

数据分析思维的养成

秦路

1

什么是数据分析的思维

数据证明

不只是数据

业务核心

好奇心



啤酒与尿布
为什么是错的？

买了尿布的人会买啤酒，但是买啤酒的人会买尿布么？

尿布旁边应该摆放其他东西么？啤酒是否是最好选择？

怎么摆放啤酒？部分还是全部品类？

场景型的摆放是否比品类摆放更好？

数据呢？



为什么

想知道数据的来源

想知道数据的准确性

想知道数据波动的原因

想知道数据优化的方法

想知道……

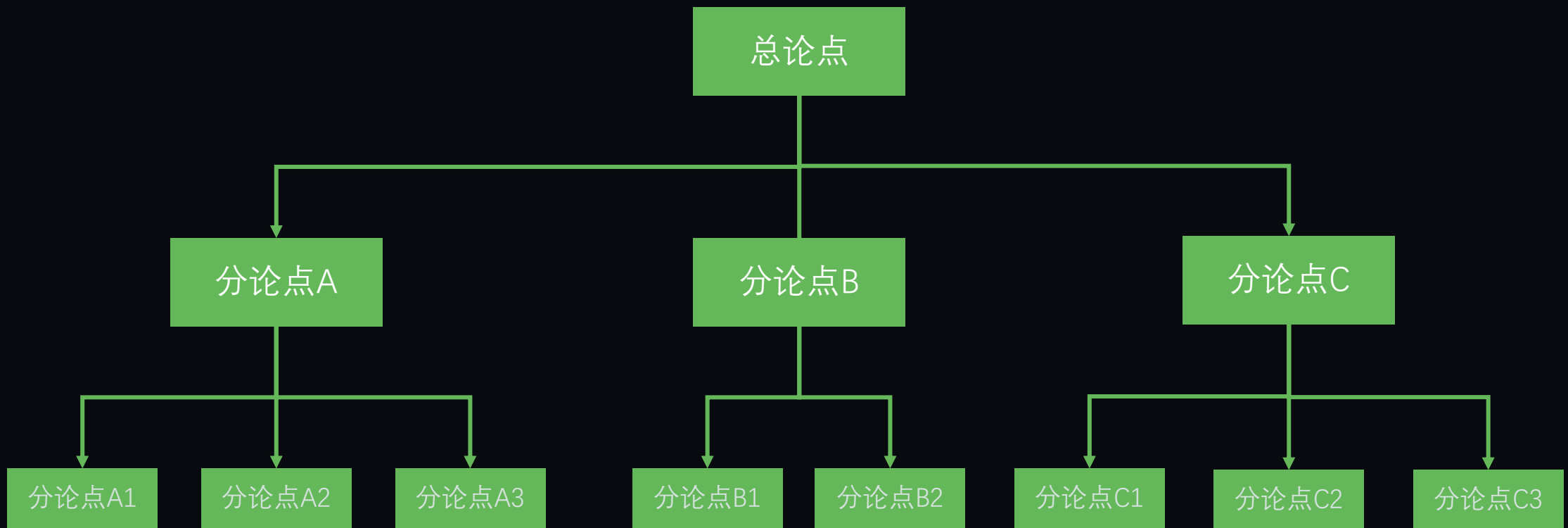
共享单车的城市投放数量多少？

2

数据分析的三种思考方式

结构化思维

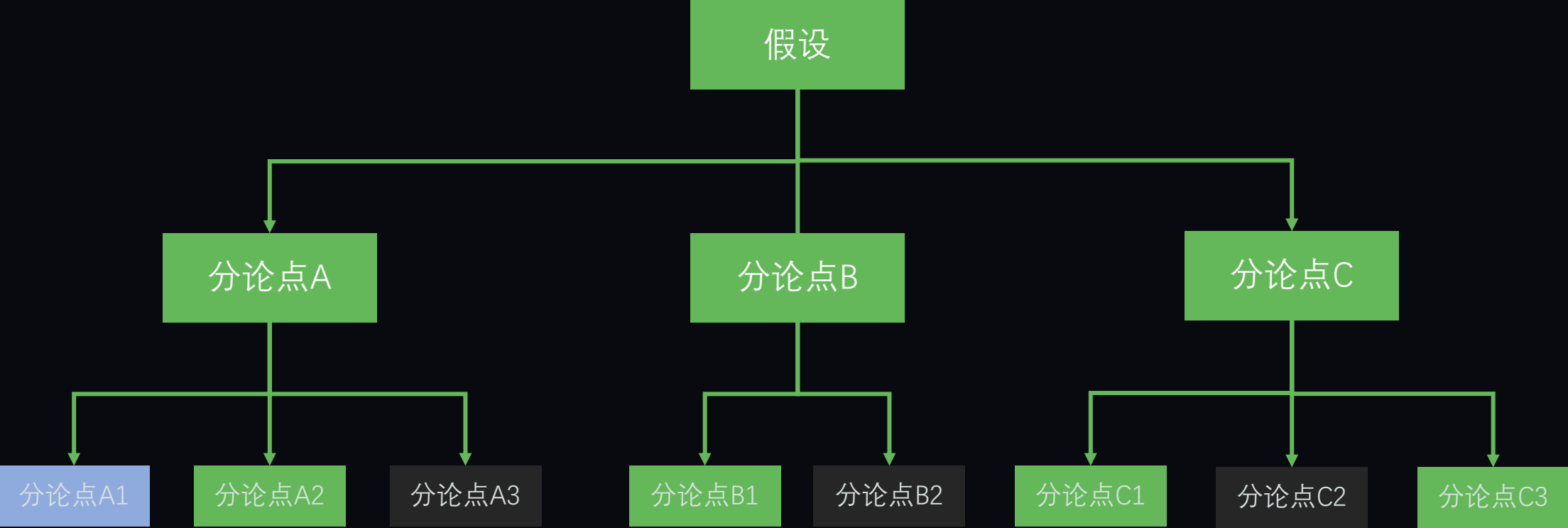
金字塔结构



MECE原则

假设原则

关键驱动原则



小问题

小问题

小问题

小问题

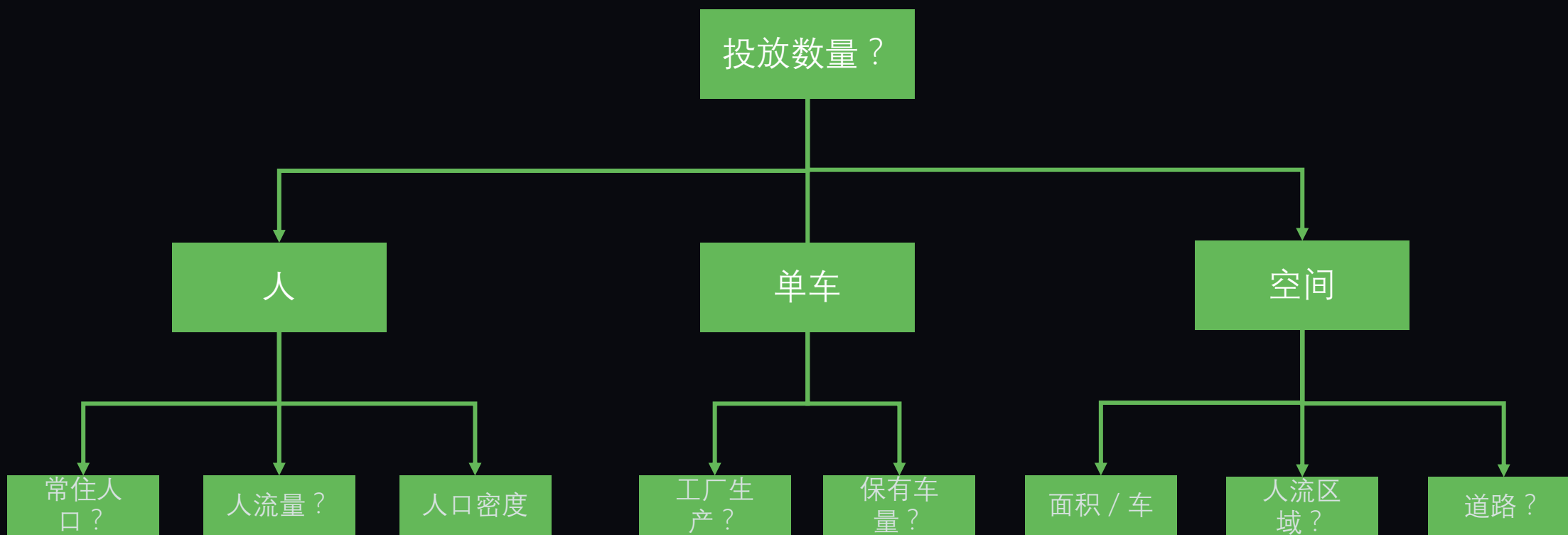
大问题

小问题

小问题

小问题

小问题



公式化思维

将结构转化成业务变量

$$\text{利润} = \text{收入} - \text{成本}$$

$$\text{收入} = \text{流量} * \text{转化率} * \text{客单价}$$

$$\text{成本} = \text{固定成本} + \text{变动成本}$$

动态化思维

数据分析应该从业务出发动态地看待数据
往前看一步，往后看一步

下载 — 注册 — 活跃 — 忠诚 — 流失

如果共享单车会报废，应该怎么修正？

3

数据分析思维的陷阱

为什么数据会欺骗你

辛普森悖论

	安卓	苹果
活跃用户数	10000	5000
付费用户数	550	200
付费转化率	5.5%	4.0%

	安卓平板	安卓手机	苹果平板	苹果手机
活跃用户数	8000	2000	1500	3500
付费用户数	500	50	100	100
转化率	6.25%	2.5%	6.67%	2.86%

三级医院死亡率10%， 二级医院死亡率5%

贝叶斯定理

如果某种疾病的发病率为千分之一。现在有一种试纸，它在患者得病的情况下，有99%的准确率判断患者得病，在患者没有得病的情况下，有5%的可能误判患者得病。现在试纸说一个患者得了病，那么患者真的得病的概率是多少？

100000

健康：99900

患病率千分之五

患者：100

测出有病：4995

无病情况下，5%的错误率

测出无病：94905

测出有病：99

有病情况下，99%的准确率

测出无病：1

真正的准确率： $99 / (99 + 4995) = 1.9\%$

颠因倒果

统计图表

