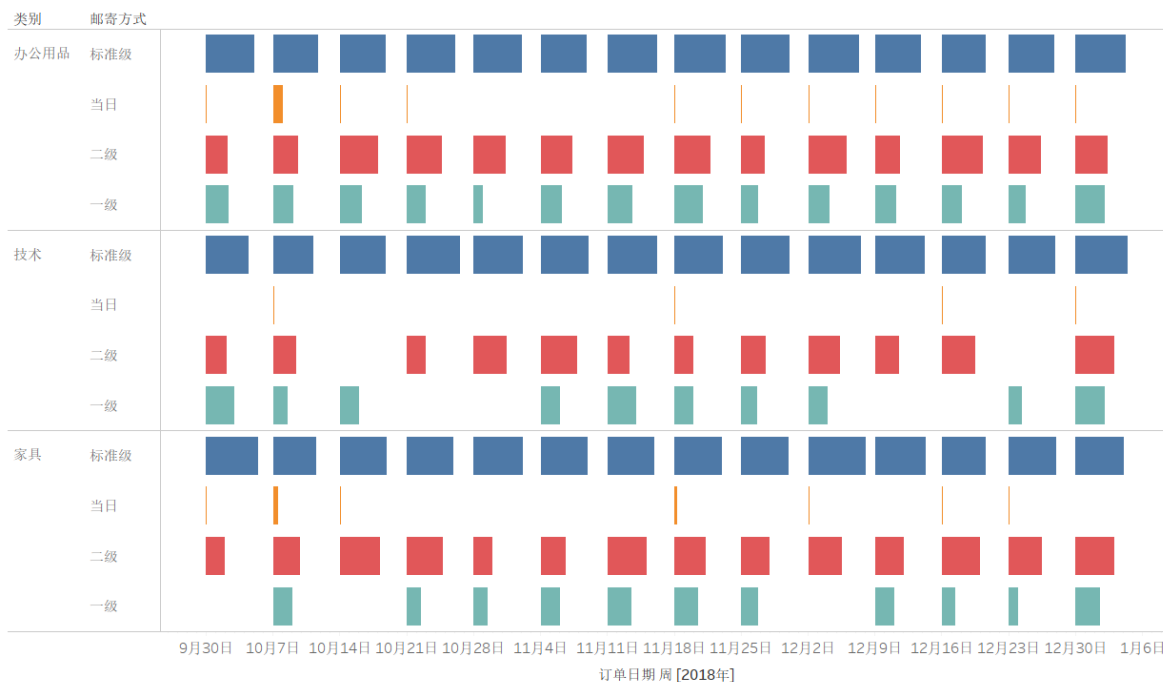
- **Step2**：将“订单ID”拖进“标记卡 - 详细信息”进行解聚，然后在智能推荐区域选择盒须图



三、甘特图

甘特图，又称横道图，是以图示的方式通过活动列表和时间刻度形象地表示出任何特定项目的活动顺序和持续时间。甘特图的横轴表示时间，纵轴表示活动（项目），线条表示在整个期间上该活动或项目的持续时间，因此可以用来比较与日期相关的不同活动（项目）的持续时间长短。甘特图也常用于显示不同任务之间的依赖关系，并被普遍用于项目管理中。

甘特图



3.1 甘特图应用场景

- **Step1:** 创建计算字段“下单到发货间隔天数”

下单到发货间隔天数

×

全部

输入搜索文本

COS

COT

COUNT

COUNTD

COVAR

COVARP

DATE

DATEADD

DATEDIFF

DATENAME

DATEDIFF('day',[订单日期],[发货日期])

计算有效。

应用

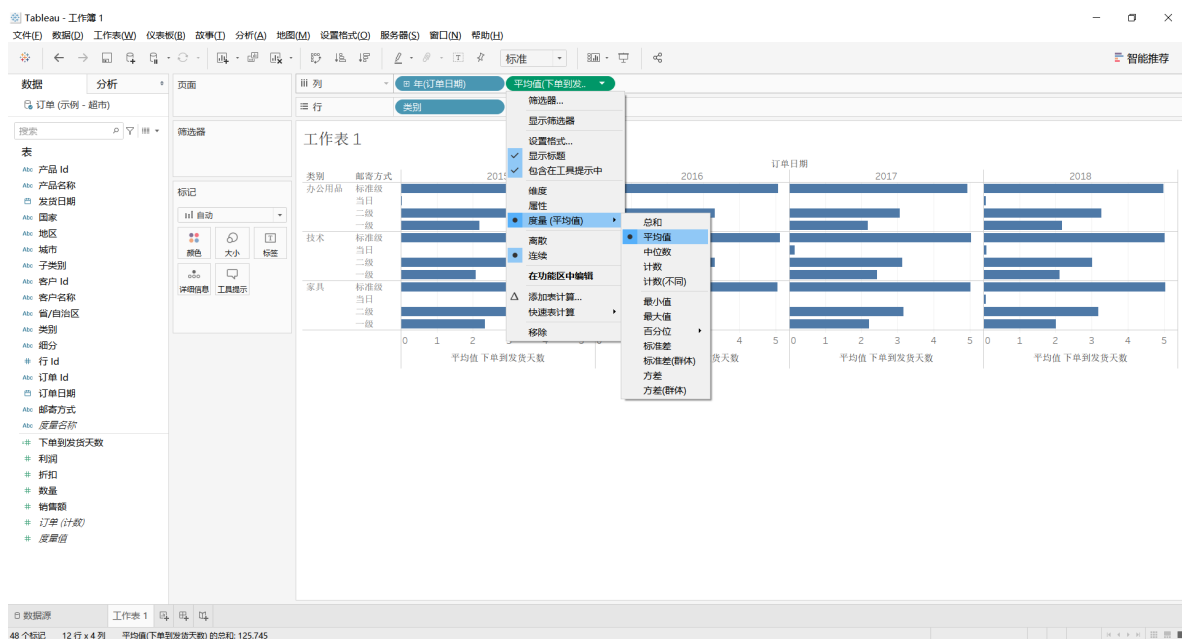
确定

DATEDIFF(date_part, start_date, end_date, [start_of_week])

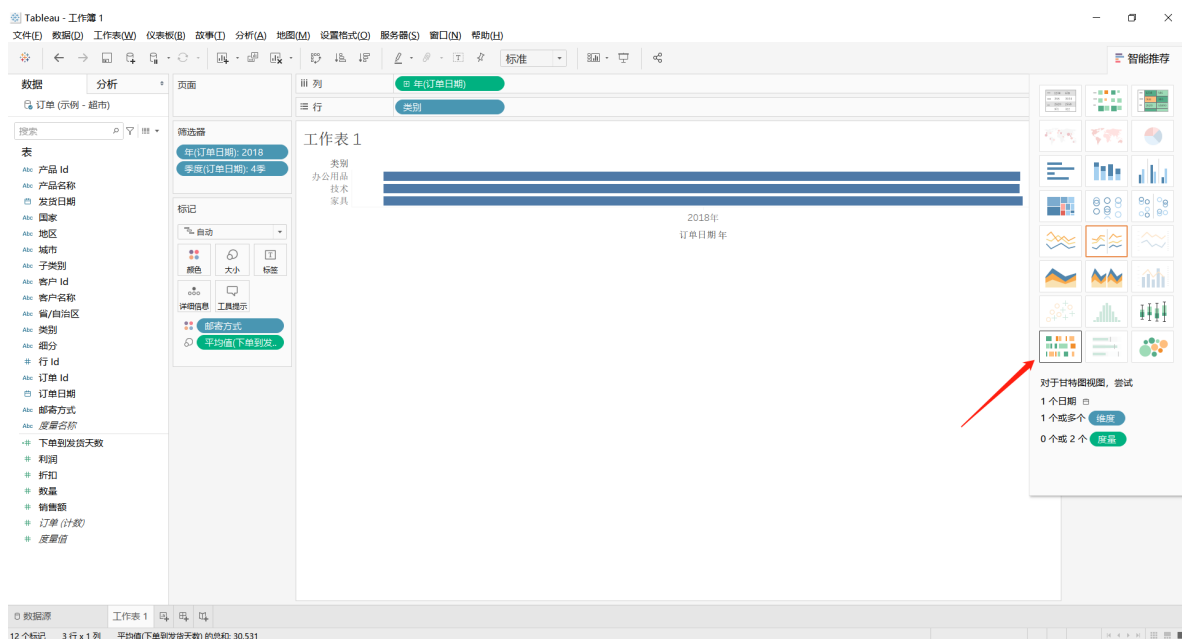
返回两个日期之差，即 end_date 减 start_date。日期差的表示单位为 date_part。如果省略 start_of_week，则周起始日由为数据源配置的起始日确定。

示例：

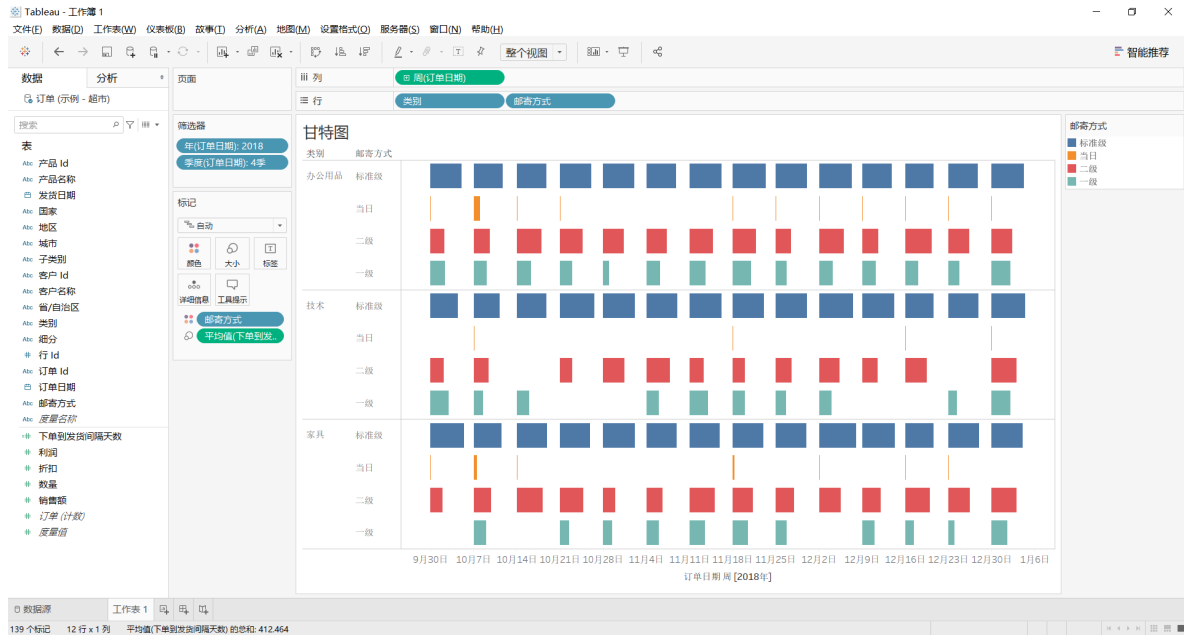
- **Step2:** 将“类别”和“邮寄方式”拖到行功能区，“订单日期”和“下单到发货时间间隔”拖到列功能区，度量方式选择平均值



- **Step3:** 将“订单日期”拖到筛选器中，筛选2018年第四季度，然后在智能推荐区选择甘特图

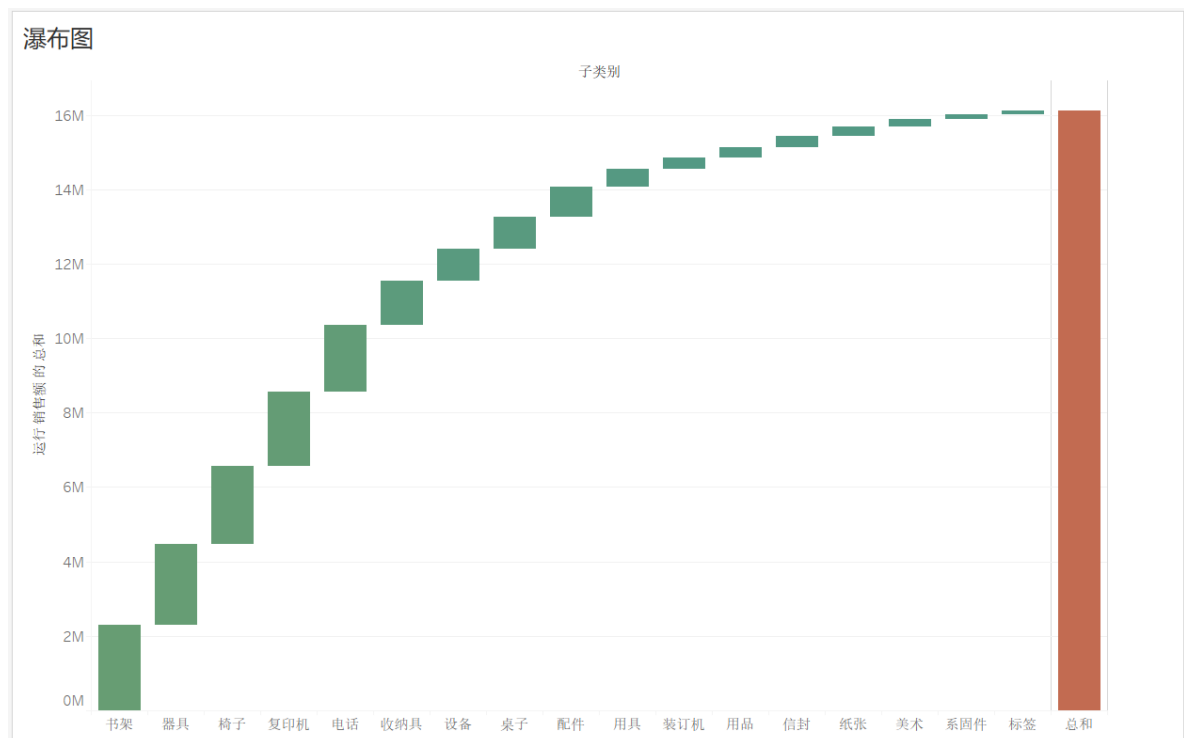


- **Step4:** 把订单日期进行下钻，下钻到周的粒度，然后将“邮寄方式”拖回行功能区



四、瀑布图

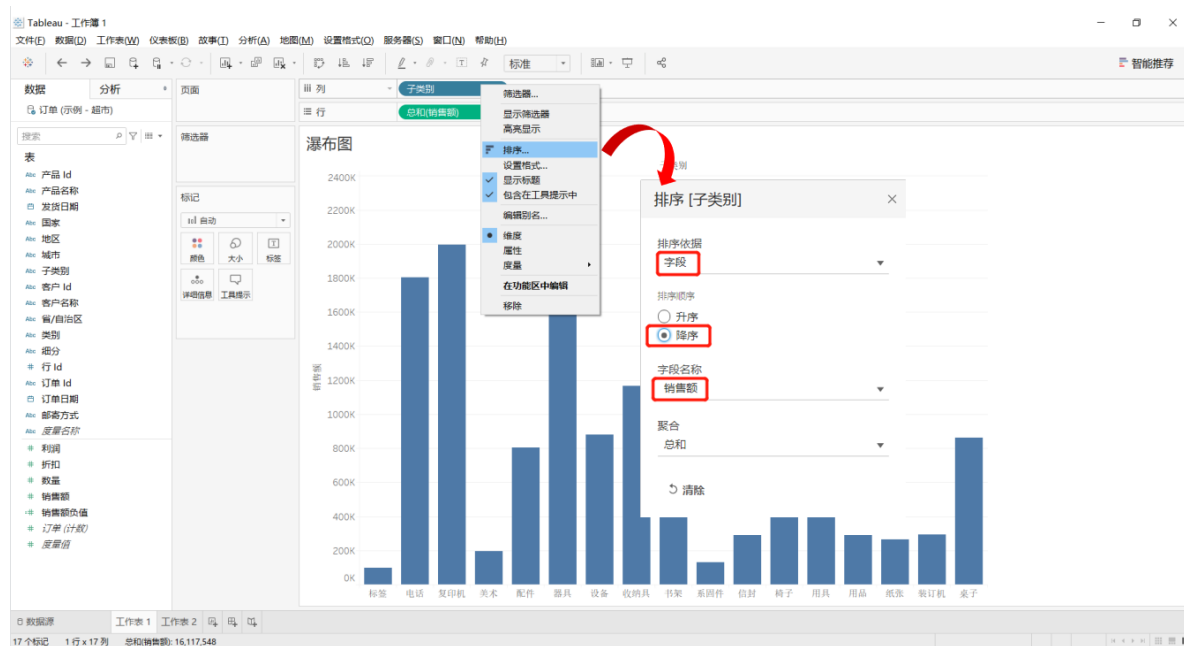
瀑布图是数据可视化分析中常见的一种图形，采用绝对值与相对值结合的方式，适用于表达数个特定数值之间的数量变化关系。对于一系列具有累计性质的正值/负值具有很好的展示功能，既可以辅助理解数据的大小，又能直观地展示出数据的增减变化，反映数据在不同时期或受不同因素的影响结果。



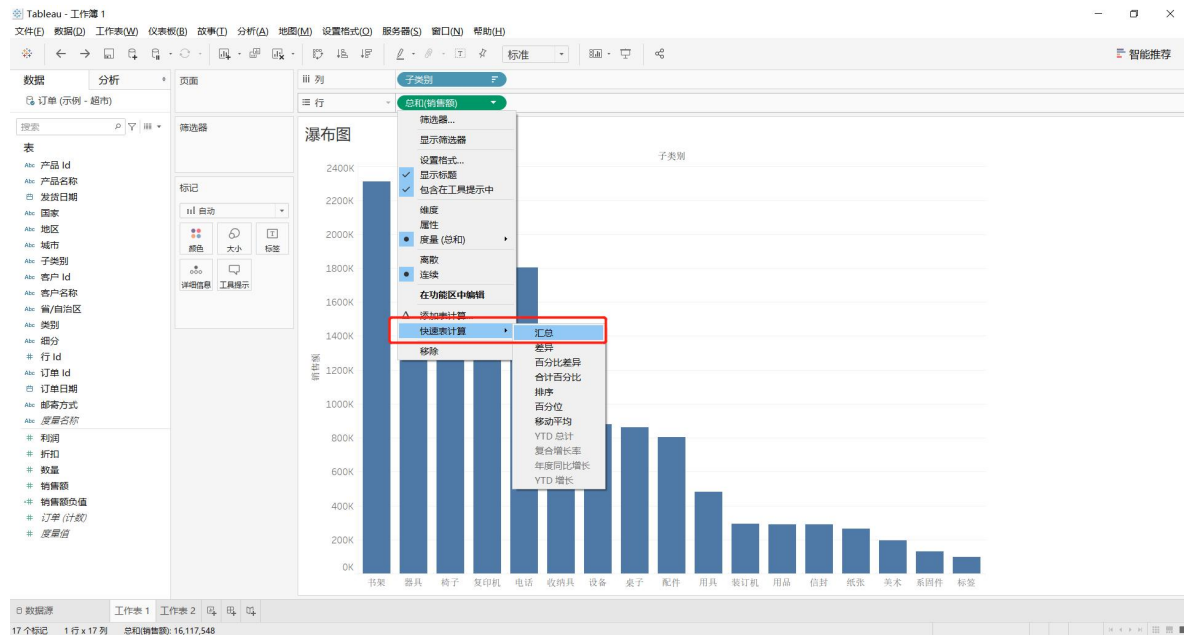
4.1 瀑布图应用场景

在《示例-超市》订单数据中，我们想要分析不同品类销售额对总销售额的贡献以及影响大小，就可以构建不同品类产品销售额的瀑布图，在该图中，从左到右代表各品类对销售额的贡献依次减少，最右边总和代表所有品类销售额的总和

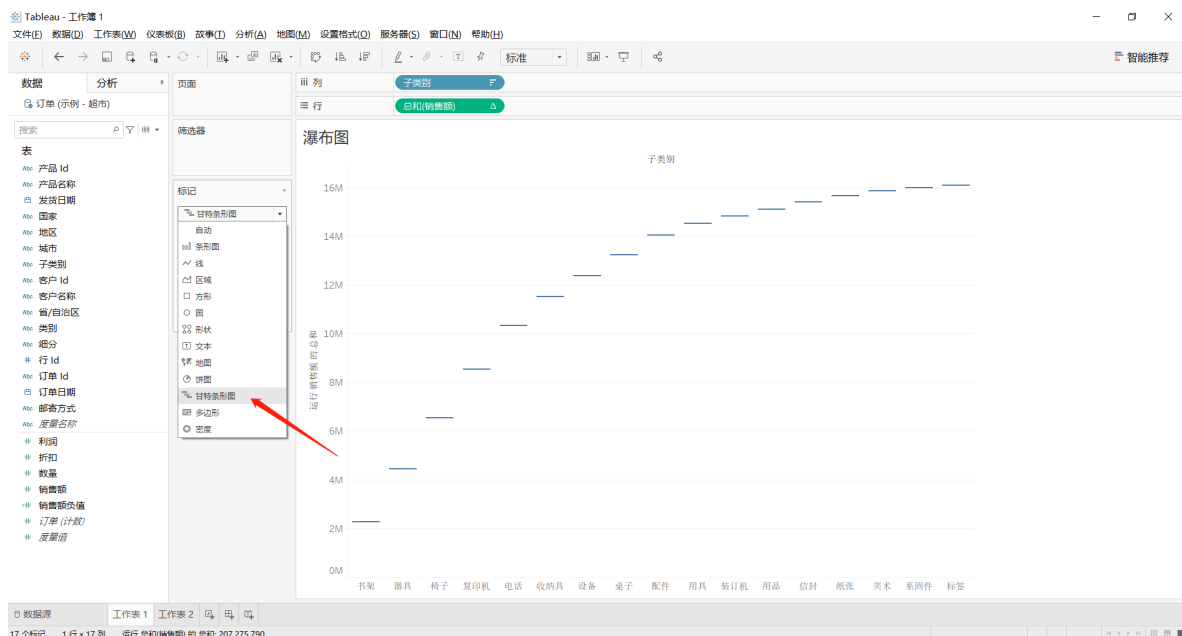
- **Step1:** 将“子类别”拖进列功能区，“销售额”拖进行功能区，对子类别按照销售额进行降序排序



- **Step2:** 在“销售额”下拉菜单快速表计算中选择“汇总”



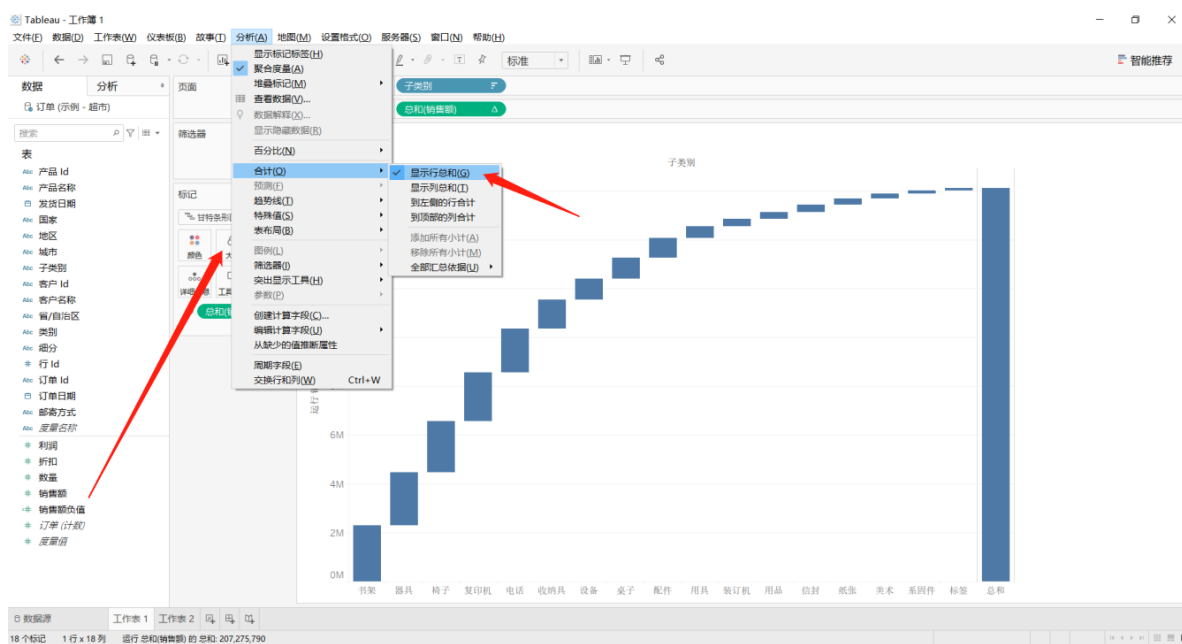
- **Step3:** 在标记卡功能区选择“甘特条形图”



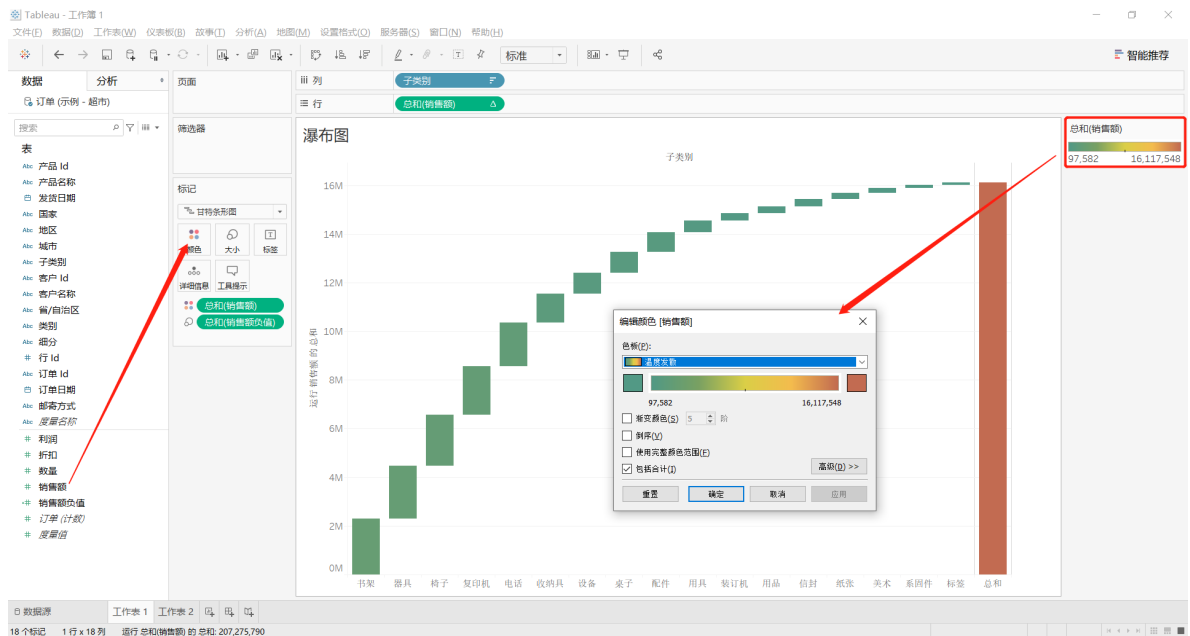
• Step4: 创建计算字段“销售额负值”



• Step5: 将“销售额负值”拖进标记卡“大小”里，在菜单栏“分析 - 合计”中选择“显示行总计”



• Step6: 将“销售额”拖进标记卡“颜色”里，然后对颜色进行编辑

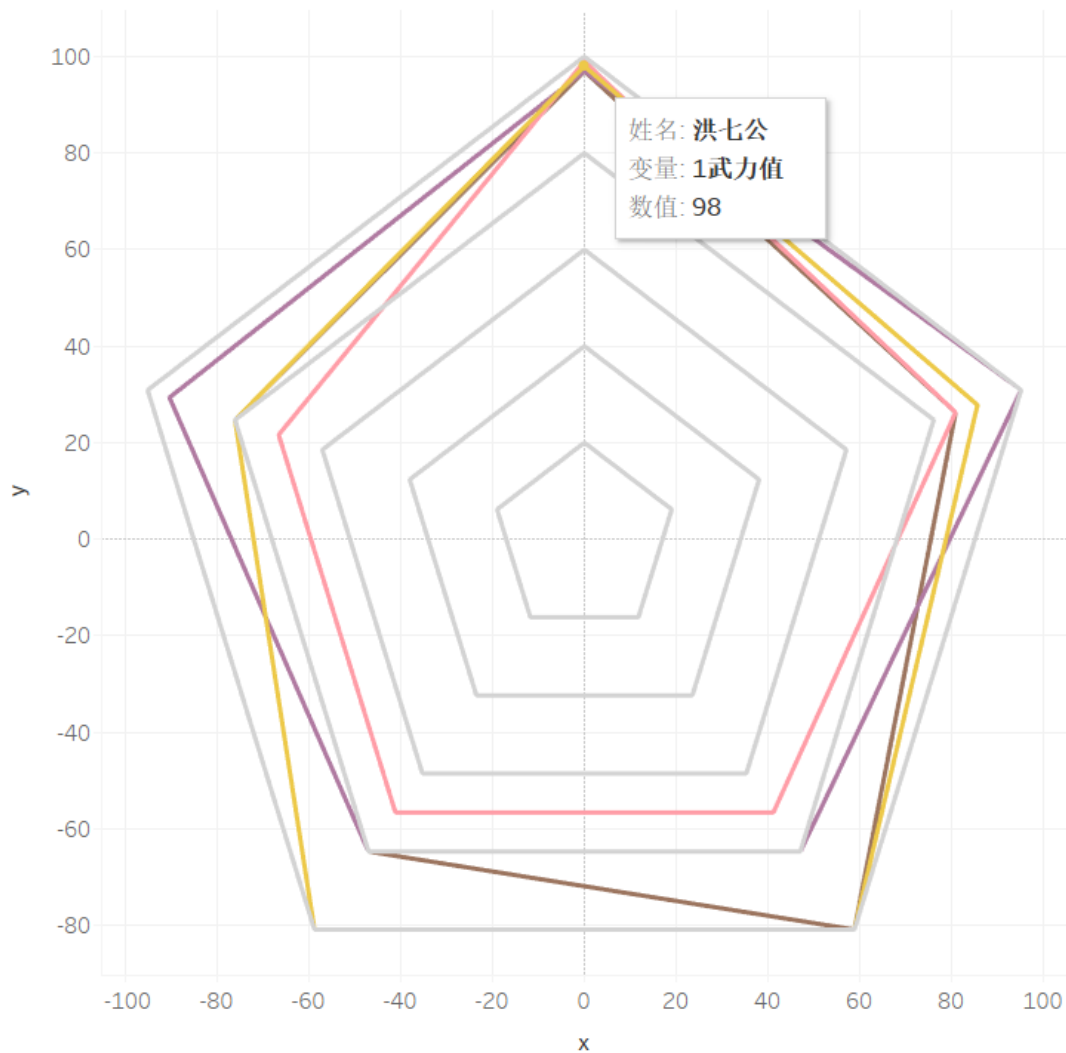


五、雷达图

雷达图是专门用来进行多指标体系比较分析的专业图表，主要应用于企业经营状况的展示——收益性、生产性、流动性、安全性和成长性的评价。其主要特点是简洁、方便、精确、直观，可以将多维数据投影到同一平面上，实现多维数据的可视化。

例如，我们想要对金庸武侠小说中中东邪、西毒、南帝、北丐四个人物进行综合评价，我们就可以选取几个有代表性的指标（武力值、智力值、权力值、魅力值、颜值），通过构建雷达图进行展示。

雷达图



5.1 雷达图应用场景

- **Step1:** 导入数据源，同时选中后六列，点击下拉菜单，选择“转置”，然后将后两列分别命名为“变量”、“数值”

Tableau - 工作簿 1

文件(F) 数据(D) 服务器(S) 窗口(W) 帮助(H)

连接 添加

雷达图数据
Microsoft Excel

工作表

使用数据解释器
数据解释器或许可以清理您的 Microsoft Excel 工作簿。

Sheet1

新建并集

Sheet1 (雷达图数据)

需要更多数据?
将表拖到此处以使其相关。 了解更多信息

转置

姓名	1武力值	1武力值 1	2智力值	3权力值	4魅力值	5颜值
黄药师	97	97	100	80	80	80
欧阳锋	99	99	85	70	70	70
一灯	97	97	85	100	80	80
洪七公	98	98	90	100	100	80
Ring1	20	20	20	20	20	20
Ring2	40	40	40	40	40	40
Ring3	60	60	60	60	60	60
Ring4	80	80	80	80	80	80
Ring5	100	100	100	100	100	100

姓名	变量	数值
黄药师	1武力值 1	97
黄药师	1武力值	97
黄药师	2智力值	100
黄药师	3权力值	80
黄药师	4魅力值	80
黄药师	5颜值	95
欧阳锋	1武力值 1	99
欧阳锋	1武力值	99
欧阳锋	2智力值	85
欧阳锋	3权力值	70
欧阳锋	4魅力值	70

- **Step2:** 转到工作表中，创建计算字段“路径”

路径

CASE [变量]
WHEN '武力值' THEN 1
WHEN '智力值' THEN 2
WHEN '权力值' THEN 3
WHEN '魅力值' THEN 4
WHEN '颜值' THEN 5
ELSE 6
END

计算有效。

4 依赖项

应用

确定

全部

输入搜索文本

ABS
ACOS
AND
ASCII
ASIN
ATAN
ATAN2
ATTR
AVG
BUFFER

ABS (number)
返回给定数字的绝对值。
示例: $ABS(-7) = 7$

- **Step3:** 创建计算字段“弧度”

弧度

Sheet1 (雷达图数据)

IF [路径]=6 THEN $PI() / 2$
ELSE $PI() / 2 - ([路径] - 1) * 2 * PI() / 5$
END

计算有效。

应用

确定

全部

输入搜索文本

ELSEIF
END
ENDSWITH
EXCLUDE
EXP
FIND
FINDNTH
FIRST
FIXED
FLOAT

IF <expr> THEN <then> [ELSEIF <expr2> THEN <then2> ...] [ELSE <else>] END
测试一系列表达式，从而返回第一个 true <表达式> 的 <then> 值。
示例: IF [利润] > 0 THEN 'Profitable' ELSEIF [利润] = 0 THEN 'Break-even' ELSE 'Loss' END

- **Step4:** 创建横坐标轴的值“x”

x

Sheet1 (雷达图数据)

$[数值] * COS([弧度])$

计算有效。

应用

确定

全部

输入搜索文本

AVG
BUFFER
CASE
CEILING
CHAR
COLLECT
CONTAINS
CORR
COS
COT

COS (angle)
返回角度的余弦。以弧度为单位指定角度。
示例: $COS(PI() / 4) = 0.707106781186548$

- **Step5:** 创建横坐标轴的值“y”

y

Sheet1 (雷达图数据)

$[数值] * SIN([弧度])$

计算有效。

应用

确定

全部

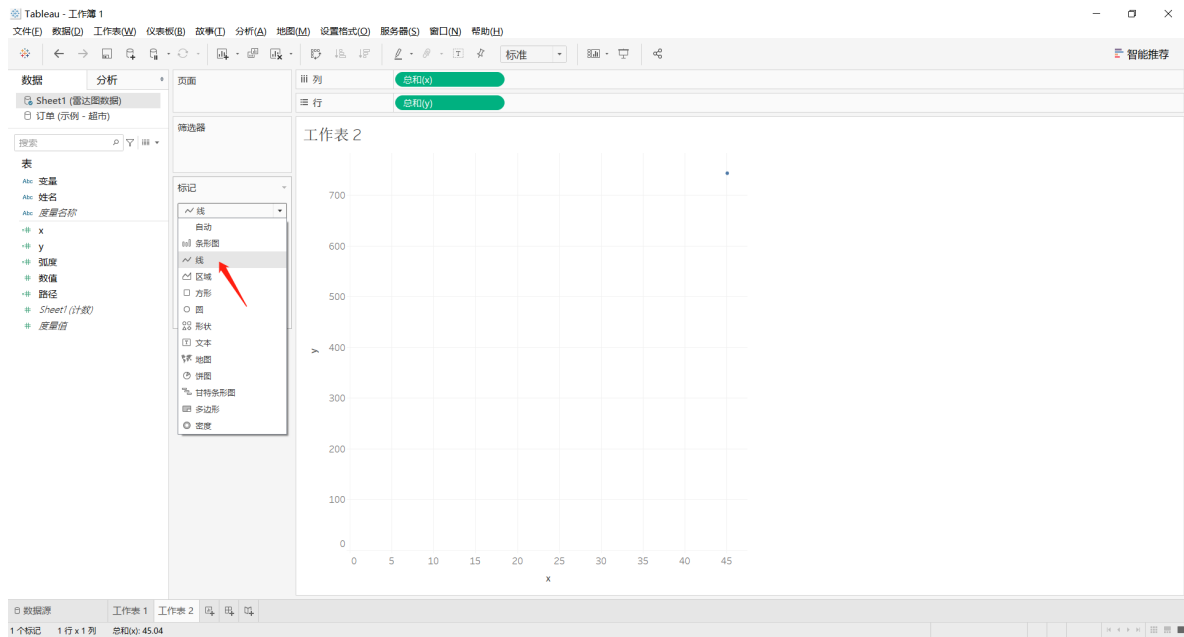
输入搜索文本

RUNNING_MAX
RUNNING_MIN
RUNNING_SUM
SCRIPT_BOOL
SCRIPT_INT
SCRIPT_REAL
SCRIPT_STR
SIGN
SIN
SIZE

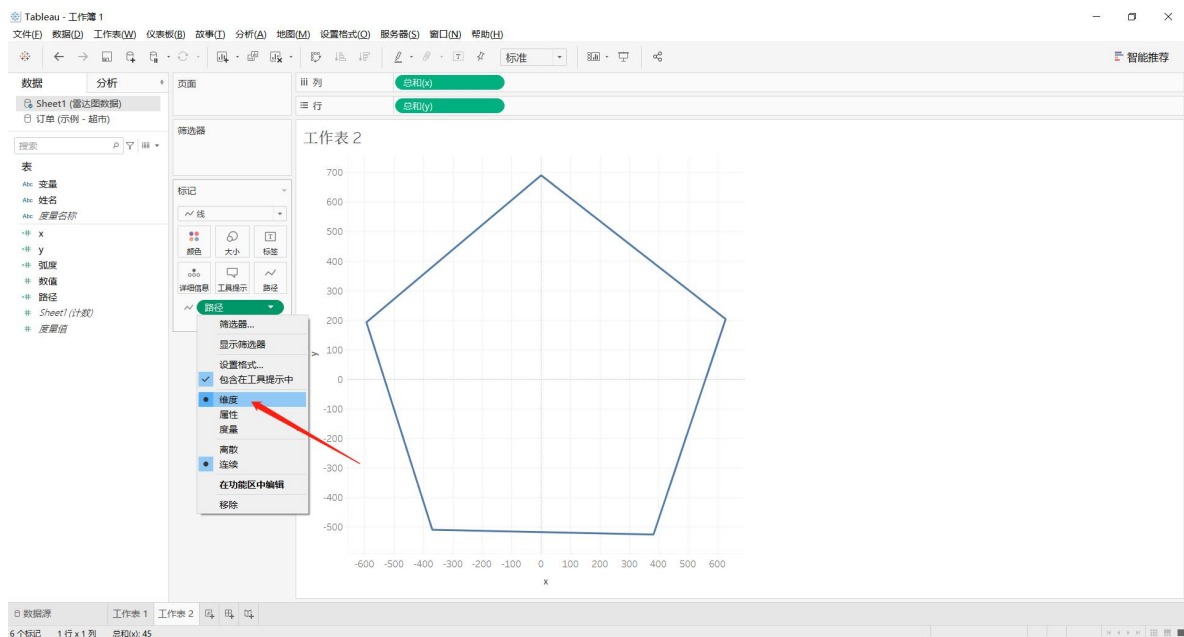
弧度
数据类型: 浮点
状态: 有效
IF [路径]=6 THEN $PI() / 2$
ELSE $PI() / 2 - ([路径] - 1) * 2 * PI() / 5$
END

描述(D)...

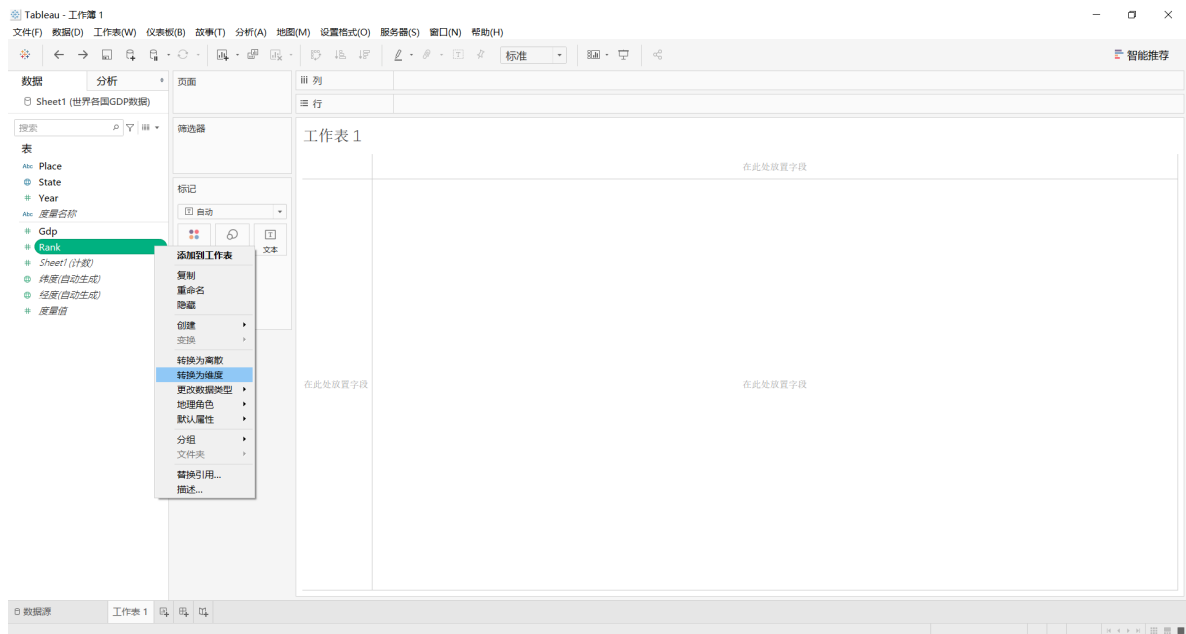
- **Step6:** 把“x”拖进列，“y”拖进行，在标记卡中选择“线”



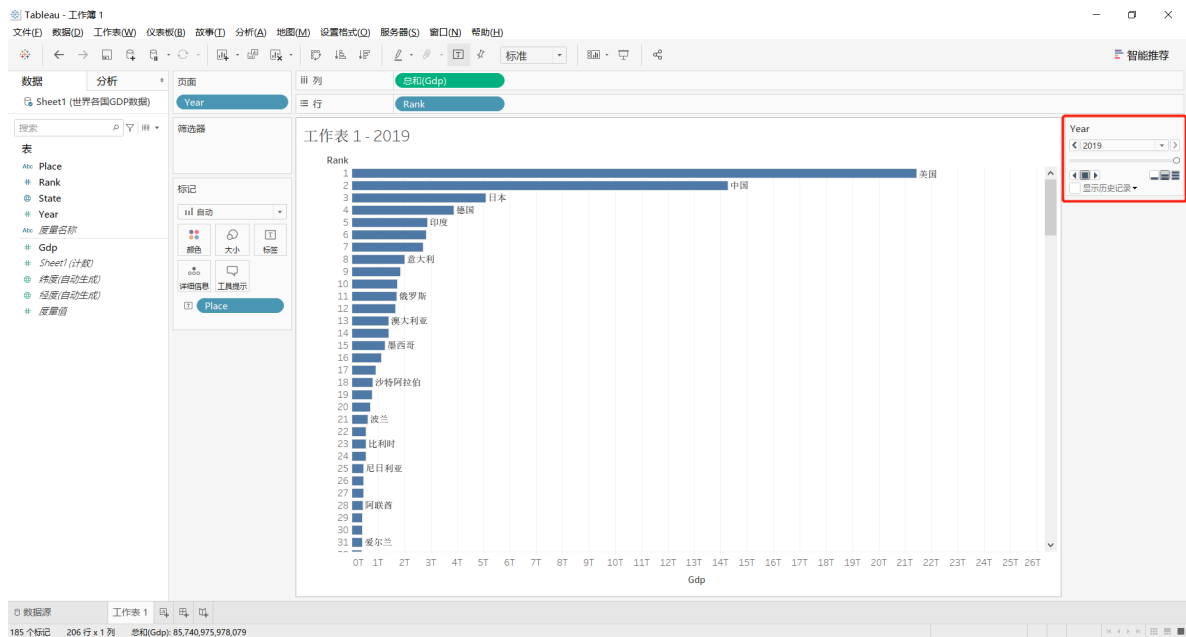
- **Step7:** 把“路径”拖进标记卡中的路径，然后点击下拉菜单，选择“维度”



- **Step8:** 把“姓名”拖进标记卡中的颜色，然后点击右上角编辑颜色，把Ring1 - Ring5置灰，点击确定



- **Step2:** 将“Rank”拖进功能区，“GDP”拖进列功能区，“Year”拖进“页面卡”，“Place”拖进“标记卡 - 标签”



- **Step3:** 创建一个维度计算字段“中国颜色”

中国颜色

```
IF [Place]='中国' THEN 1
ELSE 0
END
```

计算有效。

全部

输入搜索文本

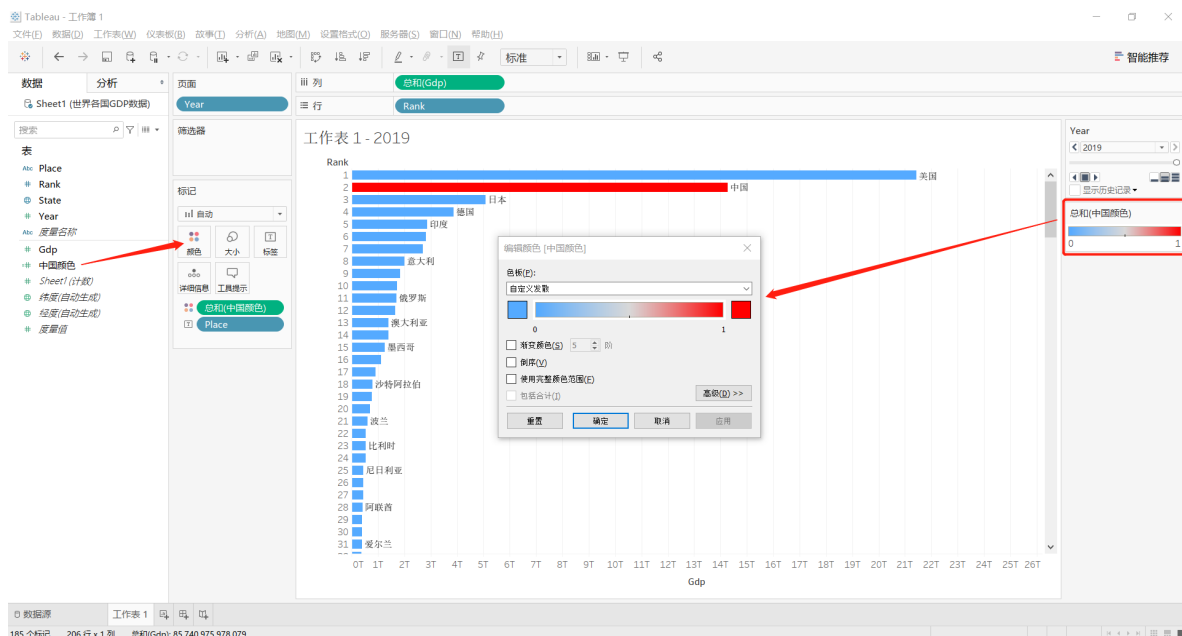
- ELSE
- ELSEIF
- END
- ENDSWITH
- EXCLUDE
- EXP
- FIND
- FINDNTH
- FIXED

IF <expr> THEN <then> [ELSEIF <expr2> THEN <then2> ...] [ELSE <else>] END

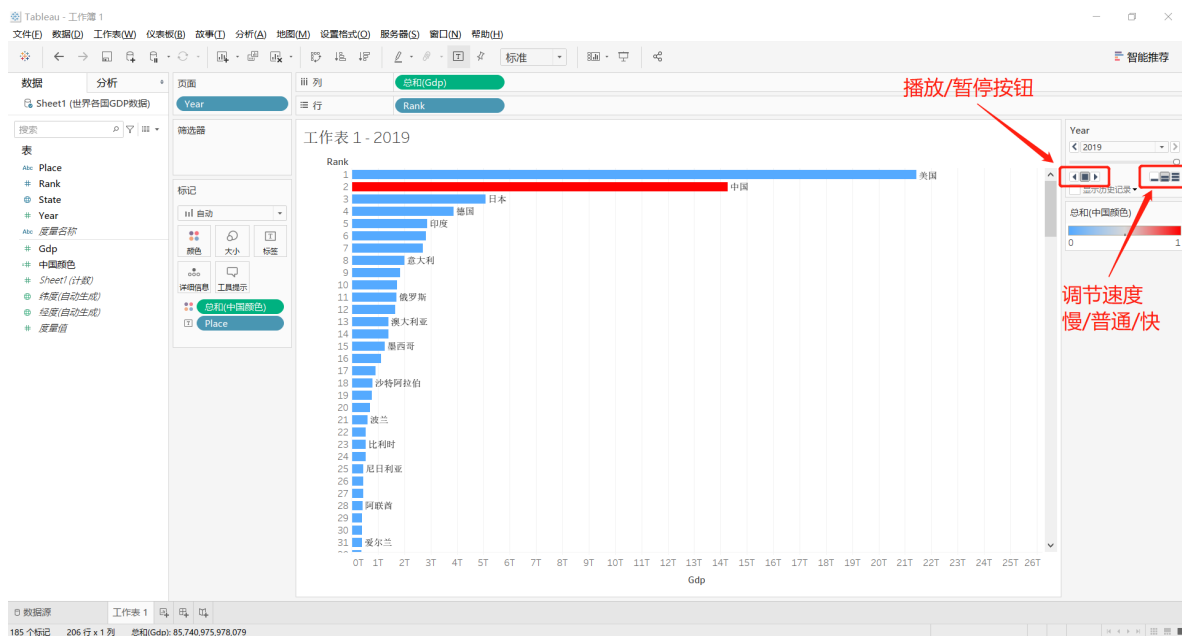
测试一系列表达式，从而返回第一个 true <表达式> 的 <then> 值。

示例: IF [利润] > 0 THEN 'Profitable' ELSEIF [利润] = 0 THEN 'Break-even' ELSE 'Loss' END

- **Step4:** 将“中国颜色”拖进“标记卡 - 颜色”，并对颜色进行编辑



- **Step5:** 在右侧“页面卡”筛选播放速度，点击进行自动播放



本章小结

第五章 高阶可视化，满足你一切想象力

帕累托图

- 概念 —— 按照一定的类别根据数据计算出其分类所占的比例，用于分析导致结果的主要因素
- 应用 —— “二八原则”的图形化体现，能够直观体现两点重要信息：“至关重要的极少数”和“微不足道的大多数”。经常被应用在销售管理中

盒须图

- 概念 —— 又叫箱线图，用以显示数据的位置、分散程度、异常值等
- 应用 —— 观测数据在同类群体中的位置，可以知道哪些表现好，哪些表现差；比较四分位全距及线段的长短，可以看出哪些群体分散，哪些群体更集中

甘特图

- 概念 —— 以图示的方式通过活动列表和时间刻度形象地表示出任何特定项目的活动顺序和持续时间
- 应用 —— 主要用来比较与日期相关的不同活动（项目）的持续时间长短。也可用于显示不同任务之间的依赖关系，普遍应用在项目管理中

瀑布图

- 概念 —— 采用绝对值与相对值结合的方式，适用于表达数个特定数值之间的数量变化关系
- 应用 —— 直观展示数据的增减变化，反映数据在不同时期或受不同因素的影响结果

雷达图

- 概念 —— 专门用来进行多指标体系比较分析的专业图表
- 应用 —— 可以将多维数据或指标投影到同一平面上，实现多维数据的可视化

动态图

- 概念 —— 根据不同的选项动态变化的图表
- 应用 —— 能够从不同维度动态交互查看复杂数据的信息