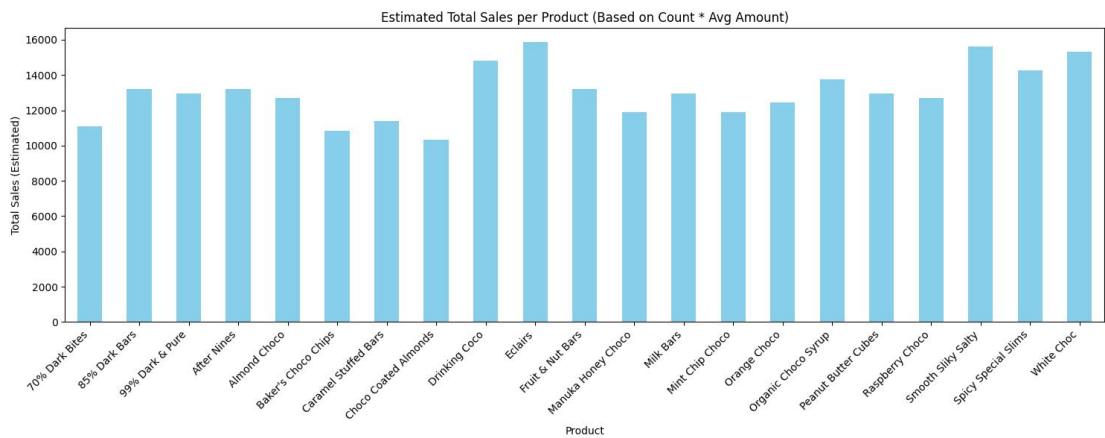


41.



**解释：** 这幅图展示了不同产品的估算总销售额。以下是详细解释：

图表标题

- 标题: "Estimated Total Sales per Product (Based on Count \* Avg Amount)"  
表示这是基于产品出现次数和平均金额计算的估算总销售额。

坐标轴

- X 轴: "Product"  
表示不同产品的名称。
- Y 轴: "Total Sales (Estimated)"  
表示估算的总销售额。

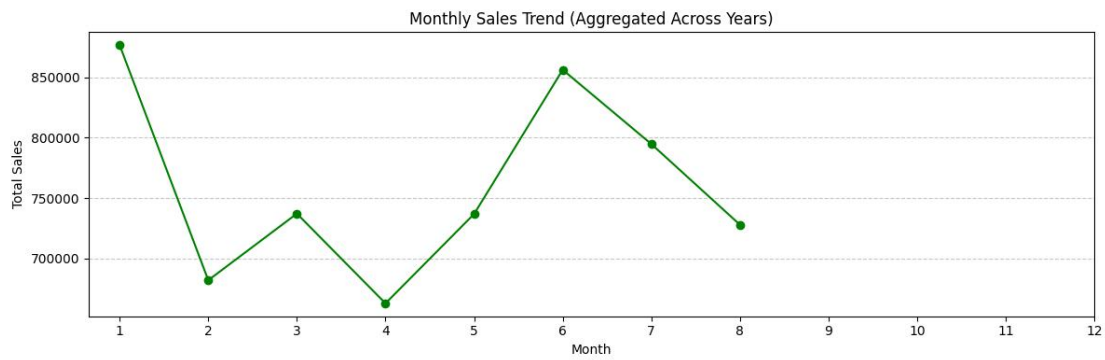
柱状图数据

- 每个蓝色柱子代表一个产品的估算总销售额。
- 从图中可以看出，不同产品的销售额有较大差异。

数据特点

- Choco Syrup 和 Spice Special Slims: 销售额明显高于其他产品。
- White Choc 和其他几款产品: 销售额相对较低。

42.



**解释：** 这幅图展示了按月汇总的年度销售趋势，横轴是月份，纵轴是总销售额。以下是对这

幅图的具体解释：

图表标题

- **Monthly Sales Trend (Aggregated Across Years):** 表示图中显示的是跨年度汇总的每月销售趋势。

坐标轴

- **X 轴:** 标记为“Month”，表示月份，从 1 月到 12 月。
- **Y 轴:** 标记为“Total Sales”，表示总销售额，单位未标明，但数值在 700,000 到 900,000 之间。

数据趋势

- 数据点以绿色折线和圆点标出，显示了每个月的销售总额。
- 整体趋势显示销售额在一年中的变化，有明显的波动。

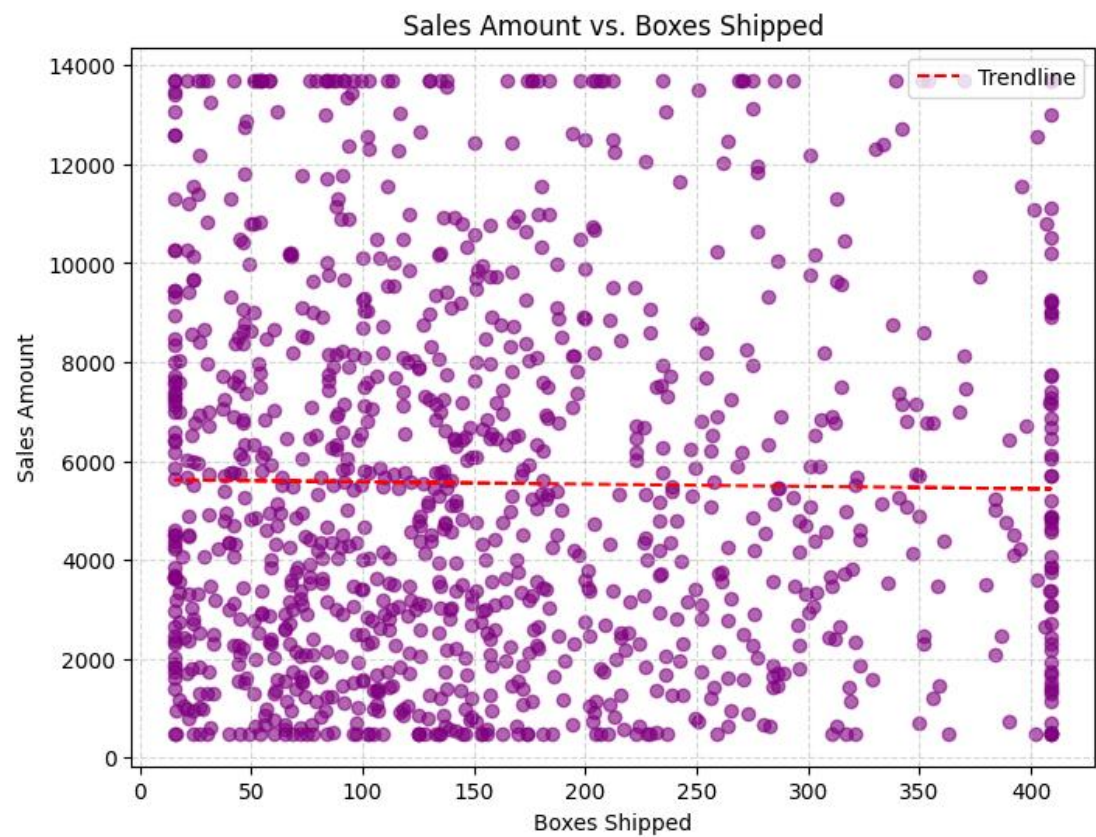
月度销售额变化

- 1 月开始销售额较高，接近 900,000。
- 从 2 月开始，销售额急剧下降，3 月略有回升。
- 4 月销售额再次下降到全年最低点。
- 5 月和 6 月销售额逐步上升，6 月达到全年最高点，超过 900,000。
- 从 7 月开始销售额开始下降，直到 8 月。
- 9 月到 12 月销售额较为平稳，没有显著的上升或下降趋势。

总结

这幅图表明销售额在一年中有明显的季节性波动，可能与节假日、促销活动或其他季节性因素有关。企业可以根据这种趋势进行库存管理、营销活动规划等，以优化销售业绩。

43.



解释：这幅图是一个散点图，展示了发货箱数（X 轴）与销售额（Y 轴）之间的关系。以下是对这幅图的具体解释：

图表标题

- **Sales Amount vs. Boxes Shipped:** 表示图中显示的是销售额与发货箱数之间的关系。

坐标轴

- **X 轴:** 标记为“Boxes Shipped”，表示发货的箱数，范围从 0 到 400。
- **Y 轴:** 标记为“Sales Amount”，表示销售额，范围从 0 到 14000。

数据点

- 每个紫色的点代表一个数据样本，显示了特定发货箱数对应的销售额。
- 数据点的分布较为分散，显示出销售额与发货箱数之间的关系并不简单。

趋势线

- 图中有一条红色虚线，标记为“Trendline”，表示通过数据点拟合的趋势线。
- 这条趋势线显示了销售额随发货箱数变化的总体趋势，尽管数据点分布较为分散，但趋势线显示出一个轻微的负相关关系，即随着发货箱数的增加，销售额有轻微下降的趋势。

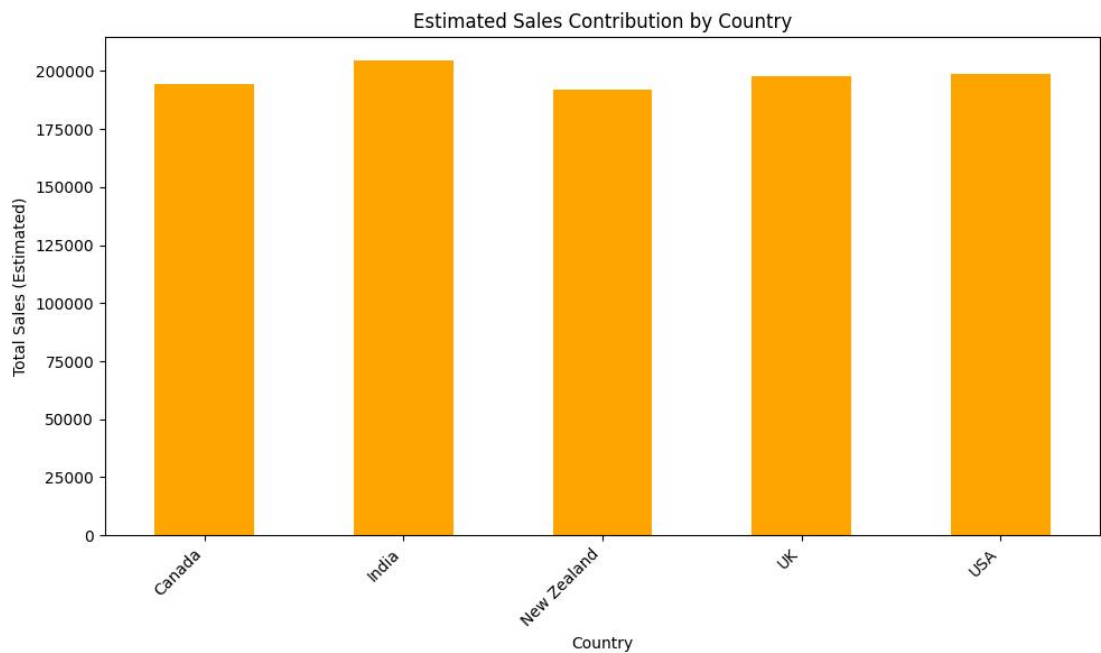
数据特点

- 数据点在各个发货箱数区间内都有广泛的分布，特别是在箱数较少（0-100）和较多（300-400）的区间内。
- 销售额的分布范围较广，从接近 0 到超过 14000，显示出不同发货箱数下销售额的变异性。

总结

这幅图表明，虽然存在一个轻微的负相关趋势，但销售额与发货箱数之间的关系并不明显或强烈。这可能意味着其他因素（如产品类型、市场需求等）对销售额有更大的影响。对于企业来说，了解这种关系有助于优化物流和销售策略。

44.



解释：这幅图是一个柱状图，标题为“Estimated Sales Contribution by Country”，表示不同国

家对总销售额的估计贡献。以下是对图中信息的具体解释：

图表标题

- **Estimated Sales Contribution by Country:** 表示图中展示的是按国家划分的估算销售贡献。

坐标轴

- **X 轴:** 标记为“Country”，显示了各个国家的名字，包括加拿大（Canada）、印度（India）、新西兰（New Zealand）、英国（UK）和美国（USA）。
- **Y 轴:** 标记为“Total Sales (Estimated)”，表示估算的总销售额，单位未标明，但数值范围从 0 到超过 200,000。

柱状数据

- 每个橙色柱子代表一个国家的估算总销售额。
- 从图中可以看出，各个国家的销售额之间存在一些差异。

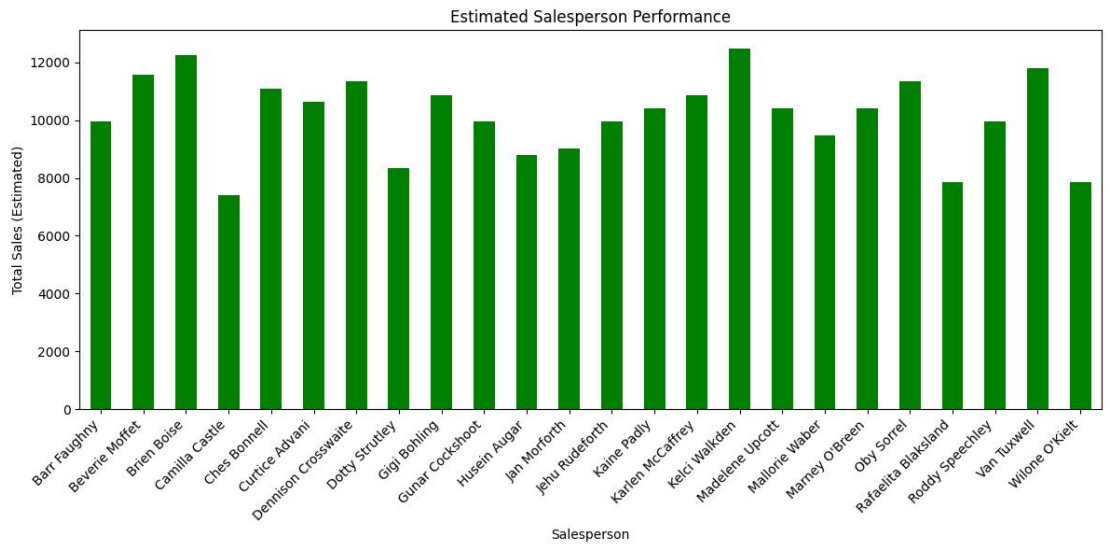
销售贡献

- 印度的估算销售额最高，其次是美国和英国，而加拿大和新西兰的销售额相对较低但相差不多。
- 估算销售额在各个国家之间分布相对均匀，但印度略高于其他国家。

总结

这幅图表明不同国家对总销售额的贡献程度。对于跨国公司来说，这种信息对于了解不同地区的市场表现和制定针对性的市场策略非常有价值。根据销售额的分布，公司可能会考虑在销售额较高的国家增加投资或在销售额较低的国家寻找增长机会。

45.



**解释：** 这幅图是一个柱状图，标题为“Estimated Salesperson Performance”，表示各个销售人员的估算销售业绩。以下是对图中信息的解释：

图表标题

- **Estimated Salesperson Performance:** 表示图中展示的是各个销售人员的估算销售表现。

坐标轴

- **X 轴:** 标记为“Salesperson”，显示了不同销售人员的姓名。

- **Y 轴:** 标记为“Total Sales (Estimated)”，表示估算的总销售额，单位未标明，但数值范围从 0 到超过 12,000。

### 柱状数据

- 每个绿色柱子代表一个销售人员的估算总销售额。
- 从图中可以看出，不同销售人员的业绩存在显著差异。

### 销售表现

- 一些销售人员（如 Bren Boise、Katia Walkden、Roddy Speechley）的销售额超过 12,000，表现突出。
- 其他销售人员（如 Camila Castle、Oby Sorrel、Wilone Oxxelt）的销售额相对较低，未达到 10,000。
- 大多数销售人员的销售额集中在 8,000 到 12,000 之间。

### 总结

这幅图清晰地展示了各个销售人员的销售业绩差异，帮助识别表现最佳和最差的销售人员。对于销售团队来说，这种可视化信息有助于发现优秀员工并提供改进方向，从而优化团队整体业绩。

## 51.

```
Decision Tree training time: 0.01 seconds
Random Forest training time: 0.68 seconds
XGBoost training time: 1.47 seconds
[LightGBM] [Warning] Found whitespace in feature_names, replace with underlines
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.000078 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 360
[LightGBM] [Info] Number of data points in the train set: 875, number of used features: 58
[LightGBM] [Info] Start training from score 5499.407537
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
LightGBM training time: 0.29 seconds
CatBoost training time: 1.64 seconds
Models trained successfully.
```

### 解释：模型训练时间

- **Decision Tree training time: 0.01 seconds:** 决策树模型的训练时间非常短，仅为 0.01 秒。这表明决策树模型训练速度快，但通常其预测性能可能不如更复杂的模型。
- **Random Forest training time: 0.68 seconds:** 随机森林模型的训练时间稍长，为 0.68 秒。随机森林由多棵决策树组成，训练时间随着树的数量增加而增加。
- **XGBoost training time: 1.47 seconds:** XGBoost 模型的训练时间为 1.47 秒。XGBoost 是一种高效的梯度提升模型，通常在处理大规模数据时表现良好。
- **LightGBM training time: 0.29 seconds:** LightGBM 模型的训练时间为 0.29 秒。LightGBM 是一种优化的梯度提升模型，通常在训练速度和内存使用方面优于 XGBoost。
- **CatBoost training time: 1.64 seconds:** CatBoost 模型的训练时间为 1.64 秒。CatBoost 对于处理分类特征有独特的优势，但训练时间可能较长。

### LightGBM 模型训练日志

- **[LightGBM] [Warning] Found whitespace in feature\_names, replace with underlines:** 警告信息，表示特征名中包含空格，建议用下划线替换。这通常不会影响模型训练，但可能会影响特征的重要性和解释性。
- **[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was**

0.000078 seconds. You can set `force_row_wise=true` to remove the overhead. And if memory is not enough, you can set `force_col_wise=true`.: 信息提示, 表示 LightGBM 自动选择了行-wise 多线程处理方式, 并提供了优化建议以减少开销。

- **[LightGBM] [Info] Total Bins 360:** 信息提示, 表示 LightGBM 将特征值分成了 360 个桶 (bins), 用于加速训练。
- **[LightGBM] [Info] Number of data points in the train set: 875, number of used features: 58:** 信息提示, 表示训练集中有 875 个数据点, 使用了 58 个特征。
- **[LightGBM] [Info] Start training from score 5499.407537:** 信息提示, 表示训练的起始分数为 5499.407537。
- **[LightGBM] [Warning] No further splits with positive gain, best gain: -inf:** 警告信息, 表示在某些情况下, 没有找到具有正增益的进一步分裂点。这通常表示模型在这些情况下无法进一步改进。

## 总结

- **训练时间:** 展示了不同模型的训练时间, 决策树模型最快, CatBoost 最慢。这种信息对于模型选择和优化非常重要, 尤其是在处理大规模数据时。
- **LightGBM 日志:** 提供了关于训练过程的详细信息和警告, 帮助开发者了解模型训练的细节并进行优化。

## 52.

```
Decision Tree training time: 0.01 seconds
Random Forest training time: 0.68 seconds
XGBoost training time: 0.10 seconds
[LightGBM] [Warning] Found whitespace in feature_names, replace with underlines
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.000055 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 360
[LightGBM] [Info] Number of data points in the train set: 875, number of used features: 58
[LightGBM] [Info] Start training from score 5499.407537
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
LightGBM training time: 0.04 seconds
CatBoost training time: 1.46 seconds
Models trained successfully.
```

## 解释: 模型训练时间

- **Decision Tree training time: 0.01 seconds:** 决策树模型的训练时间非常短, 仅为 0.01 秒。
- **Random Forest training time: 0.68 seconds:** 随机森林模型的训练时间为 0.68 秒。
- **XGBoost training time: 0.10 seconds:** XGBoost 模型的训练时间为 0.10 秒。
- **LightGBM training time: 0.04 seconds:** LightGBM 模型的训练时间为 0.04 秒, 是所有模型中最快的。
- **CatBoost training time: 1.46 seconds:** CatBoost 模型的训练时间为 1.46 秒, 是所有模型中最慢的。

## LightGBM 模型训练日志

- **[LightGBM] [Warning] Found whitespace in feature\_names, replace with underlines:** 警告信息, 提示特征名中存在空格, 建议用下划线替换。
- **[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.000055 seconds. You can set force\_col\_wise=true to remove the overhead:** 信息提示, 表示 LightGBM 自动选择了列-wise 多线程处理方式, 并提供了优化建议以减少开销。

- **[LightGBM] [Info] Total Bins 360:** 信息提示，表示 LightGBM 将特征值分成了 360 个桶 (bins)，用于加速训练。
- **[LightGBM] [Info] Number of data points in the train set: 875, number of used features: 58:** 信息提示，表示训练集中有 875 个数据点，使用了 58 个特征。
- **[LightGBM] [Info] Start training from score 5499.407537:** 信息提示，表示训练的起始分数为 5499.407537。
- **[LightGBM] [Warning] No further splits with positive gain, best gain: -inf:** 警告信息，表示在某些情况下，没有找到具有正增益的进一步分裂点。

## 总结

- **训练时间:** 展示了不同模型的训练时间，LightGBM 模型最快，CatBoost 最慢。这种信息对于模型选择和优化非常重要，尤其是在处理大规模数据时。
- **LightGBM 日志:** 提供了关于训练过程的详细信息和警告，帮助开发者了解模型训练的细节并进行优化。