

全流程讲透数据分析和指标体系

全流程讲透数据分析和指标体系.....	1
一、为什么需要数据分析.....	2
1. 量化 IT 投资成效，以数据驱动决策.....	2
2. 通过数据分析验证产品成效	3
3. 通过数据分析洞察用户.....	6
4. 通过数据分析找到机会点	7
5. 结语.....	8
二、数据分析的框架	8
1. 数据生成	11
2. 获取数据	12
3. 数据建模	12
4. 构建数据模型	13
5. 设计维度和度量指标.....	13
6. 数据分析	14
7. 数据应用	15
8. 结语.....	16
三、数据分析 数据指标体系设计	16
1. 用户数据之 AARRR 模型.....	16
2. 行为数据之 UJM+OSM 模型	17
3. 业务数据之指标分层.....	18
4. 结语.....	19
四、数据呈现之数据分析方法	20
1. 漏斗分析	20
2. 多维拆解	21
3. 趋势分析	22
4. 对比分析	23
5. 帕累托分析	24
6. 交叉分析	25
五、结语.....	26

本文分四部分内容：

- 一、为什么需要数据分析
- 二、数据分析的框架
- 三、数据指标体系搭建
- 四、数据呈现之数据分析方法

一、为什么需要数据分析

数据分析的重要性不言而喻，没有数据，就是感性呀。

你说你做完这个功能可以获得什么价值回报，没有数据支撑，太过苍白无力，没人信服你。但是只要你说，你做完这个项目，可以提升订单转化率多少个点，并且信誓旦旦讲，那没有人不会不理你。毕竟，这是赤裸裸的钱，谁不爱钱呢。

数据不会被观点打败，数据只能被数据打败。人家拿数据得出的结论，跟你的直觉再对不上，你心里再不服气，想反驳，也必须拿数据说话。

我们现在妥妥地已经进入了数据时代。

互联网公司，无一例外地强调自己的数据驱动决策；传统企业，现在最重要的战略就是数字化转型。

相信今天你能感觉得到，数据在我们的工作和生活中，已经成了空气和水一样的存在。

数据思维或者数据分析能力已经成为这个时代的必须，不是可选。

那么数据分析为何如此重要呢，我从以下 4 点来阐述。这些场景也是日常发生在我实际工作中的，我们用数据每天来做各种分析、洞察或者决策。

1. 量化 IT 投资成效，以数据驱动决策

无论哪家公司开发资源永远不够，每个产品经理都想争取开发资源做自己的项目，说自己项目更有价值。如何评估优先级，数据来说话。

站在公司或者决策者角度，数据最本质的作用，是作为资源调配的裁判，帮我们用最客观的方式将资源投到最有价值的事情上。

目前我所在公司运转方式为：年初制定年度 KPI，如转化率或者用户满意度，那么接下来一年内所有产品经理运行的项目将按照这个指标来。

产品经理 A 和产品经理 B 同时启动了自己的项目，那么 IT 资源投到谁的项目上，A 和 B 需要去做项目 ROI 的论述。谁的 ROI 更高，业务总负责人就同意将资源调配到哪个项目上，这样 A 和 B 都没什么话说，IT 资源也将被用在刀刃上，减少资源浪费。

产品经理在论述项目价值时，采用自下而上的论证方式。比如说这个项目要完成哪些功能，功能 1、功能 2、功能 3……每个功能带可以带来多大的价值，如转化率可以提升多少，这些都来自数据。

有时候我们以为的很大项目价值经这种层层灵魂拷问，最后验证预估价值不大，在早期就能减少资源浪费，而不是真的要等到功能上线才发现对业务价值帮助不大，那时候开发资源已被浪费。

如果你一直抱着项目上线后看真实数据反馈，除非你家开金矿的，可以容忍不断做尝试创新类实验。运气好成功了，大家都开心；运气不好，大家失败了，不怕，反正家里开金矿，大不了重头再来。但是，我想这应该是个伪命题。

今日头条系产品为何能快速发展？它们的产品方法论是什么？

今日头条是一家流量运作公司，对流量 ROI 的运用纯熟度与效率非常高。依赖强大的数据和算法系统，通过 AB test 同时运行几千条功能测试，用最短的时间去找出最有价值的业务方向，得到结论后，快速推出市场，抢占先机。

中国互联网已经进入到一个获取流量成本很高的时代，企业并没有那么多容错机会给到大家不断尝试试错。要提高决策准确性，主要依赖数据论证。

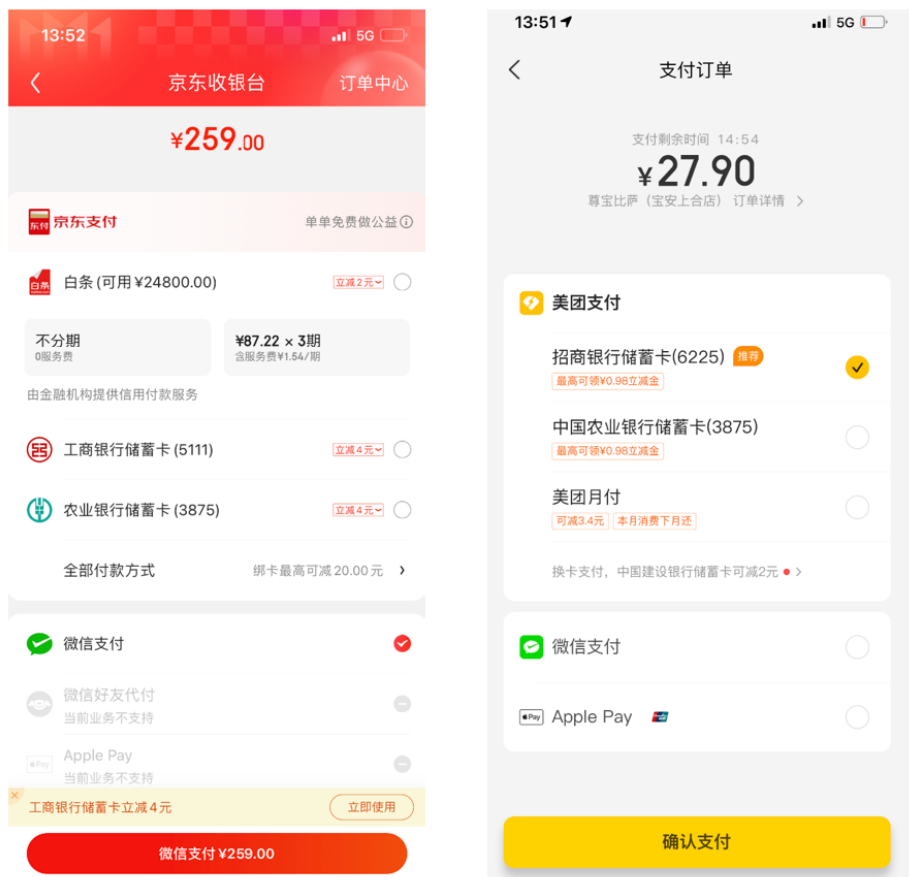
2. 通过数据分析验证产品成效

你在面试时，你说你牛逼。如何叫人信服你？你说你牛逼，你就真牛逼？

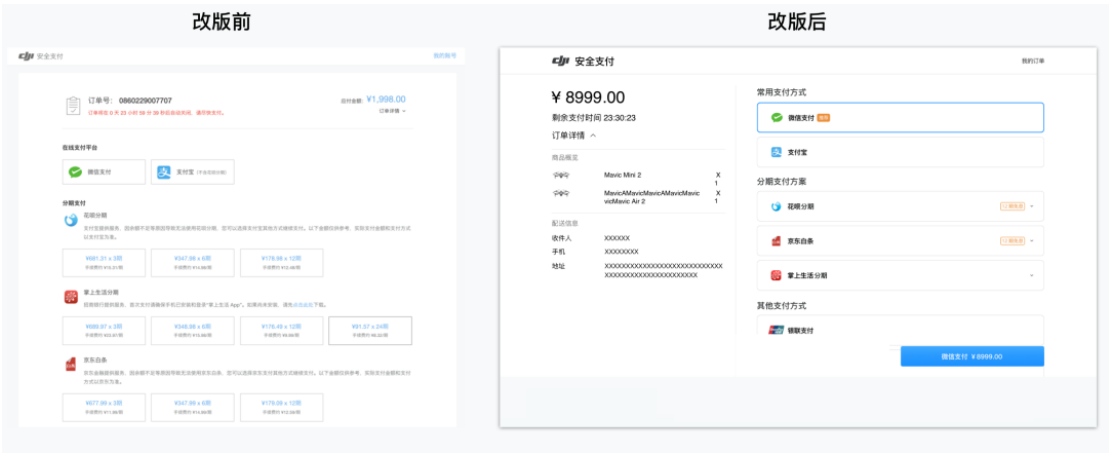
相信大家已经习惯了小套路，说自己曾经做的某个项目转化率提高了多少个点之类的。

实际上，在我们做的每个项目或者功能中，都应该用数据来验证产品成效。

电商网站的收银台大家耳熟能详，比如以下：



参照国内竞品调研和产品交互设计师自以为的产品目标：给用户传递安全感——通过增加底部“确认支付”按钮来实现。



实际上线效果却不尽如意，新版本比老版本跌了 3 个点左右。

事后我们反思了整个过程，得到一些启发：

支付需要用户冲动型消费，不需要用户那么理性思考。按钮操作，让用户有种仪式感，反而增加了用户犹豫心理。

比如淘宝换成了指纹支付，或者刷脸支付，转化率会提高。我作为用户，使用了淘宝指纹支付后，下单支付快了很多。我也有同事，一不小心刷脸支付后，就懒得再发起退款了~

支付产品的核心是：安全和快捷。如果品牌背书，用户已经觉得安全，那么接下来就是快！去掉按钮才是让用户有快速支付的感觉。

于是，我们又花了一些时间，去掉底部“确认支付”按钮，用户选择支付方式就可以进到支付环节，转化率竟然又提升了！



去掉了底部“确认支付”按钮

整个团队从这个案例中都得到较大感触，充分体现了我们自以为更好的方案对用户或者从业务角度并不是最好的。如果没有数据来验证，我们还会一直停留在自己觉得设计很好的功能中自嗨！

用 AB test 测功能，用数据来验证功能成效是最有说服力的手段。

3. 通过数据分析洞察用户

用户研究是产品经理必须要去做的一件事，懂用户，挖痛点，给方案。

用户研究常用的方法除了用户访谈、调查问卷等定性研究外。从已存数据中发现用户的行为偏好，建立数据与用户画像之间的关联，针对不同人群需求或者痛点给出合理的产品解决方案也是数据驱动决策的手段。

比如从历史订单数据中，挖掘新老用户购买商品的偏好，可以针对新老用户群体做个性化的商品推荐，这也是大家能够感知到的电商产品推荐。你去逛淘宝首页，是不是会觉得淘宝比自己还懂自己？给自己推荐的大部分商品是自己喜欢的。

通常对用户历史行为数据收集越详细，越是能够了解用户，为用户做出合理的产品设计。

为什么人们总说作为电商平台和支付平台的阿里有最完整的行为数据乃至最完整的人群画像，那是因为根据生活消费的商品和服务，几乎能够推断一个人全部的特征，而越是习惯网上购物的用户，衣食住行都用同一个支付手段的用户，就越能够被电商和支付平台完整描述。

产品在做产品功能设计时，并不只是单纯为用户提供功能和服务就可以，理所当然以为用户会来使用。定要对用户需求或者痛点挖掘地足够深，才能精准提供服务。从过去已存数据中挖掘用户行为偏好或者痛点，是产品设计的第一步。

这里借用《产品思维》一书中提到的例子。

我们是一个创业小团队，正在做一个 P2P（个人对个人）金融产品。平台上已经有了一定量的用户，他们购买了我们的理财产品。我们正在考虑要不要增加 VIP（贵宾）套餐服务，定位高价值用户，定向提供理财顾问服务。

我们会请许多专家，提供很多额外的分析工具，让这些用户享受高端服务，赢利方式就是 VIP 年费。单拿这个服务来看，肯定是没有问题的。

“许多类似的产品都有这样的服务。”也许老板就会这么跟你说。但这不能成为我们就一定要提供 VIP 套餐服务的理由，我们还是要看看我们的用户是什么样的。

可以先统计下当前用户购买理财产品的行为数据，看看大部分买的额度有多大。其实，额度背后代表的是这些用户的收入水平和对理财的态度。

比如粗暴一点假设，我们看用户的行为，过去 99%的用户都是奔着投资 5 万元，锁定期 6 个月，年化收益率 5%的产品去的。

这些行为代表什么呢？从生活经验判断，对年化收益率要求不高，对流动性也要求不高，对安全性要求比较高的人，应该都是普通的上班族。这些人在理财方面非常保守，而且还比较年轻，理财额度并不高。

具体的验证可以通过访谈和调研来完成。假如结果显示，这些用户群体基本都是毕业三年内的职场新人，理财行为非常保守，而且还比较年轻，理财额度并不高。

这时回过头来看，这个 VIP 套餐服务的吸引力就特别有限了。仅从用户特征来看，几乎就可以给这个创意判死刑了。

从已存数据中分析用户需求或者痛点，找到用户行为偏好，精细化用户群体，了解用户，才能判断业务模式的可行性。而不是理所当然觉得用户会使用这个功能。

4. 通过数据分析找到机会点

在刚做支付业务时，首先就对 Top10、各端做了支付成功率分析，很快就发现某些国家的转化率是低于其他国家的，自此这些国家被列为重点和困难国家。我们的机会点也是优先提高这些国家的支付转化率。

再比如做商城购物流程优化项目时，首先拉取了商详页-购物车-结算页-收银台-支付成功 发现商详页和结算页转化率两个环节在整个路径中转化率最低，因此马上定出了事项优先级，优先解决这两个页面的转化问题。

职场中如何价值最大化？可不就是发现最严重的问题，找到最大的机会点，把资源用在刀刃上。

5. 结语

管理大师德鲁克说：“不能衡量，就无法管理。”

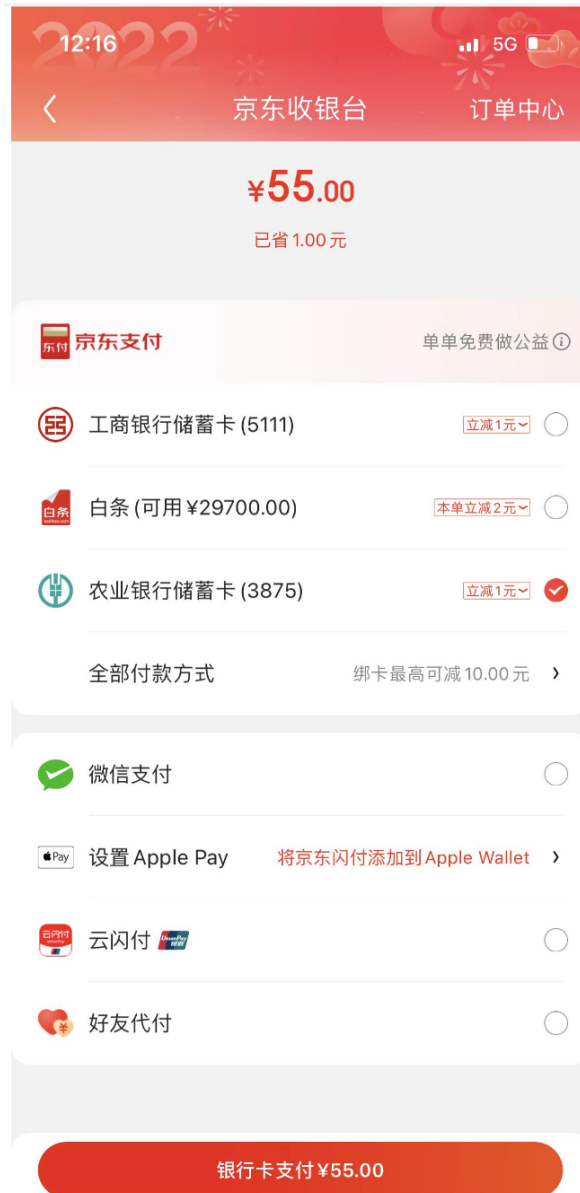
产品经理完拍脑袋、凭感觉、凭经验做决策的时代已经过去了。如果你还没有数据思维或者数据分析相关的能力，被时代淘汰真的是，早晚的事！

产品经理不需要成为数据分析方面的专家，但什么时候分析数据、分析哪些数据、如何分析数据、如何用数据辅助决策、如何用数据驱动业务，这些问题是产品经理必须要回答的。

二、数据分析的框架

我以支付业务为例来讲解。

用户来到支付收银台后，在页面上有很多点击行为，比如选择各种支付方式，微信支付、ApplePay 支付等最后完成支付，也有可能点击左上角返回键或者右上角订单中心离开当前页面。



这个过程会产生很多数据，从数据大类上分成：用户数据、行为数据和业务数据。

谁（用户数据）做了什么（行为数据）结果如何（业务数据）？



用户数据指用户本身的特性，如用户画像，使用你产品的用户男性多还是女性多，年龄多大等。

行为数据指用户使用产品在页面上的各种点击行为，在页面上停留时长等。

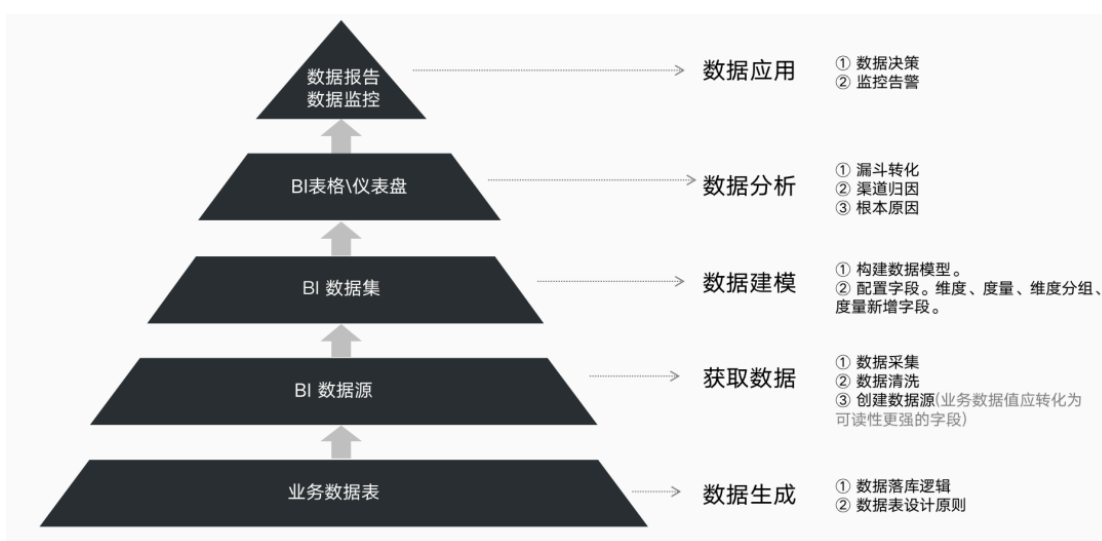
业务数据指用户行为之后，实际产生的结果，业务数据会落库业务数据表。分析业务数据的意义，可以衡量商业价值，是业务最终呈现结果，用以推动公司业务的发展。

用户数据和行为数据通常可以从第三方数据工具，如友盟、Google Analytics 直接获取，业务数据一般要内部建设。



今天重点讲业务数据搭建完整过程，以阿里云的 Quick BI 为例。

在整个数据分析的框架中，分为五大层次，依次是：数据生成、获取数据、数据建模、数据分析和数据应用。



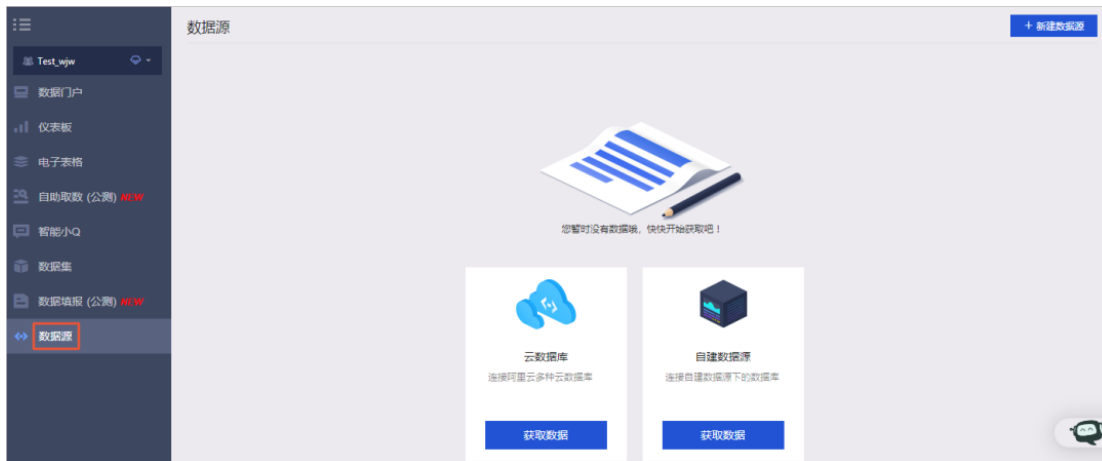
1. 数据生成

还是以支付业务为例，用户选择支付方式完成支付后，落库核心的两张业务表：订单表和交易表。一个订单会对应多笔交易（每选择一种支付方式生成一笔交易，一笔订单可以使用多个支付方式尝试支付），其实还会产生其他表，比如收货地址表等。

订单表					
order id	shopper email	device	status	created at	
000001	xxxxxx@gmail.com	pc	success	20210312 23:12:30	
000002	xxxxxx@gmail.com	h5	pending	20210312 23:12:30	
000003	xxxxxx@gmail.com	app	pending	20210312 23:12:30	
交易表					
uuid	order_id	payment_method	status	device	created_at
111	000001	weixin	success	pc	20210312 23:12:30
222	000001	alipay	pending	pc	20210312 23:12:30
333	000002	weixin	pending	h5	20210312 23:12:30
444	000003	weixin	pending	app	20210312 23:12:30

2. 获取数据

通常使用第三方工具如 ETL 将业务系统的数据经过抽取(Extract)、清洗转换 (Transform) 之后加载 (Load) 到数据仓库的过程，数据呈现在 BI 的数据源。



3. 数据建模

所有数据进到数仓以后，需要根据实际想要看的业务数据进行数据建模，建模后的数据呈现在数据集。数据集作为数据源和可视化展示的中间环节，承接数据源的输入，并为可视化展示输出数据表。



4. 构建数据模型

数据建模是什么含义呢？

底层的业务数据表其实很多，几十张上百张都有，但到了业务数据分析阶段，当需要分析的数据存储在不同的表，可以通过数据关联，把多个表连接起来，形成模型进行数据分析。

比如上述的业务底层订单表到了数据分析阶段衍生的订单表字段发生变化，name 和 city 是从业务地址表取来的数据。

总的来说，数据模型是完全面向数据分析的业务场景形成的新表。以支付业务为例，我构建的数据模型有：用户表、订单表和交易表。

业务订单表						
order_id	shopper_email	device	status	created_at		
000001	xxxxxx@gmail.com	pc	success	20210312 23:12:30		
000002	xxxxxx@gmail.com	h5	pending	20210312 23:12:30		
000003	xxxxxx@gmail.com	app	pending	20210312 23:12:30		
BI 订单表						
order_id	shopper_email	device	status	name	city	created_at
000001	xxxxxx@gmail.com	pc	success	zhangsan	shenzhen	20210312 23:12:30
000002	xxxxxx@gmail.com	h5	pending	zhangsan	shenzhen	20210312 23:12:30
000003	xxxxxx@gmail.com	app	pending	zhangsan	shenzhen	20210312 23:12:30

5. 设计维度和度量指标

对数据字段可以进行下一步分类：

- 维度（Dimensions）
- 度量（Measures）

在统计学中，单一数据字段可以被分为离散和连续。离散通常是维度，比如城市名称、用户名字，特征是有限数量的值；连续通常是度量，比如销量、利润或成功率，特征是不可罗列，可能为任一数值。维度和度量中有许多灰色区域，比如金额，可以做维度，也可以做度量。

在上述订单表中，device、city 等是维度，对 order_id 计数的总订单数、对 status = success 计数的成功订单数是度量。

维度							度量		
order_id	shopper_email	device	status	name	city	created_at	总订单数	成功订单数	订单成功率
000001	xxxxxx@gmail.com	pc	success	zhangsan	shenzhen	20210312 23:12:30	1	1	1
000002	xxxxxx@gmail.com	h5	pending	zhangsan	shenzhen	20210312 23:12:30	1	0	0
000003	xxxxxx@gmail.com	app	pending	zhangsan	shenzhen	20210312 23:12:30	1	0	0

度量可以再分原子度量和派生度量。

原子度量指从维度里直接获取到，上表中的总订单数和成功订单数。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/995240030230011044>