

# 目录

一. 题目: .....	3
二. 数据: .....	3
三. 可视化工具.....	3
四. 可视化方案或可视化实现过程.....	4
1.导入数据并进行规范化.....	4
2.数据连接与整理.....	4
3.可视化方案: .....	6
五. 可视化结果.....	6
1. GMV 指标情况以及门店曝光量之间的关系可视化 .....	6
2. 外卖平台利润率关系可视化.....	8
3. 不同门店的营业状况可视化 .....	8
4. 商户补贴和门店曝光量对于商家营收的影响可视化 .....	10
5.研究 CPC 相关数据对平台门店 GMV 的可视化分析 .....	11
6.数据指标周报看板 .....	13
六. 总结.....	16

一. **题目：**网络上下载一组数据（自行获取），选择一种可视化工具（Excel、Tableau、Matlab、Echarts 等），设计一种可视化方案实现该数据的可视化，并做适当的数据分析（或挖掘）。

二. **数据：**包括数据来源、数据内容和数据处理，如果数据量比较大，数据内容可列出小部分则可

本次实验中，我设计的是关于外卖信息的数据可视化。其中数据通过爬虫在网站上进行爬取。数据整合为一张表，这张表格记录了“蛙小辣火锅杯”品牌在 2020 年 1 月至 8 月期间，位于上海的多家门店（如宝山店、五角场店、龙阳广场店等）在饿了么和美团两大外卖平台上的运营数据。数据涵盖每日的销售表现（GMV、商家实收）、用户行为（门店曝光量、访问量、下单量）、订单情况（有效订单、无效订单）、推广费用（CPC 总费用、曝光量）以及补贴信息（商户补贴、平台补贴）。下面是部分数据截图，具体数据见随文档上交的excel 文件。  
其中数据：



三. 可视化工具

在本次作业中，我选择 **Tableau** 作为主要的数据可视化工具，**EXCEL** 作为数据看板的辅助工具主要基于以下关键原因：

高效的交互式探索：**Tableau** 卓越的拖放式界面和即时交互能力（如筛选、下钻、工具提示）使我能够快速探索数据，动态地发现模式、异常值和潜在关系，这对于深入理解作业数据集至关重要。

强大的可视化表现力与灵活性：**Tableau** 提供极其丰富的图表类型和高度定制化选项（颜色、大小、标签、计算字段等），使我能够根据作业需求精确地创建清晰、美观且信息量丰富的可视化，有效传达分析结果。

便捷的仪表板构建与故事叙述：Tableau 的仪表板功能让我能轻松整合多个相关视图，并通过筛选器动作实现联动；其‘故事点’功能则帮助我将分析过程和关键发现组织成一个逻辑连贯、引人入胜的叙事结构，完美呈现作业成果。

卓越的数据处理与连接能力：Tableau 能无缝连接作业中使用的多种数据源（如 Excel, CSV），其高效的数据引擎确保了操作的流畅性，内置的数据准备功能（如数据透视、混合）也简化了前期数据处理步骤。

基于函数与迷你图的 Excel 数据看板，能以轻量化方式实现高效分析。通过‘SUMIFS’、‘AVERAGEIFS’等函数可快速完成多条件数据汇总，结合‘TEXT’函数规范显示格式，显著提升数据可读性。迷你图直接在单元格内生成趋势图表，配合条件格式的数据条/色阶功能，无需占用额外空间即可直观呈现关键指标波动与异常值。利用‘INDIRECT’、‘CHOOSE’等函数与数据验证下拉菜单联动，可轻松实现动态维度切换，满足多场景分析需求。文件无需依赖复杂插件，仅靠公式自动更新即可保持数据时效性，兼容所有 Excel 版本。典型应用如销售看板中，函数计算实时业绩，迷你图对比月度趋势，条件格式标出达成率，形成"计算-展示-预测"一体化解决方案，特别适合需要快速响应的业务场景。

四. 可视化方案或可视化实现过程(可以采用图示方式，如附录图 2)a.报告应至少包含 3 种以上类型的图（可以使用子图、混合子图等）且美观（配色、图例、标注文字说明）；要能正常显示中文；（PS：可以只有一副图，但采用了多种类型的图，见附录图2-3） b. 报告应能体现数据的采集与选取(可略)、数据预处理、数据展示与分析； c. 每小部分的数据分析内容应说明：为什么要使用某种类型的图，该类型的图有什么特点？

## 1.导入数据并进行规范化

选择导入 excel 文件，然后整理文件的格式，让系统可以正常识别。

## 2.数据连接与整理

把 excel 表中的表和详细描述进行连接，然后用 spsspro 数据分析工具进行**数据清洗**。清除掉多余或者无用的项，异常值设置为中位数填充，缺失值设置为 0，如下：



### 3.可视化方案:

- (1) GMV 指标情况以及门店曝光量之间的关系可视化
- (2) 外卖平台利润率关系可视化
- (3) 不同门店的营业状况可视化
- (4) 商户补贴和门店曝光量对于商家营收的影响可视化
- (5) 研究 CPC 相关数据对平台门店 GMV 的可视化分析
- (6) 数据指标周报看板（自动化-EXCEL）

## 五. 可视化结果

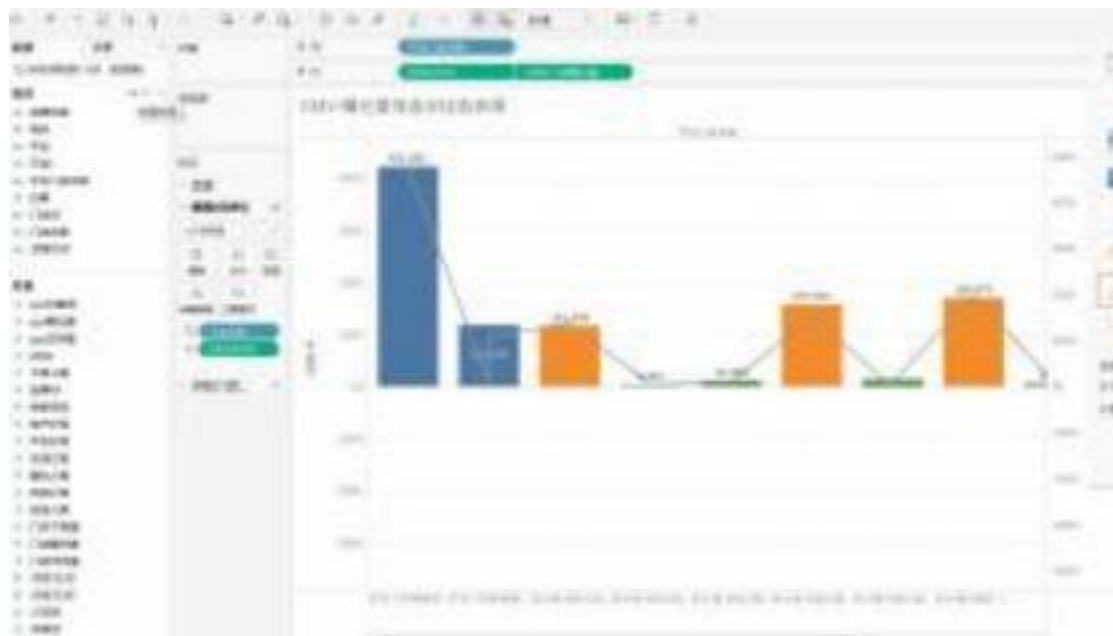
### 1.GMV 指标情况以及门店曝光量之间的关系可视化

使用双轴图的模式，显示出 1 到 8 月各个门店的 GMV 指标情况以及门店曝光量之间的关系。

基本结构：

双轴组合图（左侧 Y 轴-GMV 指标总和的条形图 + 右侧 Y 轴-门店曝光量的折线图）

X 轴为门店名称（蛙小辣·美蛙火锅杯(宝山店)、蛙小辣·美蛙火锅杯(宝山店)等）。条形图特点是能直观对比不同类别（门店）的数据大小，便于快速看出 2020 年 1 到 8 月各门店的门店曝光量的差异，比如哪家流量高、哪家流量低，一目了然。折线图擅长展示数据的变化趋势、波动情况，能清晰反映各门店的曝光的高低起伏，辅助分析 GMV 与曝光量的关联逻辑。（GMV 是数据分析中的重要指标，是指一定时间内通过平台或门店达成的所有订单的总金额）



可以看到，左轴的GMV有正负值，表示盈利和亏损。纵观整个图，可以发现门店曝光量与门店 GMV 存在正相关关系。高 GMV 的门店往往有着高曝光（如干拌麻辣烫（武宁路店）与蛙小辣火锅杯（宝山店）等），但是，有异常情况如第二组干拌麻辣烫有相对较高的 GMV，但是低曝光，同时第三组的蛙小辣宝山店有这高曝光但是低 GMV，所以门店运营存在“流量陷阱”——高曝光不等于高转化，需优化用户路径。但是整体遵循曝光与GMV 正相关的规律。

## 2. 外卖平台利润率关系可视化



基本结构：柱状图

首先创建利润率计算字段， $\text{利润率} = (\text{商家实收} - \text{总成本}) / \text{商家实收}$

$([\text{商家实收}] - [\text{cpc 总费用}] - [\text{商户补贴}] - [\text{平台补贴}]) / [\text{商家实收}]$

创建平台（美团，饿了么）的筛选器，以及日期（月为单位）筛选器，根据图，可以看出任何平台的各个门店的利润率都为负，可能是收到疫情影响。这张平台门店利润分布图通过对 437 家门店的利润数据可视化，清晰揭示了三个核心发现：首先，利润分布呈现显著头部效应，前 50 家门店贡献了绝大部分利润，而 60% 以上的长尾门店利润表现疲软，反映出资源分配不均的问题。其次，图中潜在的负利润门店暴

露了补贴政策失效风险，需重点核查高成本低转化的门店运营数据。其次，我在把门店曝光拉到大小，发现总体上还是曝光越大，利润率越大。（下图）



根据图，可以看出较为优质的门店数量较多，集中在左上方。代表高流量、高利润、高订单，有一个门店进行低调盈利，低流量但高转化率。有两个门店陷入流量陷阱，高流量但低订单/利润。同时，值得关注的是最右边的离群点。该店是武宁路的干拌麻辣烫店，亏了 13 万，-79%的利润率，但是进店人数高达 3 万每月，说明该店的有效订单不高，转化率低。需要优化运营策略。





### 3. 对比与结论

高离散 & 极值型：“拌客干拌麻辣烫（武宁路店）”，GMV 波动极大，需关注极值订单（如大额团购、异常交易），或分析高业绩时段 / 活动的可复制性。

低离散 & 稳定型：多数“蛙小辣美蛙火锅杯”门店，GMV 分布集中，业务稳定性强，可挖掘共性运营策略（如品控、营销复用）。

一方面箱线图可以快速定位“异常门店”（如拌客武宁路店的高极值），排查数据真实性或业务特殊性；同时为门店分层管理提供依据，例如重点优化离散高的门店、复制稳定门店经验等。

### 4. 商户补贴和门店曝光量对于商家营收的影响可视化

基本结构:填充气泡图

填充气泡图特点：

是一种直观且信息丰富的数据可视化形式，结合了位置、大小、颜色三个视觉通道传递信息。

多维度展示能力：大小（气泡尺寸）：绑定一个数值字段，气泡越大代表该字段数值越高，直观呈现数据量级差异，像本文中展示门店曝光量。颜色（气泡填充色）：能关联分类（区分不同类别，如产品类型）或数值（通过渐变色体现数据大小、占比等，如利润率高高低），进一步丰富信息层次。本文用来展示不同门店商户补贴数据。

配合 Tableau 的交互功能（筛选器、工具提示等），还能叠加更多隐藏维度，鼠标悬停气泡时，工具提示可显示详细字段（如具体日期、客户名称），点击筛选器切换分析视角（不同业务线、不同时间区间），让可视化具备动态、多层次分析能力。



选择研究商户补贴与门店曝光量对于商家营收的影响关系。

「总计（门店曝光量）」：驱动气泡大小，体现曝光维度对商家的影响。

「总计（商家实收）」：以标签形式呈现，直接展示营收结果。

「总计（商户补贴）」：控制气泡颜色，反映补贴投入维度。

用 Tableau 气泡图的典型优势，把 “补贴、曝光、营收 ” 三个关键运营指标，通过 颜色、大小、标签 三个视觉通道融合展示 → 无需复杂计算，一眼识别：

1. 哪些门店 “补贴高 + 曝光高 + 营收高 ”（优质运营组合）；
2. 哪些门店 “投入（补贴 / 曝光）” 与 “产出（实收）” 不匹配（需优化策略）；
3. 不同门店在 “补贴 - 曝光 - 营收 ” 体系中的差异化表现。

分析与总结：门店 ID “337460136 ” 的营收最高，而且是断层式排列首位，同时还能发现该门店气泡大小最大，且颜色最深。说明一个门店的营收收入与这个门店的商户补贴及曝光量正向相关。其次营收较高的是门店 ID “2001104355 ”，该门店在可视化图中展示出如同门店 ID “337460136 ” 一样的信息。又证明了三者的正相关关系，该图表给与的指示即如果需要提高门店的营收，可以从门店曝光量及增加商户补贴方面入手。

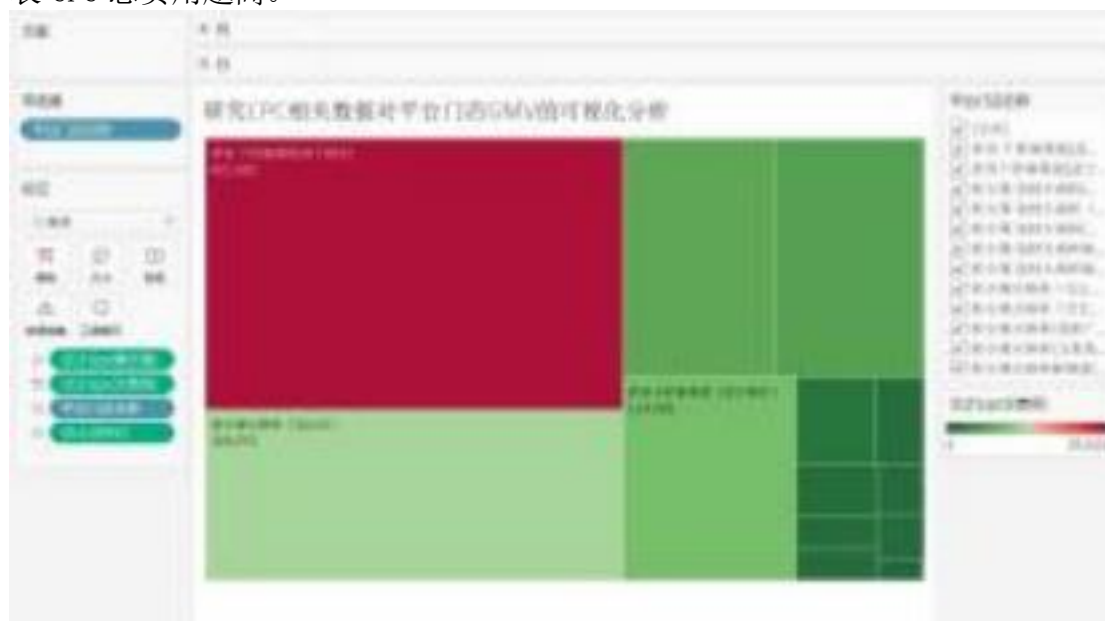
## 5. 研究 CPC 相关数据对平台门店 GMV 的可视化分析

基本结构：树状图

树状图特点：由嵌套矩形构成，用矩形展示选定维度内的类别，呈现分层 “树 ” 状结构，能在有限空间对比数量与模式，展现部分与整体关系。

维度与度量：用维度定义结构，度量定义矩形大小或颜色。像 “平台门店名称 ” 维度构建层级，“CPC 曝光量 ” “CPC 总费用 ” 等度量决定矩形大小、颜色。

大小与颜色：大小一般关联度量，体现数值规模；颜色可关联维度（用分类调色板区分类别，如不同运输方式设不同颜色 ）或度量（用连续调色板，如按销售数字、利润设颜色渐变，深色突出极端值 ）。此处选择红绿发散，红色越深代表 CPC 总费用越高。



1. 门店 GMV 规模差异：

“拌客干拌麻辣烫（武宁路店）” 矩形面积最大，GMV 最高（421,832），是当前分析中营收核心门店。

“蛙小辣火锅杯（宝山店）”“拌客干拌麻辣烫（武宁路店小矩形）”等面积较小，GMV 相对低（如 169,975、114,008）。

## 2. CPC 数据的辅助分析：

虽未直接用颜色区分，但可通过筛选器 + 工具提示挖掘：

筛选特定门店（如“拌客干拌麻辣烫（武宁路店）”），查看其“cpc 曝光量”“cpc 总费用”，分析“高 GMV 是否伴随高 CPC 投入”。

对比不同门店的 CPC 数据与 GMV 关系，找“CPC 投入效率差异”（如小 GMV 门店是否因CPC 不足导致，或高 CPC 投入但 GMV 低的优化空间）。

3. 直观定位重点门店：通过面积快速识别GMV 头部门店（如“拌客干拌麻辣烫（武宁路店）”），优先分析其 CPC 策略复用性。

4. 关联 CPC 投入产出：结合标记卡的 CPC 字段，延伸分析“CPC 曝光、费用如何影响 GMV”，辅助优化投放策略（如调整低效率门店的 CPC 预算）

## 6. 数据指标周报看板搭建（EXCEL）

首先确定数据源，即源数据。然后通过 EXCEL 的跨表协作。在数据源表上另开新表，即看板表。然后通过 EXCEL 的函数公式引用，以周为单位进行时间的索引，然后运用 sumifs, vlookup, 和 if 等相关数据分析常用 Excel 函数对各个指标进行相关联，使得数据指标看板可以根据时间进行实时变换，便于后续进行指标因动归因分析。（数据看板样例如下）



实时化体现:

由于指标函数共同引用 A13, 所以改动 2020/8/10 时, 整个数据看板就会对于源数据的数据进行变化。如下图:



同时，在数据指标看板的右上角还可以对平台进行筛选，洞察外卖平台不同的指标变化，如下：

1.观察美团平台 2020/6/16——2020/6/22 的指标





2. 观察饿了么平台 2020/6/16——2020/6/22 的指标



(如果需使用数据看板验证, 在报告同一文件下有 Excel 数据源与看板文件。

六. 项目总结

该项目深入挖掘了“蛙小辣火锅杯”品牌 2020 年 1-8 月的外卖运营数据, 揭示了隐藏在数据背后的关键故事。首先, 曝光量与 GMV 的关联分析(双轴图)显示二者存在显著正相关性, 但发现“武宁路店”等高曝光门店出现 GMV 低迷的“流量陷阱”, 说明平台引流需与门店转化能力匹配; 其次, 利润率可视化(散点图+平台筛选)暴露出所有门店均处于亏损状态(利润率全为负值), 结合疫情时间背景, 推测外部环境对餐饮业造成系统性冲击。进一步通过门店营业状况箱线图分析, 发现“武宁路店”GMV 离散度极高(存在 422K 异常值), 而其他门店业绩稳定, 提示需排查异常订单或优化头部门店运营策略。

在归因分析中, 气泡图清晰呈现商户补贴、曝光量与营收的三维关系: ID “337460136” 门店通过高补贴与高曝光的协同实现断层式营收领先, 验证了资

源投入对业绩的正向驱动。而树状图对 CPC 费用的解析表明，“武宁路店”虽 CPC 投入最高，但未有效转化为 GMV 增长，指向其转化链路存在瓶颈。最终，通过 Excel 动态看板的周维度监控，发现美团平台在 6 月第三周订单量骤降 15%，而饿了么同期补贴成本上升 20%，为平台策略调整提供了实时依据。

数据可视化不仅验证了“曝光-转化-营收”的基础逻辑，更暴露了疫情下外卖行业的经营困境。异常门店的深度剖析（如武宁路店的高投入低转化）说明精细化运营的必要性，而动态看板则为资源调配建立了敏捷响应机制。未来可结合用户评论数据进行情感分析，进一步优化产品与服务的薄弱环节。