

# 打造更“亲民”的数据分析工具

浅谈探索式数据分析与可视化

马珂 - DataHunter



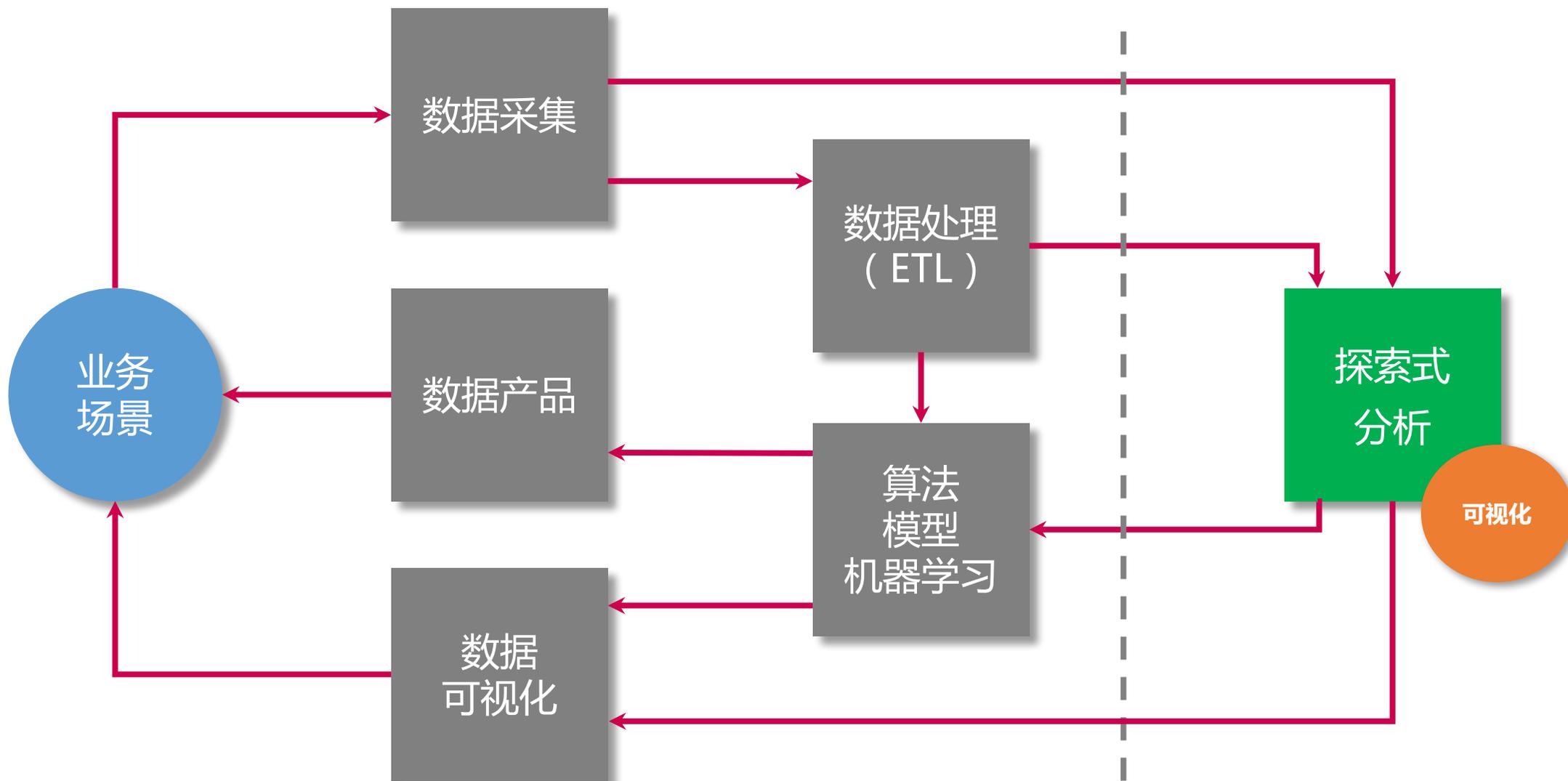


# 大数据分析



| Item | Price | Quantity | Revenue | Profit | Loss | Margin |
|------|-------|----------|---------|--------|------|--------|
| 1    | 2.00  | 50       | 100.00  | 100.00 | 0.00 | 100%   |
| 2    | 2.00  | 50       | 100.00  | 100.00 | 0.00 | 100%   |
| 3    | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |
| 4    | 2.00  | 50       | 100.00  | 100.00 | 0.00 | 100%   |
| 5    | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |
| 6    | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |
| 7    | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |
| 8    | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |
| 9    | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |
| 10   | 0.00  | 50       | 0.00    | 0.00   | 0.00 | 0%     |

# 探索式数据分析



# 数据可视化





## ● 简单对比

对分类的量化数据进行简单的对比，从而更直观的了解两者的量化对比情况，通常用来发现问题。



## ● 偏差性

通过观察数据点之间的关系，发现一些特殊的，与普通数据有明显不同的情况，用于观察数据偏离度。



## ● 时间序列

显示同一维度下数值随时间的变化，通常可以帮助人们发现趋势，进行预测。



## ● 分布情况

描述数据围绕核心数值的一个分布情况，用于观察数据的分布方式和情况。



## ● 相关性

同一维度下两个数值的变化关系对比，从而发现正相关性或负相关性，通常用于因果关系的发现。



## ● 局部与整体

用于对比部分数据和整体的关系，比如：小学生在“王者农药”玩家群里面的分布比例。



## ● 分级排序

两个以上的数值互相之间的关系，通常用于排序分级，从而查看顺序和数量。

\*\* 还有更多，七种武器仅供参考 \*\*

# 数据识别



Drop directly in the above box  
or click to upload your own images

confidence : 99% ( value : 0.99 )  
pose :  
roll(38.15) ,yaw(14.91) ,pitch(5.16)  
race : white(0.78)  
emotion : happy:100%,confused:5%  
age : 29.62 ( value : 29.62 )  
smile : true ( value : 0.97 )  
glasses : with glass ( value : 1 )  
eye\_closed : open ( value : 0.09 )  
mouth\_open\_wide : 8% ( value : 0.08 )

## ● DH 数据分析系统 (DH Data Analytics)

- 多数据源汇聚, 轻松数据建模
- 探索式, 交互式的数据分析模式



## ● DH 数据大屏幕 (DH Data Screen)

- 用于业务指挥中心, 会议室, 展厅, 展会等场景
- 整合多种业务数据, 用最合适的方式进行实时动态的可视化



## ● 大数据技术服务 (Data Service)

- 大数据系统咨询规划以及建设
- 数据清洗抽取 (ETL)

**THANKS**

