

## “给咖啡店提供下个月经营策略”（报告最后两页）

1, 推出

【coffee coffee】

【coffee bread】

【coffee pastry】

【coffee cake】

【coffee medialuna】

五种 combo 并给予 95 折优惠。

2, 可以考虑增加 pastry、cake、medialuna 三类商品的进货。

3, 细分套餐时段：

早间套餐（8am-12pm）：coffee + medialuna、coffee + pastry

午间套餐（1pm- 3pm）：coffee + sandwich、coffee + cake

全天套餐：                    coffee + bread

## PART 1) 数据理解:

sql 部分,

已有 2 张销售数量的统计表; 分别是 by hour 和 by month 两个 table, 经过简单验证, 总数一致。

SQL:	BY HOUR	BY MONTH
ITEM	Sales_count	Sales_count
the next two queries have the same output	<pre>SELECT   SUM(CAST(sales_volume as int)) as all   , item FROM hour_item_sales GROUP BY item ORDER BY item</pre>	<pre>SELECT   SUM(CAST(sales_volume as int)) as all   , item FROM month_item_sales GROUP BY item ORDER BY item</pre>
query used	<pre>SELECT * FROM hour_item_sales ORDER by item</pre>	<pre>SELECT * FROM month_item_sales ORDER by item</pre>

-----

excel 部分, 有3张表格,

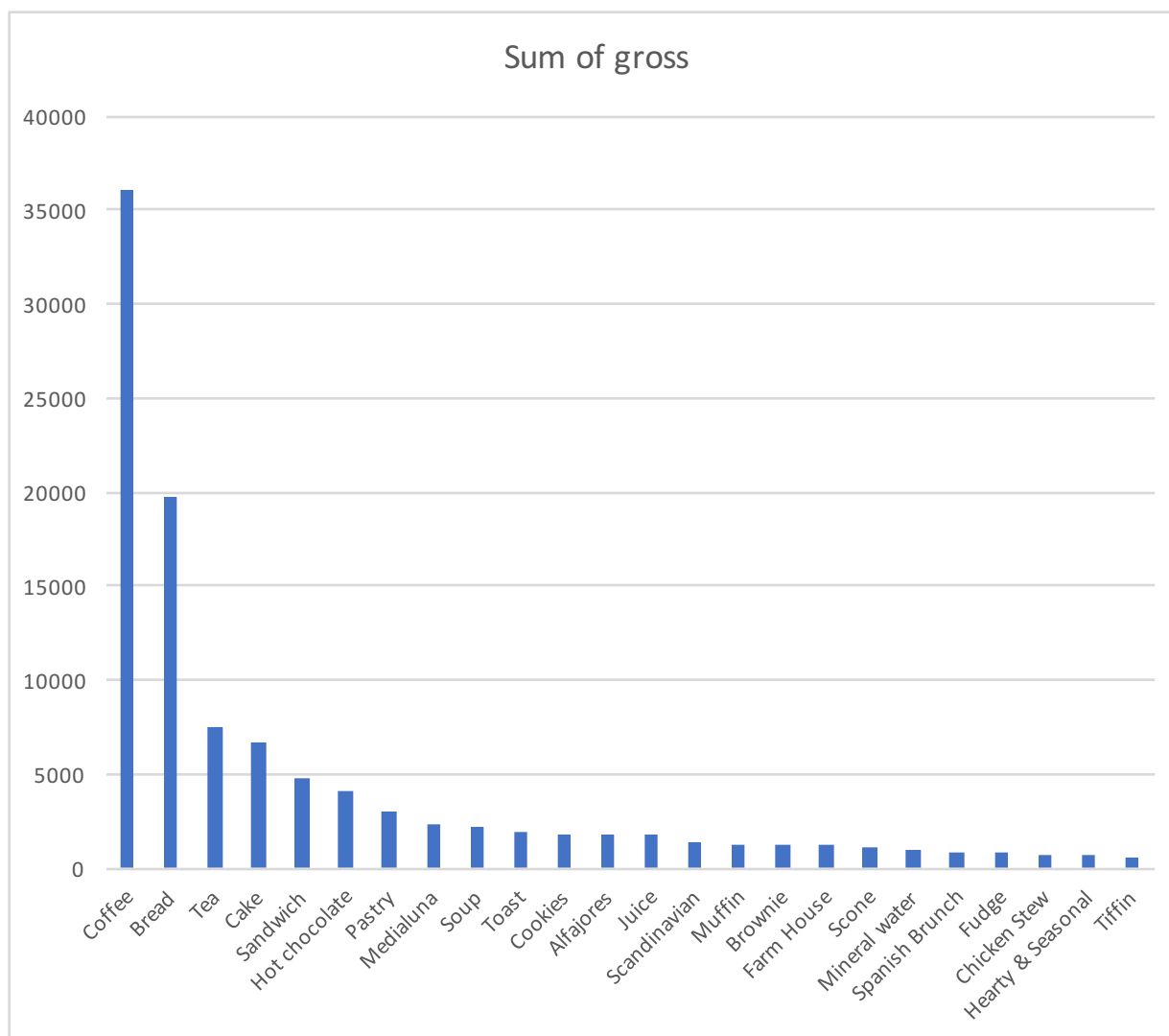
- 第一张为 单品 销售数量,
- 第二张为 商品单价,
- 第三张为 订单 时间 ID 和 清单。

EXCEL	Item_Total_Sales_Count	Item_Price	Order_info
ITEM	Sales_count	Price	Date , Time , OrderID , # of Items , List of Items

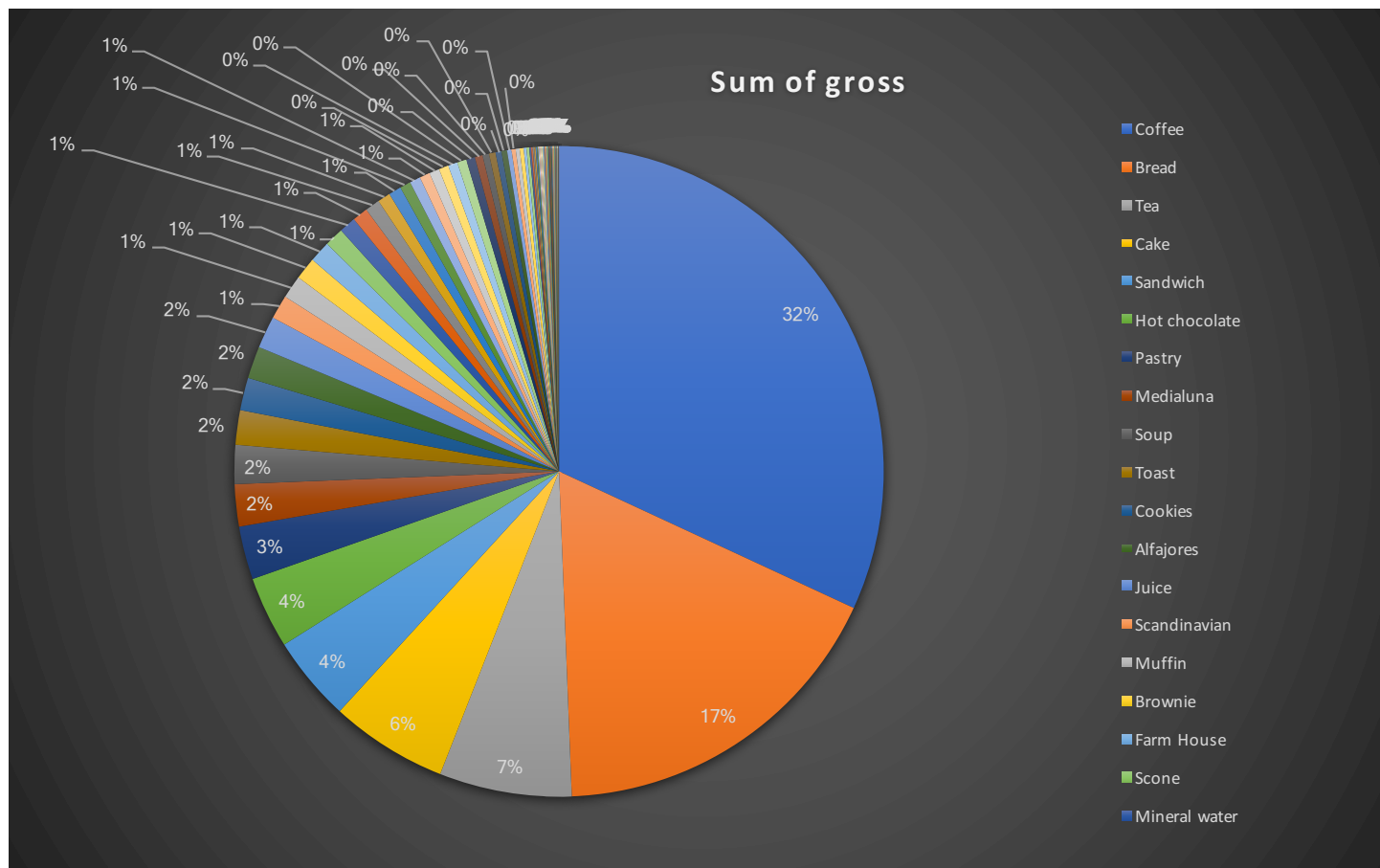
## PART 2) Excel 数据处理

通过 vlookup function 汇总 表一 sales count 和 表二 item price 的 单价 与 销售数量 相乘, 得出 单品总销售额 Gross per item。

- 各单品销售额前 25 的 Bar Chart 分布情况如下:
- 销售分布近似 Geometric Distribution, 其中 Coffee 和 Bread 占据了总销售额的绝大部分。



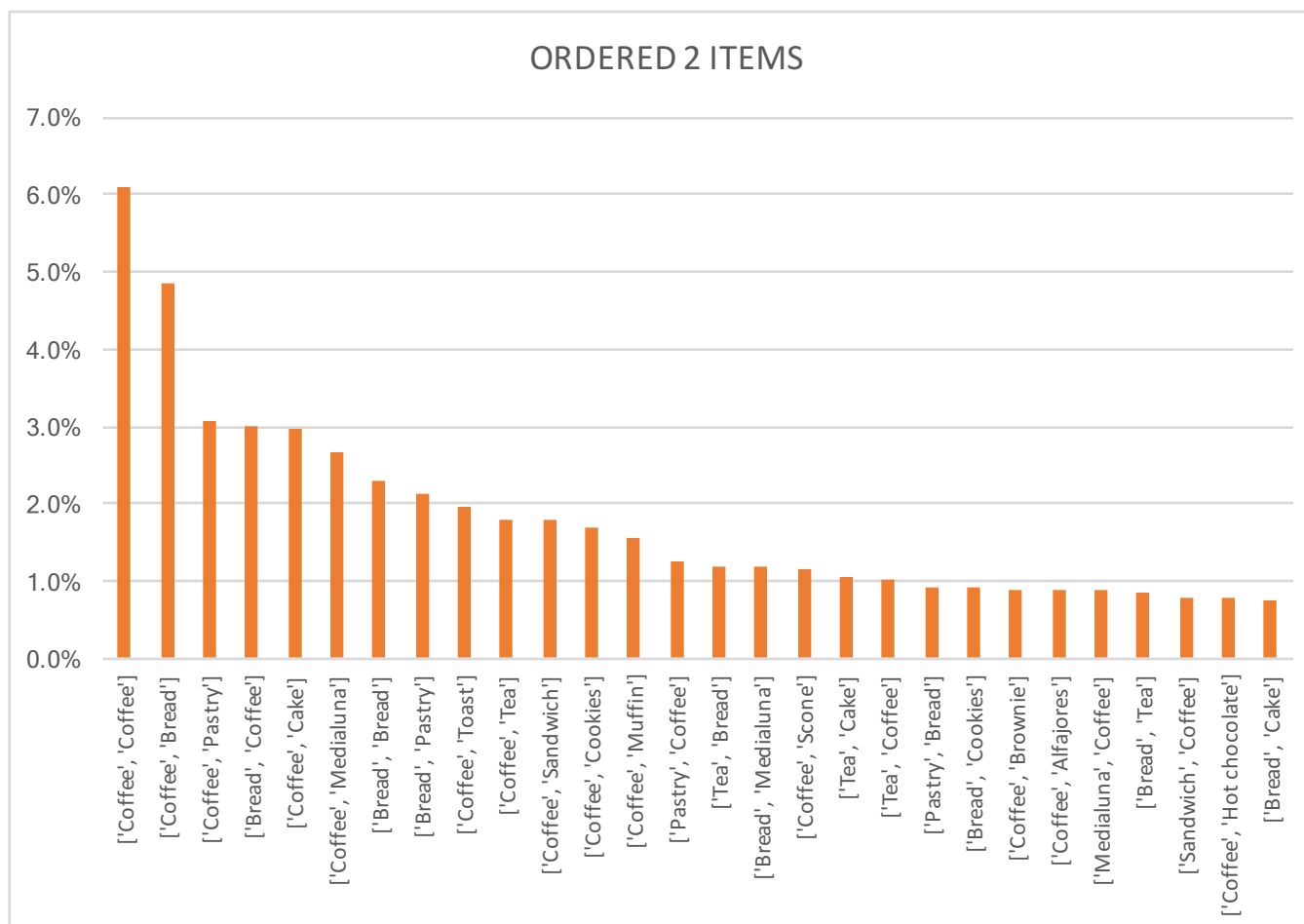
为了更直观的表达销售额分部情况，我们这里用到 Pie Chart。  
根据如下表格得出初步结论：49% 的销售额是由 coffee 和 bread 提供的。



- 表三，根据 order info 生成 pivot table，找出最多的 2 item combo，并且计算最大的 combo 的售价。  
将订单按照每个订单包含的商品数量分类：获得出售最多的 n 个 combo 组合 (Number\_of\_item = 1, 2, 3....)  
为简化方案的复杂度，这里只讨论单次购买小于等于 3 个商品组合的订单。



(方案1: ) 前五个 2 item combo: 【coffee coffee】 【coffee bread】 【coffee pastry】 【coffee cake】 【coffee medialuna】  
 因此, 可以推出这五个套餐一起搭售。又, 根据前表, coffee bread 单卖原本销量就很好, 可以考虑 (方案2: ) 增加 pastry cake medialuna 三类商品的进货, 以备可能的需求增长。



PART 3) SQL - CSV 数据处理

	Error	Comment
By hour	Bread 这个商品 1am 有 1 个销售记录， 1am 不营业所以剔除	medialuna 和 pastry 早间售出更多， sandwich 和 cake 午间售出更多。 <b>(方案3: )</b> 细分套餐时段
By month		趋势平稳， 过去六个月没有异常波动

