



项目二 可视化数据的准备

【能力目标】

1. 能够理解不同数据源类型特点，会进行各类数据源连接的配置。
2. 能够根据数据分析需求进行数据筛选处理，创建数据分析所需的新字段。
3. 能够理解多表关联的意义，可以对不同来源的多表数据进行不同关联的组合。

进行可视化任务首先要进行数据的准备，根据分析任务的要求，数据的来源可能不同，有可能来自外部网站或者企业内部的业务系统，数据准备的任务就是将不同的数据源整合到一起。

在信息化管理的企业中，绝大多数的数据都是存储在关系型数据库中的，所以需要可视化分析软件连接数据源，获取分析所需的数据；Excel 类型的文档同样在企业中应用广泛，也是数据来源的主要途径。在获得相应的数据后，还需要通过数据可视化分析软件对数据进行整理、合并、创建新字段等操作，形成分析所需要的数据集。

FineBI 提供了一个统一的数据连接接口，通过选择不同的参数，就可以完成各类数据库的连接，从而获取数据库内的各个数据表。



任务 2.1 连接数据源

【任务描述】

要开展分析任务，首先要确定数据源的类型，然后根据不同类型的分别进行预处理，将不同来源的数据导入 FineBI 分析软件的数据库，而 Excel 文件和关系型数据库需要采用不同的方式。

【知识准备】

1. Excel 表格

在日常工作中，经常使用 Excel 工作簿来管理各种数据，一个工作簿可以包括多个工作表。在 FineBI 中可以选择 Excel 数据表作为分析的数据源。在 FineBI 中选择 Excel 数

导入 Excel
来确定其类型。

2. 结构化

(1) MySQL
分社区版（免
之一，在 Web
关系数据库管
MySQL

就加快了数
最常用标准
一般中小型

(2) SQL
具有使用万

SQL Serve
Micro

了企业级
安全可靠

(3)

数据库关
据管理工

实现了分

Ora

承租方

提高资

Data O

虽

可用性

(

应用

有常

可恢

命令

使得

地化

多任

表,通过上传导入 FineBI 的工作库中,数据只需要导入一次就可以在后续环节中多次使用。
导入 Excel 数据时要求每个工作表有一个标题行, FineBI 会自动根据表格每列的数据来确定其类型。

2. 结构化数据

(1) MySQL 数据库。MySQL 是一个关系型数据库管理系统,属于 Oracle 旗下产品,分社区版(免费)和企业版(收费)两种类型。MySQL 是最流行的关系型数据库管理系统之一,在 Web 应用方面,MySQL 是最好的 RDBMS (Relational Database Management System, 关系数据库管理系统) 应用软件之一。

MySQL 数据库将数据保存在不同的表中,而不是将所有数据放在一个大仓库内,这样就加快了数据存取速度并提高了灵活性。MySQL 所使用的 SQL 语言是用于访问数据库的最常用标准化语言。由于其体积小、速度快、总体拥有成本低,尤其是开放源码这一特点,一般中小型网站和企业项目都选择 MySQL 作为数据库。

(2) SQL Server 数据库。SQL Server 是 Microsoft 公司推出的关系型数据库管理系统,具有使用方便、可伸缩性好、与相关软件集成程度高等优点。依靠微软的 Windows 平台,SQL Server 在众多的企业系统中被广泛用于数据存储。

Microsoft SQL Server 是一个全面的数据库平台,使用集成的商业智能 (BI) 工具提供了企业级的数据管理。Microsoft SQL Server 数据库引擎为关系型数据和结构化数据提供了更安全可靠的存储功能,使用户可以构建和管理用于业务的高可用和高性能的数据应用程序。

(3) Oracle 数据库。Oracle 数据库系统是美国 Oracle 公司(甲骨文)提供的以分布式数据库为核心的一组软件产品。作为一个通用的数据库系统,Oracle 数据库具有完整的数据管理功能;作为一个关系型数据库,它是一个完备关系的产品;作为分布式数据库,它实现了分布式处理功能。

Oracle 数据库最新版本为 Oracle Database 19c。Oracle Database 19c 引入了一个新的多承租方架构,使用该架构可轻松部署和管理数据库云。此外,一些创新特性可最大限度地提高资源的使用率和灵活性,如 Oracle Multitenant 可快速整合多个数据库,而 Automatic Data Optimization 和 Heat Map 能以更高的密度压缩数据和对数据进行分层。

虽然相对于前面两款数据库系统,Oracle 数据库的商业售价较高,但是 Oracle 数据库可用性强、可扩展性强、数据安全性强、稳定性强,在国内大中型企业中应用广泛。

(4) DB2。IBM DB2 是美国 IBM 公司开发的一套关系型数据库管理系统。DB2 主要应用于大型应用系统,具有较好的可伸缩性,可支持从大型机到单用户环境,应用于所有常见的服务器操作系统平台下。DB2 提供了高层次的数据利用性、完整性、安全性、可恢复性,以及小规模到大规模应用程序的执行能力,具有与平台无关的基本功能和 SQL 命令。

DB2 采用了数据分级技术,能够使大型机数据很方便地下载到 LAN 数据库服务器,使得客户机/服务器用户和基于 LAN 的应用程序可以访问大型机中的数据,并使数据库本地化及远程连接透明化。

DB2 以拥有一个非常完备的查询优化器而著称,其外部连接改善了查询性能,并支持多任务并行查询。DB2 具有很好的网络支持能力,每个子系统可以连接十几万个分布式用

户,可同时激活上千个活动线程,对大型分布式应用系统尤为适用。
 不过由于 DB2 早期面向 IBM 大型机,在国内的知名度不及 Oracle 和 SQL Server,及硬件、售价等因素,在国内企业中使用并不多,大多集中在金融领域内。

3. FineBI 组合数据的方式

在连接数据源之前,需要了解一下 FineBI 中数据表的存放方式。分组和业务包是 FineBI 的数据管理方式,如图 2.1 所示,可以理解为文件夹,分组和业务包的关系相当于上层文件夹和下层文件夹,业务包中存放着用户定义数据连接从数据库中获取的数据或者上传的 Excel 数据,也就是用户需要使用进行分析的数据表。一个分组可以包含多个业

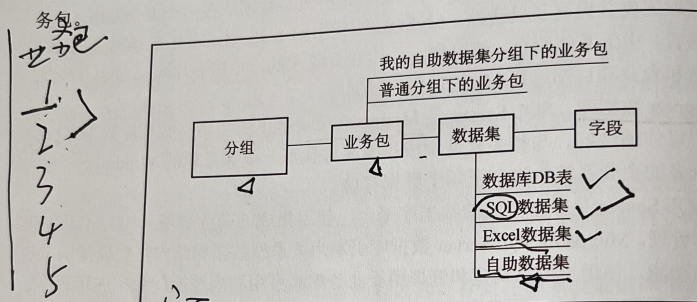


图 2.1 FineBI 数据表的组织结构

业务包通过 FineBI 定义的数据连接从数据库中获取数据。业务包中包含了连接数据库所获取的数据表。若为非实时数据表,业务包在数据更新以后将获取到的数据保存在本地,BI 分析则从本地读取数据。实时数据的数据表中则保存了获取连接数据库数据的一系列 SQL 配置等,在模板分析时生成相应的 SQL 语句向数据库查询。

【任务实施】

任务 2.1.1 连接 Excel 数据源

Excel 采用了工作簿形式存储数据,在使用时可以直接在“数据准备”菜单中将其导入到分析软件中,不需要额外建立数据链接。

(1) 在 FineBI 中选择“数据准备”菜单,右侧出现数据列表,显示当前用户可以使用的分组和业务包,如图 2.2 所示。

(2) 接下来以创建 GDP 分组为例,说明 Excel 数据源的使用。在图 2.2 所示界面中单击“添加分组”按钮,输入分组名称为“GDP”,然后单击该分组名称,再单击“添加业务包”按钮,输入业务包名称为“全国 GDP”(业务包名称不可重复)。光标位于分组或业务包时,右侧显示“...”,可以对它进行重命名或删除操作,如图 2.3 所示。

(3) 单击业务包名称,即可进入数据列表界面,如图 2.4 所示。单击“添加表”按钮,在弹出下拉菜单中选择“Excel 数据集”命令。



连接 Excel 数据源视频讲解

图 2.2

如选
单击“上
据的字段
据,其余
定”按钮

据有
开对

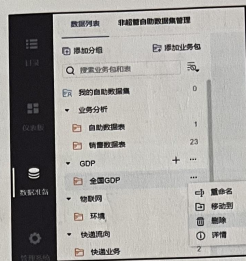


图 2.3 分组与业务包的管理

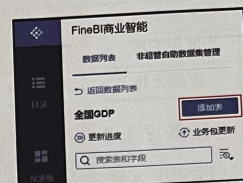


图 2.4 数据表界面

定”按钮即可完成数据表的创建。

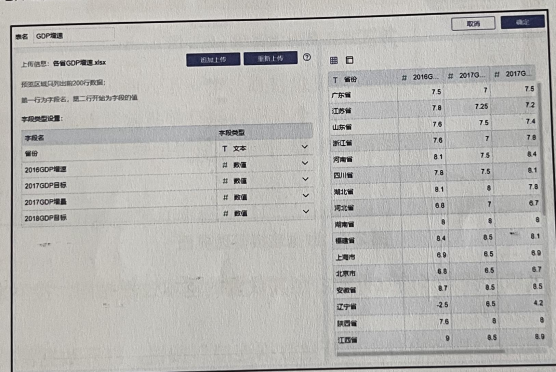


图 2.5 上传 Excel 文档并预览

开对话框，对 Excel 数据集进行追加或更新。

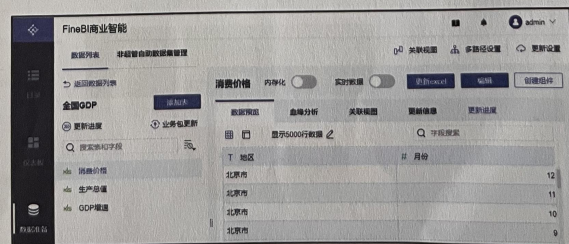


图 2.6 上传多个 Excel 数据集

FineBI 支持追加上传与重新上传的操作，如表 2.1 所示。

表 2.1 上传类型

上传类型	解释
追加上传	在原先 Excel 数据表的基础上追加 Excel 数据
重新上传	此次上传 Excel 数据会替换原先的数据

任务 2.1.2 连接 MySQL 数据源

FineBI 支持从 MySQL 这类关系型数据库中获取数据进行分析，并且为这些数据库系统提供了一个统一的管理和设置界面。在进行“业务包”添加“数据集”之前，需要通过“管理系统”菜单中的“数据连接”下的“数据连接管理”进行数据源设置，方便数据集获取。数据连接的管理界面如图 2.7 所示。

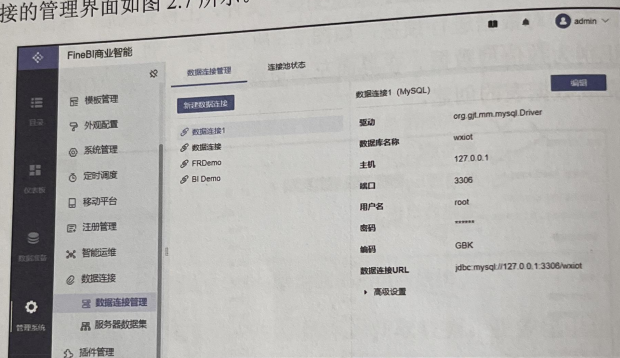
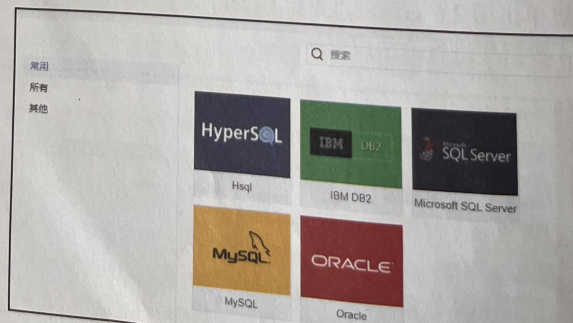


图 2.7 数据连接管理界面

在 FineBI 中可以选择众多的数据源，但其设置的基本过程相同，接下来以 MySQL 数据源为例，说明连接过程。

(1) 在图 2.7 所示的界面中单击“新建数据连接”按钮，打开数据源类型选择界面，如图 2.8 所示。常用选项卡列表中包括了 HyperSQL、DB2、SQL Server、MySQL、Oracle 五种类型，选择左侧的“所有”选项，可以查看系统支持的所有数据库类型，如图 2.9 所示，用户通过单击图标可以选择相应的数据库类型。



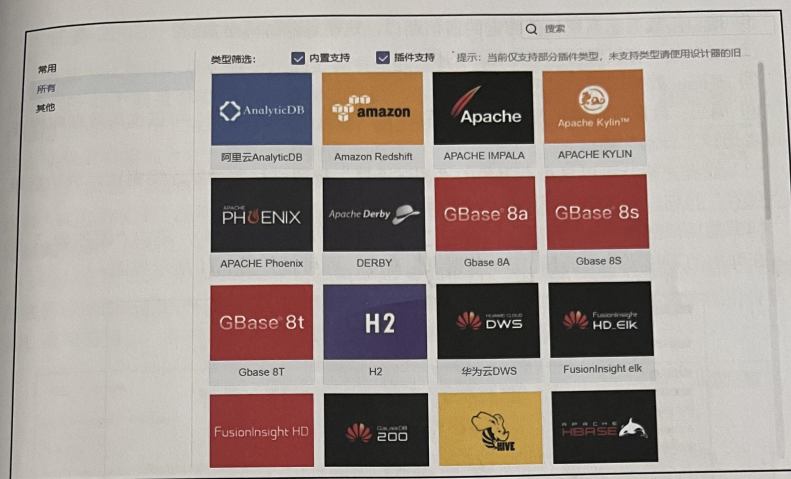


图 2.9 所有数据库类型列表

(2) 选择“MySQL”，进入下一步，进行数据库连接的相关参数设置界面，如图 2.10 所示。一般设置数据库连接的参数包括以下几个。

- ① 数据库名称：一个数据库管理系统可以包括多个数据库，因此连接时需要知道具体的数据库名字。
- ② 主机：如果分析数据来自本机，可以填 localhost；如果是网络中的其他机器，则填该机器 IP 地址。

图 2.10 展示了 MySQL 数据库连接参数设置界面。窗口顶部有“取消”、“测试连接”和“保存”按钮。下方是“MySQL”配置区域，包含以下字段：

- 数据连接名称：数据连接2
- 驱动：com.mysql.jdbc.Driver
- 数据库名称：database
- 主机：hostname
- 端口：端口
- 用户名：用户名
- 密码：密码
- 编码：自动
- 数据连接URL：jdbc:mysql://hostname:port/database

底部有一个“高级设置”链接。

图 2.10 数据库连接参数设置

因此,用于表示数据构成关系的图表,可以有饼图、环形图、玫瑰图,以及堆积柱形图、堆积面积图等。饼图、环形图、玫瑰图适合表示数据静态构成的情况,当数据构成随时间动态变化时,可以用堆积柱形图或堆积面积图来表示。



任务 3.3 数据联系和分布可视化

【任务描述】

针对 2018 年 1~5 月中国空调销售的数据,现在需要对下面的情况进行分析:

- (1) 产品价格与销量间的关系。
- (2) 各类产品在各价格段的销售量分布情况。
- (3) 销量前 10 的品牌及其利润关系。

【知识准备】

1. 数据的联系和分布

数据的联系主要是分析数据中各个变量之间的关系,比如散点图主要是分析因变量随自变量而变化的大致趋势,由此趋势可以选择合适的函数进行经验分布的拟合,如线性关系、指数关系、对数关系等。数据的分布主要是分析数据分布的规律,比如正态式分布、线性分布。数据的分布和联系可视化图表如图 3.37 所示。

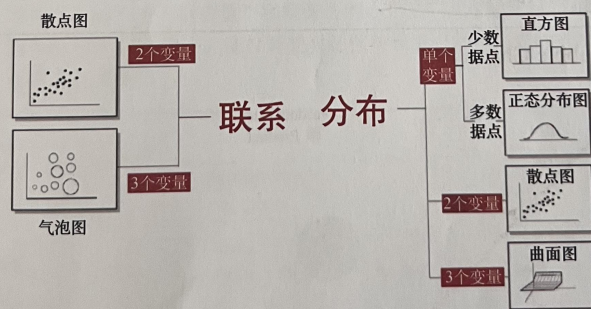


图 3.37 数据的分布和联系可视化图表

2. 表示数据联系和分布的可视化图表

表示数据联系的可视化图表有气泡图和散点图。其中散点图表示数据中两个变量之间的联系,如图 3.38 所示,而气泡图可以表示三个变量之间的联系,如图 3.39 所示。散点图是指在数理统计回归分析中,数据点在直角坐标系平面上的分布图,散点图核心的价值在于发现变量之间的关系。

同时，散点图可以用于表示数据的分布。另外直方图也非常适合表示数据的分布情况。

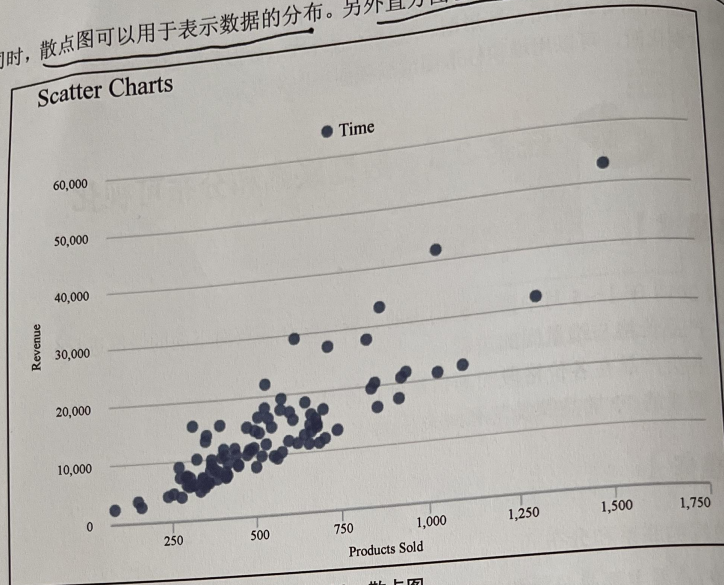


图 3.38 散点图

由于散点图一般研究的是两个变量之间的关系。因此，气泡图就是在散点图的基础上增加变量，提供更加丰富的信息，点的大小或者颜色可以定义为第三个变量。所示气泡图可以看作是散点图的变形。气泡图通常用于展示和比较数据之间的关系与分布，一般用颜色映射类型，用圆圈大小映射数值。

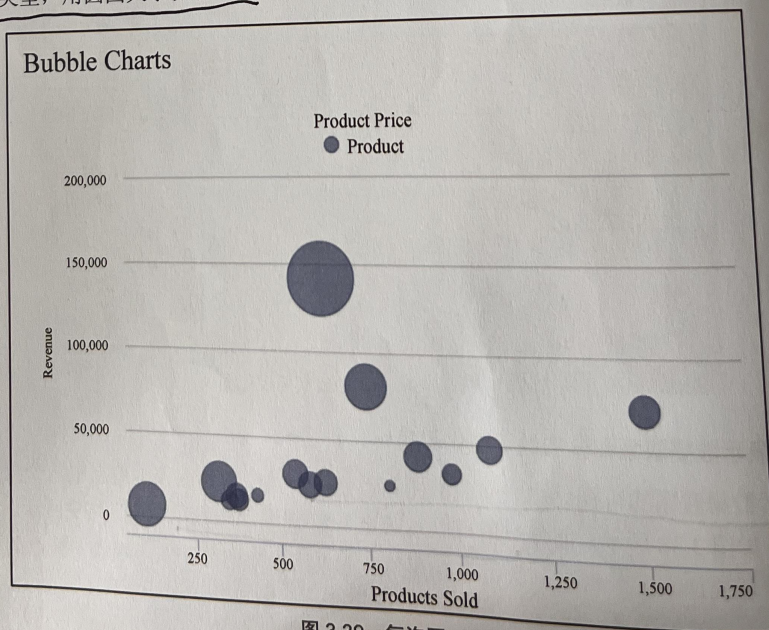


图 3.39 气泡图

【任务实施】

任务 3.3

在本任务中，我们可以用散点图来展示数据。

- (1) 将数据导入到 Excel 窗口中的“数据”选项卡。
- (2) 然后，选择“数据”选项卡中的“数据”选项。

图表

变化的
分析：
“指数

【任务实施】

任务 3.3.1 产品价格与销售量的关系

在本任务中,需要分析产品的价格与销售量这两个变量之间的关系,可以用散点图来表示。

(1) 将“维度”窗口中的“价格段”字段拖放到“横轴”,将“指标”窗口中的“销售量(求和)”字段拖放到“纵轴”,此时生成默认的柱形图。

(2) 然后在“图表类型”中选择“散点图”,则生成价格与变量关系的散点图,如图 3.40 所示。



产品价格与销售量的关系
视频讲解

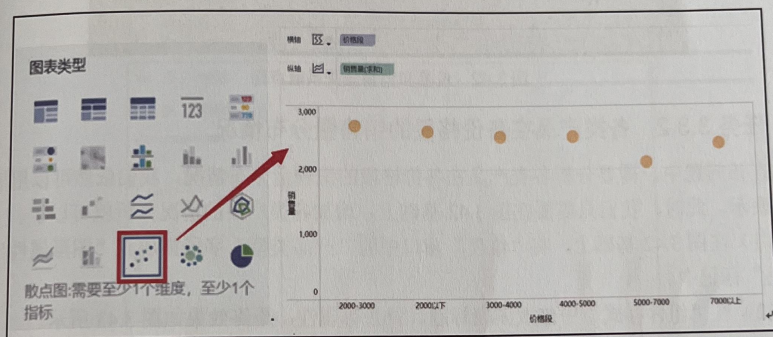


图 3.40 基本散点图

(3) 为了分析价格与销售量的关系,我们为图表添加一条分析线,分析销售量随价格变化的趋势。单击“纵轴”的“销售量”字段旁边的小三角,在弹出的菜单中选择“设置分析线”→“趋势线(纵向)”命令,在打开的对话框中设置“趋势线”的“拟合方式”为“指数拟合”,如图 3.41 所示。

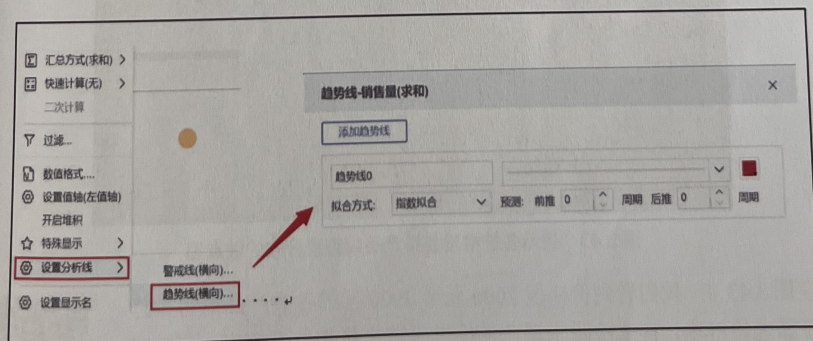


图 3.41 添加趋势线

(4) 最后设置标题、背景、轴线等组件样式,最终效果如图 3.42 所示。

从图 3.42 中可以看到,随着价格的上涨,产品销售量在逐步降低,其中 4000 元至 5000 元价格段和 7000 元以上的价格,销售量高于拟合线。而 5000 元至 7000 元价格段销售量低于拟合线。

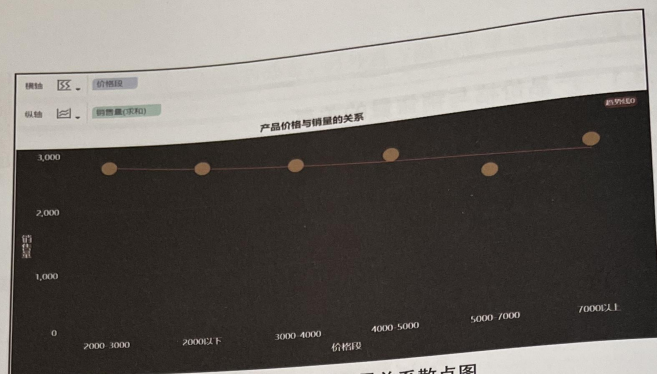


图 3.42 价格和销售量关系散点图

任务 3.3.2 各类产品在各价格段的销售量分布情况

在该问题中，需要分析各类产品在各价格段的销售量分布情况，我们依然可以用散点图来表示。此时，我们只需要在图 3.42 基础上，增加各类产品的情况分析就可以了。

(1) 在图 3.42 基础上，将“维度”窗口中的“产品类型”字段拖放到“图形属性”→“颜色”标记中。

(2) 在“组件样式”中修改图表标题、背景等属性，最终效果如图 3.43 所示。

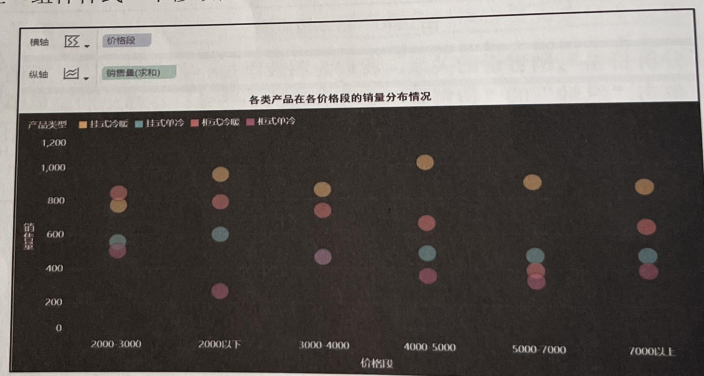


图 3.43 表示各价格段各类产品销售量分布的散点图

在图 3.43 中，我们看到价格段 2000 元至 3000 元的各类产品销售量总体高于其他价格段。挂式冷暖空调在各价格段销售量都位于前列，而柜式单冷空调在各价格段的销售量表现都不佳。

任务 3.3.3 销售量前 10 的品牌分布及其利润关系

在该任务中，需要分析销售量前 10 的品牌分布及利润关系，涉及到销售量、品牌及利润三个变量，因此我们选择气泡图来表示这三者数据间



销售量前 10 的品牌分布及其利润关系视频讲解

的联系。

(1) 将“维度”窗口中的“品牌”字段拖放到“横轴”，将“指标”窗口中的“销售量”字段拖放到“纵轴”。单击“纵轴”中的“销售量(求和)”字段旁边的小三角，在弹出的菜单中选择“过滤”命令，为“销售量(求和)”添加过滤条件，如图 3.44 所示，从而生成销售量前 10 的品牌柱形图，如图 3.45 所示。

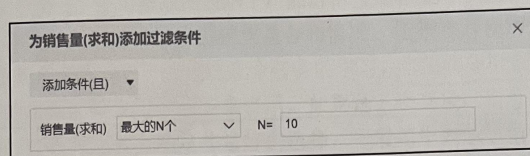


图 3.44 为销售量添加过滤条件

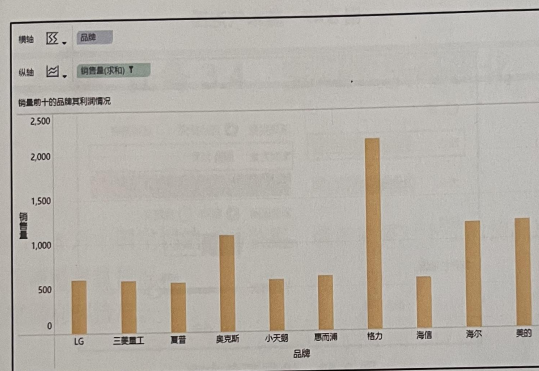


图 3.45 销售量前 10 的品牌

(2) 在“图表类型”中选择“气泡图”，如图 3.46 所示，然后呈现表示销售量前 10 的气泡图，如图 3.47 所示。



图 3.46 选择气泡图

(3) 用气泡的颜色深浅来映射利润的大小。将“指标”窗口中的“利润”字段拖放到“图形属性”→“颜色”标记，并设置颜色“渐变方案”为“炫彩”，如图 3.48 所示。

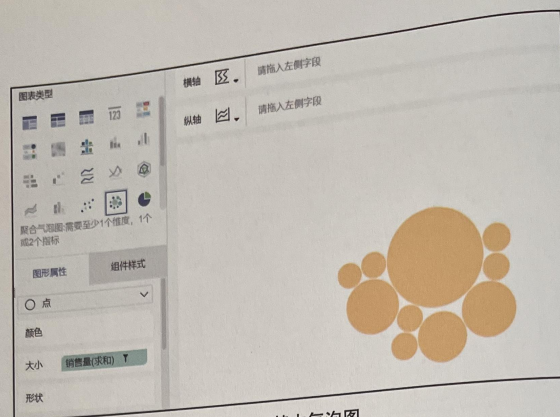


图 3.47 基本气泡图

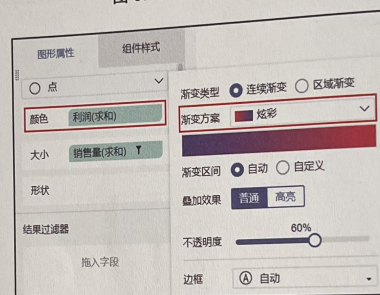


图 3.48 设置颜色方案

(4) 为了增加图表可读性，将“维度”窗口中的“品牌”字段和“指标”窗口中的“利润”字段拖放到“图形属性”→“标签”标记。最后在“组件样式”中设置图表的标题、背景、图例等，最终效果如图 3.49 所示。

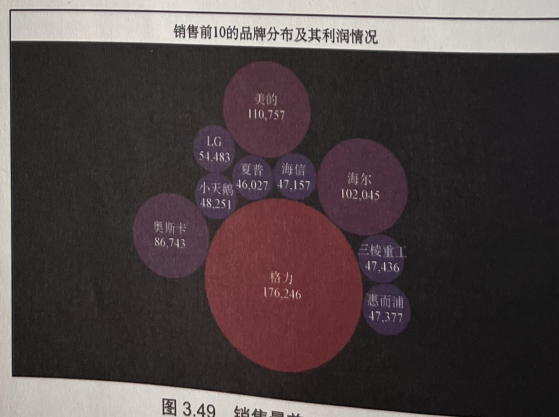


图 3.49 销售量前 10 的气泡图

从图 3.49 中看到，销售量排名前 10 的品牌有格力、美的、奥克斯、海尔、LG、小天鹅

鹅、海信、三菱重工、惠而浦、夏普。其中格力利润最高，美的、奥克斯和海尔的利润比较接近。

【归纳总结】

在本项目中，任务 3.1 使用散点图表示了产品价格与销售量两个变量间的关系。任务 3.2 使用散点图表示各类产品在各价格段的销售量分布情况。任务 3.3 使用气泡图表示销售量前 10 的品牌分布及其利润关系。通过这几个任务的学习，当需要分析数据的联系时，如果是两个变量之间的关系，可以用散点图来表示，如果是三个变量，则可以用气泡图表示。当需要分析数据的分布时，散点图是非常好的选择，一般用于表示数据分布的趋势。直方图也可以表示数据的分布，但一般用于数值数据分布的精确统计。



任务 3.4 空间数据可视化

【任务描述】

针对 2018 年 1~5 月中国空调销售的数据，现在需要对下面的情况进行分析：

- (1) 各地区空调销售量情况。
- (2) 各地区空调利润情况。

【知识准备】

1. 与地理位置相关的数据表示

数据地图是可视化数据的强大方式。当我们面对的数据包含空间位置和地理信息，或者业务要求按空间位置和地理信息分析数据时，这就涉及空间数据可视化问题。通过交互式空间数据可视化的方法，在地图上绘制数据帮助业务人员发现与数据相关的特点并诊断业务问题，普遍用于分析零售、物流、交通、粮食、能源、环保等领域。

很多场景中经常看到用在地图层上叠加与地理位置相关的数据，从而更好地展示与位置相关的数据的特征。我们将这样的可视化方式为数据地图。商业数据时代大量数据是与地理位置相关的，例如，各区域的销售额和利润数据等。用数据地图来反映这些信息比表格要直观形象，且更具交互性。

2. 常用的数据地图图表

填充地图即根据某个度量值，对图中不同区域用不同颜色进行填充所生成的数据地图。利用填充地图区域辨识率高这一特点可以一目了然看到各个地理区域的分布情况。

符号地图是在地图上标记信息的一种数据地图。符号地图适合显示各个位置的定量值。每个位置的定量值可以是一个或两个，一个用颜色进行编码，一个用大小进行编码。

流向地图是显示起点和一个或多个终点位置之间的路径的方法，通常用于表示数据流动的方向和路径。

策是否真正能够驱动业务产生价值。

如果做出相关业务决策之后，业务数据的确发生了改观，那么说明之前的数据分析工作确实找到了业务的实际问题所在。否则，就需要返回到第一步继续进行思考，分析之前是否考虑不周或者是存在偏差。

【任务实施】

根据数据分析的基本流程，在该任务中，首先我们需要熟悉问题的本质。其实超市购物篮系数偏低，我们需要做的就是找到原因，提升购物篮系数。

然后确定分析计划，分解出分析的维度和指标，可以具体分解为三个主要维度：购物篮系数和时间的关系；购物篮系数和顾客购买行为的关系；购物篮系数和商品缺货的关系。

接着是执行分析计划，根据前面得出的分析维度，以周和日为单位，观察购物篮系数的变化规律；调查低购物篮系数的顾客群体和购买行为；并调查商品缺货是否会影响购物篮系数。在此基础上，确定真正需要的数据为商品订单明细、顾客消费数据、每日商品缺货率和购物篮系数数据。

有了这些工作基础，下面就要采用相关的数据分析方法进行具体分析，并得出最终的决策。



数据分析的流程视频讲解

【归纳总结】

通过该任务，我们明确了数据分析的 7 个步骤。根据这些步骤能够帮助我们快速搭建一个清晰的数据分析思路框架。其实这些步骤，可划分为以下 3 个阶段：

- (1) 构建问题，也就是根据业务场景，识别关键问题。
- (2) 解决问题，也就是通过建模收集数据，并分析数据。
- (3) 传达结果并基于结果采取行动，比如可通过可视化图表展现分析结果，并根据分析结果做出决策。

当然，真正的数据分析没有一个统一的标准步骤，需要我们结合具体的分析任务灵活加以应用和处理。



任务 6.3 常用数据分析方法

【任务描述】

某大型零售集团，商品的销售额和毛利率一直表现良好。但是集团经理发现 8 月的销售额环比提升了 12.31%，毛利率却下降了 11.9%，毛利率环比更是下降了 21.41%。作为该大型零售集团的数据分析师，我们要完成经理交给的任务：根据商品和订单的历史数据，找出影响毛利率的关键要素，从而解决集团遇到的困境。在完成该任务的数据分析中，可以从哪些角度进行分析，用到哪些常用的数据分析方法呢？

【知识准备】

在进行数据分析时，数据分析的7个基本步骤辅以常用的经典数据分析方法，可以让我们更加灵活地应对不同业务场景下的数据分析问题。常见的数据分析方法主要有趋势分析法、对比分析法、细化分析法、象限分析法等几种。

1. 趋势分析法

趋势分析法是日常数据分析工作中最常用的方法，它按照时间的维度，对某一数据或者不同数据变化趋势进行差异化研究，以及对数据的下一步变化进行预测。趋势分析法一般而言，适用于对产品核心指标的长期跟踪。因此，趋势分析法的要点是建立一张数据指标的折线图或者柱状图，然后结合不同维度数据进行原因分析，最后进行有效预测，得到趋势分析结果。

2. 对比分析法

数据的趋势变化独立地看，其实很多情况下并不能说明问题，比如如果某个企业盈利增长10%，我们并无法判断这个企业经营的好坏，如果这个企业所处行业的其他企业的盈利普遍为负增长，则盈利增长5%也显得格外亮眼，而如果行业其他企业的盈利增长平均为50%，则盈利增长5%则是个很差的数据。也就是说，数据的一个非常重要的特点是相对性，对比分析正是利用相对性找到数据的变化特点和发展趋势，简单来说就是找差异，以及找出影响这种差异的原因、优化差异的方法。

那怎么找差异呢？一般来说需要比较两个或者多个具有关联的数据，可以是比较多少、比较大小、比较快慢，等等。但是这些数据必须要基于统一的指标，最好是在更多的维度下进行对比。

3. 细化分析法

数据一般都是多维的，因此，我们在分析数据过程中，只有通过钻取、切片、切块、旋转等细分方式对多维度指标进行拆分，才能找到产生数据变化的原因，为数据分析决策提供基础，这就是细化分析法。

4. 象限分析法

象限分析法是通过对数据进行两个及两个以上维度的划分，运用坐标的方式将图表区域分为4个象限，画出四象限图，将每个象限的数据表现作为一个类别，图形形状主要以点图呈现，它能够帮助我们快速地将多个分类下的数据按照不同指标进行归类划分，然后非常直观和快速地进行比较和获得分析结果，并让我们针对不同类别的数据制定最佳策略。

【任务实施】

任务 6.3.1 采用趋势分析法分析月毛利率变化趋势

在该任务中，大型连锁企业的商品销售毛利率的趋势变化可采用趋势分析法进行分析。图6.7呈现了商品销售的月度毛利率的变化趋势。在图中，用折线图表示商品的毛利率的变化，而柱形图呈现了每月毛利率环比的变化。



采用趋势分析法分析月毛利率变化趋势
视频讲解

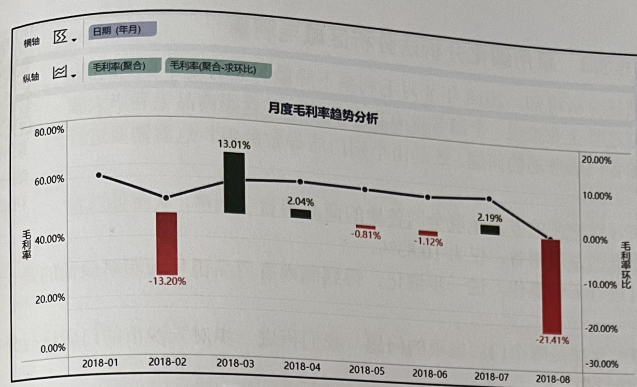


图 6.7 月度毛利率趋势分析

从图中可以很容易看到，2018年2月及2018年8月毛利率环比大幅下降，特别是2018年8月下降幅度更是达到了21.41%，这就是趋势分析的最大优势。当然，简单地画出趋势图，并不算是趋势分析，对于趋势图中明显的拐点，需要我们结合其他维度的数据进行横向、纵向的差异化对比，才能进行完整有效的趋势分析。

任务 6.3.2 采用对比分析法分析不同商品毛利率变化

在图 6.7 中，我们看到某大型连锁企业的商品销售毛利率环比增长率在2018年8月大幅下降，到底是什么原因造成了这种情况呢？因此，可以对不同商品的毛利率变化进行对比分析，从而进一步发现问题的根源。图 6.8 给出了不同商品类别的毛利率变化情况，可以看出，这几种商品的毛利率在2018年8月都呈现了下降的趋势，但是通过对比分析发现，零食类商品的降幅最为明显，因此，我们就需要重点分析零食类商品毛利率下降的关键原因。



采用对比分析法分析不同商品毛利率变化
视频讲解

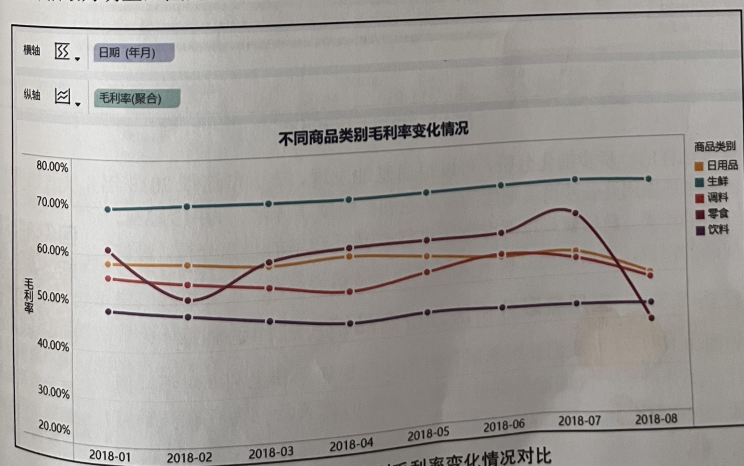


图 6.8 不同商品类别毛利率变化情况对比

任务 6.3.3 采用细化分析法分析区域毛利率

根据前述分析可知, 2018 年 8 月毛利率下降最明显的是零食类商品。对于该大型连锁企业来说, 全国各地有很多门店, 零食类商品毛利率大幅下降是全国各门店普遍的问题, 还是由个别门店导致的呢? 这就需要进行细化分析。

因此可以以地图的方式呈现全国各地的商品销售毛利率, 从而可以看出毛利率最低的是湖南省, 仅为 10.83%。再通过对湖南省商品销售利润率最低的是长沙市, 仅为 -9.76%。

为了探究到底是哪个门店出现的问题, 我们进一步对长沙市的门店进行分析, 得到图 6.9 所示的结果。

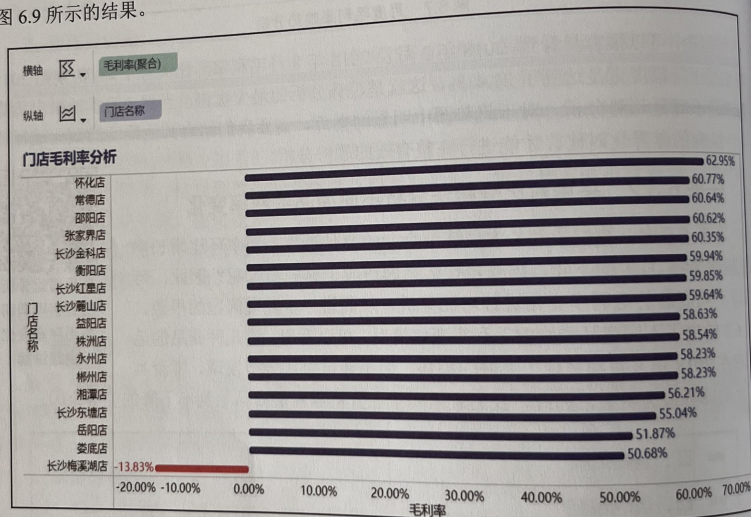


图 6.9 门店毛利率分析

通过这样的一步细化分析, 可以很清楚地发现, 该大型连锁 2018 年 8 月毛利率大幅下降的最主要原因是长沙梅溪湖店的商品毛利率出现了异常, 为 -13.83%。根据分析的结果, 我们就可以再对该店进行产品毛利率的监控分析, 最终找出问题的根源, 为后期提升商品毛利率提供解决方案。

任务 6.3.4 采用象限分析法分析商品销售额和毛利率

如图 6.10 所示, 我们使用象限分析法按照销售额和毛利率对不同商品进行分类分析后, 快速定位到处于第二象限的德芙巧克力存在高销售额、低毛利率这一异常情况。



采用细化分析法分析区域毛利率视频讲解

【归纳总结】

任务 6.3.3 事实, 任务 6.3.4 初步分析来源, 分析毛利率, 可以立存在的, 对具体分析和结合场



采用象限分析法分析商品销售额和毛利率视频讲解

【任务】

在商业 (1) (2)

【知识】

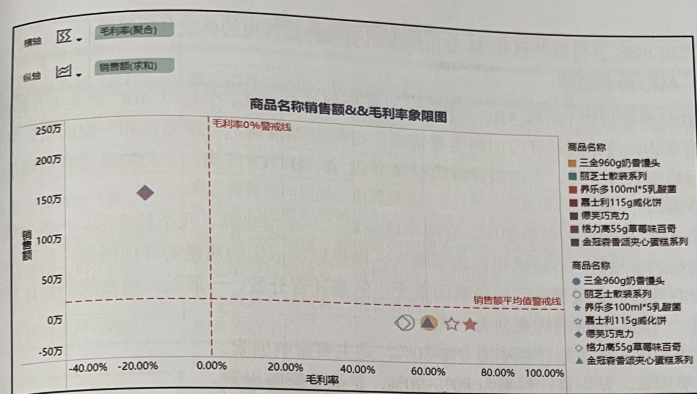


图 6.10 商品名称销售额与毛利率象限图

【归纳总结】

任务 6.3.1 利用趋势分析法分析了月毛利率的变化情况，得到 8 月毛利率严重下滑这一事实，任务 6.3.2 利用对比分析法分析了各种不同商品毛利率的变化趋势，通过对比分析，初步分析得到影响 8 月毛利率的商品类别，任务 6.3.3 利用细化分析法进一步探索问题的根源，分析影响毛利率的具体区域和门店，任务 6.3.4 利用象限分析法分析具体商品的销售额和毛利率，最终确定问题的根源。

可以发现，在本任务中，我们采用了多种数据分析方法，这些数据分析方法并不是孤立存在的。当然，数据分析方法还有很多，它们通常都是以不同的形式展现出来的，在面对具体分析场景时，我们需要清晰知道应用哪一个或几个方法来分析实际问题最为有效，并结合场景灵活运用，没有最好的分析方法只有最适合的。



任务 6.4 常用商业分析模型

【任务描述】

在商业应用领域，比如零售行业，商品和客户是影响利润的最重要的两个因素。因此，在商业数据分析中，需要重点分析下面两个方面的问题：

- (1) 如何区分商品的重要程度？
- (2) 如何根据客户特征对客户进行分类？

【知识准备】

在数据分析工作中，除了常见的数据分析方法，也会经常用到一些经典的商业分析模