

我会按照以下五点的顺序分享我在工作中的实践，如果您有更多的细节想了解，欢迎继续联系我。

1. 数据的收集
2. 数据的分析
3. 数据的转化
4. 数据库的使用
5. 实践对应的课程基础

1. 数据的收集

用户场景	用户需求分析	我的处理	效果收益
空值 - 线下门店没有做好数据维护。如商品条码，名称，品牌，类别，是否含税，都有缺失。查询热销品牌，商品，做报表的时候，会不准	增加新产品 补全旧数据 删除重复值	建立统一产品数据库25000条用于新增。另建常用产品数据库1100条，用于查缺，和日常使用	节省时间。 不用人工重复录入
偏差 - 物流公司每天代发200双UGG, 每个包裹要人工扫码，称重，跟多个UGG公司结算。耗时两小时	不同尺码，填充物，批次会有重量偏差	每隔一周调出物流数据，建立并更新完整重量信息数据库。相同货号取更重值。 另建立收费模板	模板自动结算。减少重复称重，计算的工作
特殊值 - 礼品店售货员每天要用一个半小时分类300多张不同公司的物流运单，每次手抄一张单号，分别算付给物流的成本。耗时长易算错，店长无法核准，只能估计。	不同公司，不同线路，不同重量区间，价格相差很多	用红外扫枪五分钟录入excel. 编写模板筛选多个公司，多个线路特殊字段，按重量区间，自动算出成本	节省人力成本。每天一人一个半小时变成五分钟。 店长感激，拉近客户关系
这些是数据的预处理，清洗挖掘。从原始数据到程序处理，最后成为有效数值其中也有难点。例如上述特殊值里，会有发到新疆，台湾的包裹，面单条码和普通的包裹规律一样，无法处理。咨询店长后得知，每三百单可能只有一两单。从概率模型上来说，可以覆盖99.9%的使用场景。另外提醒了店长记得特殊处理。			

2. 数据的分析

用户场景	用户需求分析	我的处理	效果收益
拉新 - O2O 引流, Burwood驻店推广，新关注公众号并当场下单用户，即减免 5 澳元	要体现当天活动各店代购流量，用户可能取关后隔段时间来重复领取，	发放实体券。推广并统计数量 留意小程序里的历史订单记录。	算出自己负责店面的流量 算出驻店拉新的获客成本
返利 - B2B 宣传, 商家自行推广的每个新用户，每单返利 1 澳元	后台的数据订单，有大量重复数据，或已取消，未付款，未寄出的包裹	查重，筛选干扰值 根据用户ID，确定为新用户	准确返利。 算出返利商家的获客成本
高价值用户 - B2B渠道分销, 物流公司代理6个 UGG，B 端用户粘性和购买的品牌不同，厂家补贴不同	Cross Sell 推广更多品牌和购买比例；发掘高价值用户；获得更多 UGG 厂家补贴	建立客户信息数据库，建立订单数据库，加权处理品牌价值，给活跃用户赋值	高资源回报率。高价值用户的满意度提高。厂家补贴边际利益增加
高价值商品和品牌 - 同一个品牌，在不同城市的热销度不同。 同一个城市，品牌的销售占比不同，商品利润不同	找出热销商品分配对分销渠道的广告提醒	根据物流数据，算出一段时期内每个品牌的销量，每个城市的购买力以及品牌的影响力和占比	增加利益，减少成本。合理分配广告资源
DDoS 攻击 - 小程序和网站的服务器试过被恶意攻击	容灾，备份，备用服务器	用了双服务器，也做了容灾备份，配置 wpDefence 及防火墙黑名单	数据安全，完整性
数据的分析中，我一直很注意Statistics Bias, 例如“ 拉新 ”里有数据来源的偏差，代购本身带有对低价格的偏向性，一旦没有减免，就发现这个小程序的用户留存率还是低的。即新增客户数不一定是真实用户数，我也咨询过店家的意见，提议后来做 B2B 的宣传，不做 O2O 的引流。“ 高价值商品 ”里有数据样本的偏差，即样本量不够大，无法代表所有城市。也可能有厂家的优惠活动，造成购买的偏向性。			

3. 数据的转化

用户场景	用户需求分析	我的处理	效果收益
网站使用率 - 监控网站登入数据, 用户留存率, 页面点击率, 用户组成, 消费习惯	自动监控 图象化数据	配置Google Analytics, Jetpack, 各种数据的离散图, 城市分布图,	直观体现用户数据
商品SEO - 小程序内搜索的商品, 匹配度低, 商品名称不规律	提高搜索精准度	比对统一产品库 编排名称, 使用淘宝爬虫工具优化搜索排名	优化用户体验
网页SEO - 检测网站的搜索排名, 内容易读性	提高搜索排名	配置 Yoast SEO 等排名工具, 增加搜索关键词	优化可读性, 更直接的可购买广告露出
图像化 - 数据图像化	可修改, 可交互的图像布局	使用Google charts, High Chart, Char.js等工具连接 SQL	图像自动根据数据转换
数据的转换, 感觉Google Analytics是最好用最清晰, 另外使用过微信公众号的后台, 淘宝的卖家后台, 物流的 ERP. 云服务的图像化后台使用过Yealink 的物联网, WordPress 的 DashBoard, API工厂等。			

4. 数据库的使用

用户场景	工具	示例
本地数据库的搭建	MySQL, Navicat, Linux MAMP(Macintosh, Apache, MySQL, PHP), Excel	我的本地开发环境
服务器数据库的搭建 双服务器容灾 数据自动备份到 Google Drive 和本地备份	服务器的管理工具 c-Panel, Linux, MySQL 索引的CRUD, Trigger ARI Adminer, phpMyAdmin FileZila	我的个人主页 senlin.com.au
云端服务器的使用	Python + api JavaScript + api AWS 云	IoT物联网 微信小程序 GitHub

5. 实践对应的课程基础

对应实践	描述	课程
1数据的收集 2数据的分析	数据库的理论基础 SQL语句的实践 数据的筛选	Introduction to Information Systems Introduction to Data Analysis Database Fundamentals
2数据的分析 3数据的转化 4数据库的使用	程序对数据的处理 SQL数据库的链接 AWS的云数据部署	Advanced Internet Programming Programming on the Internet Applications Programming Programming Fundamentals
4数据库的使用	数据的分布式存储 服务器的容灾备份 数据的机密性完整性 可用性，用户组策略	Network Design Network Servers Fundamentals of Security Network Security Networking Essentials Routing and Internetworks
2数据的分析 3数据的转化	网络数据及 应用程序的数据处理	Network Application Development in the iOS Environment Mobile Networking Web Systems Web Services Development Advanced Interaction Design
1数据的收集	用户需求的分析 企业的项目需求分析	Project Management and the Professional E-Commerce Business Requirements Modeling Communication for IT Professionals Collaborative Business Processes

以上就是我的分享。自知还有很多需要学习和继续实践的方面。

如果您有更多的细节想了解，欢迎继续联系我。

Kind Regards,

Sen

senlin.com.au

senlin1003@gmail.com

04 8202 6582