

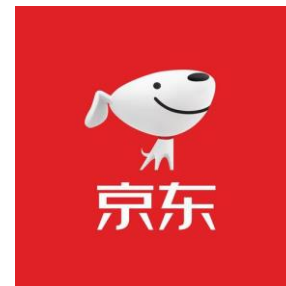


京东

京东消费者数据分析

李怡凡

分析思路



目标

作为国内知名的电商平台，京东多年来在保持高速发展的同时，沉淀了数亿的忠实用户，积累了海量的真实数据。我们希望通过商家，商品和用户的数据分析，找到影响消费的关键因素，提升订单量

问题

What: 用户的购买趋势是怎样的？

Why: 影响用户购买行为的因素都有哪些？

How: 京东的业务部门可以怎么做，提升订单量？

数据收集



Field	Type	Definition
customer_id	Bigint	客户编码
age_range	Int	年龄分段
gender	String	性别
customer_register_date	Date	客户注册日期
customer_level	Int	会员级别
city_level	Int	会员城市级别
shop_id	Bigint	店铺编码
product_id	Bigint	产品编码
brand	String	产品品牌
category	String	产品类别
product_market_date	Date	产品上市日期
vender_id	Bigint	商家编码
fans_number	Int	粉丝数
vip_number	Int	会员数
shop_register_date	Date	开店时间
shop_category	String	店铺主营
shop_score	Float	店铺打分
action_date	Date	行为时间
action_id	Bigint	行为编码
type	String	行为类别

数据观察



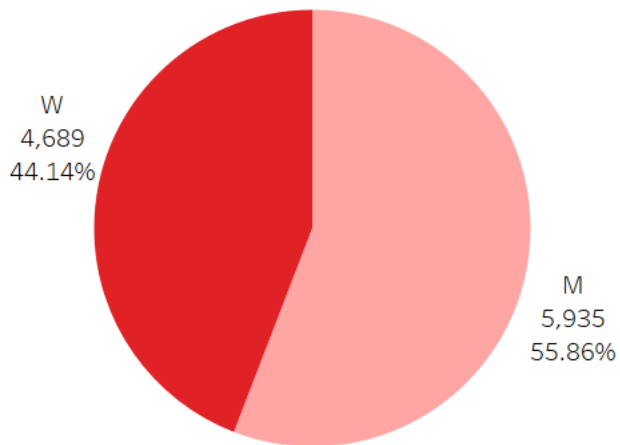
- 收集到的数据有三大维度: 用户数据, 用户购买行为, 产品/店铺数据。
- 数据大小: 183,829行
- 空值/缺失值: age range空值 (83行), gender – Unknown值 (473行), 城市会员级别 - 空值 (125行), shop_register_date大量空值 (71354行)

消费者分析

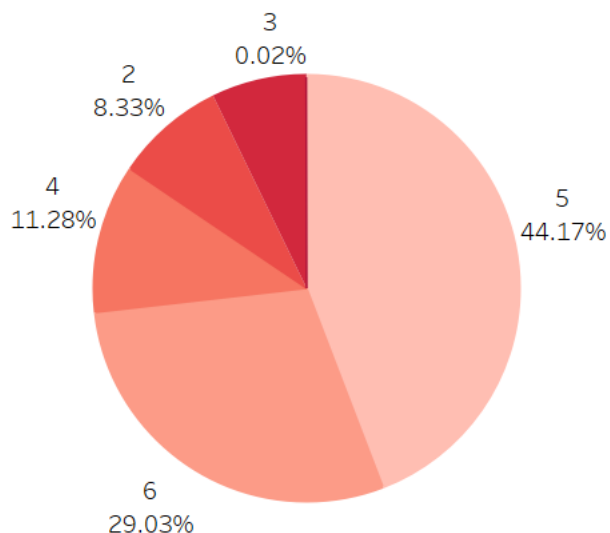


- 55.85%的消费者为男性，44.14%的消费者为女性
- 超过半数的消费者的年龄在第5组和第6组
- 消费者较为平均的分布于第4，3，5，1组城市

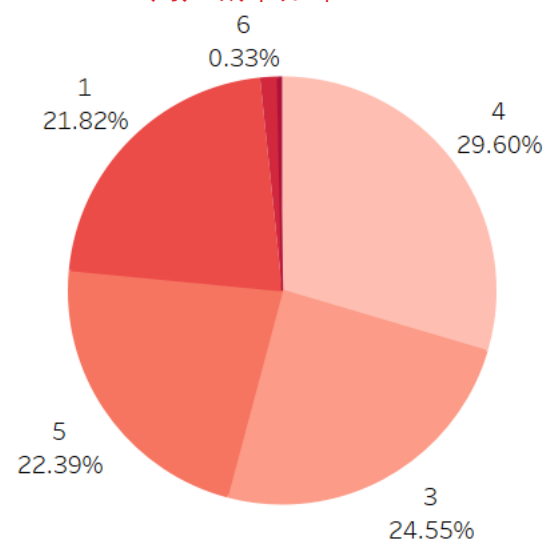
用户性别分布



用户年龄分布



用户城市分布

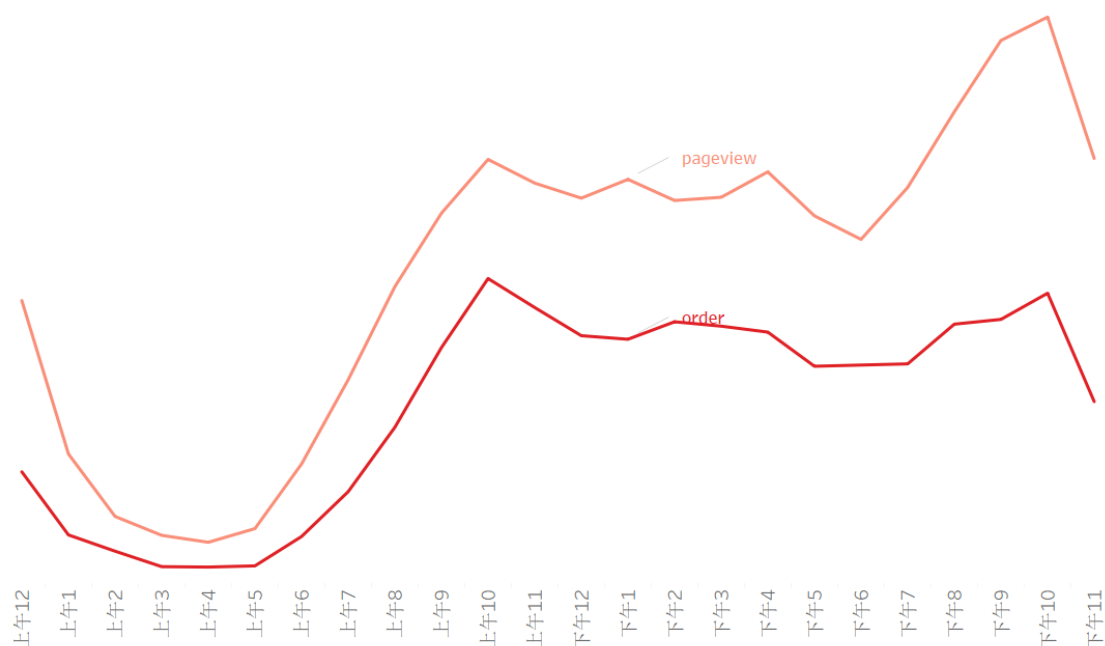


用户购买趋势

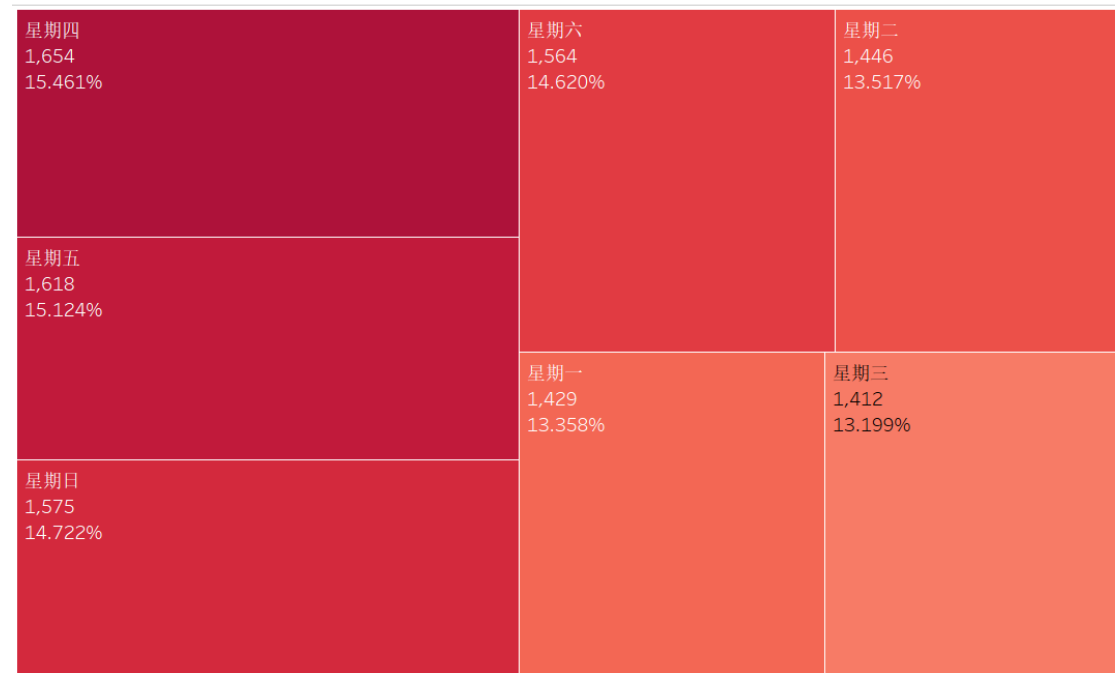


- 白天为用户高度活跃时间，晚8点至10点为页面浏览与消费的高峰时间段
- 周四，周五下单量最高，其次是周日与周六，周一，二，三订单量下降

用户下单趋势（小时）



用户下单分布（一周）



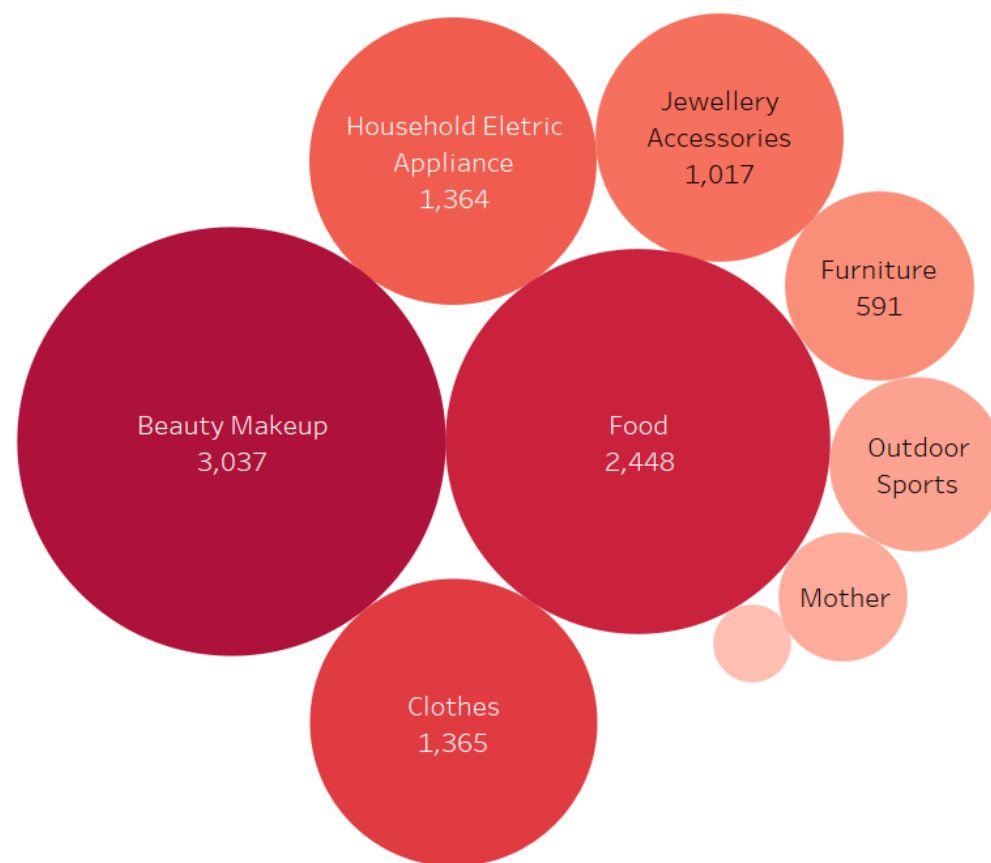
店铺种类订单量分布



- 订单量最高的店铺种类为：

- 美容护肤产品
- 食物
- 衣物
- 家用电子

店铺种类分布



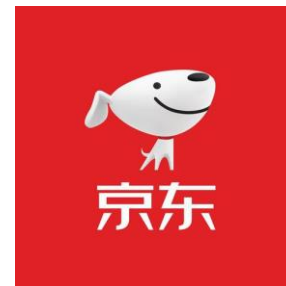
产品种类订单量分布



- 订单量最高的产品种类为：外套，茶叶，项链，脸部精华，面霜



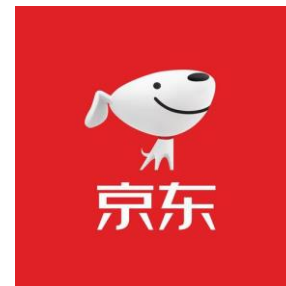
品牌订单量分布



- 订单量最高的品牌有：Lipton, Innisfree, LukFook, Li-Ning, Midea

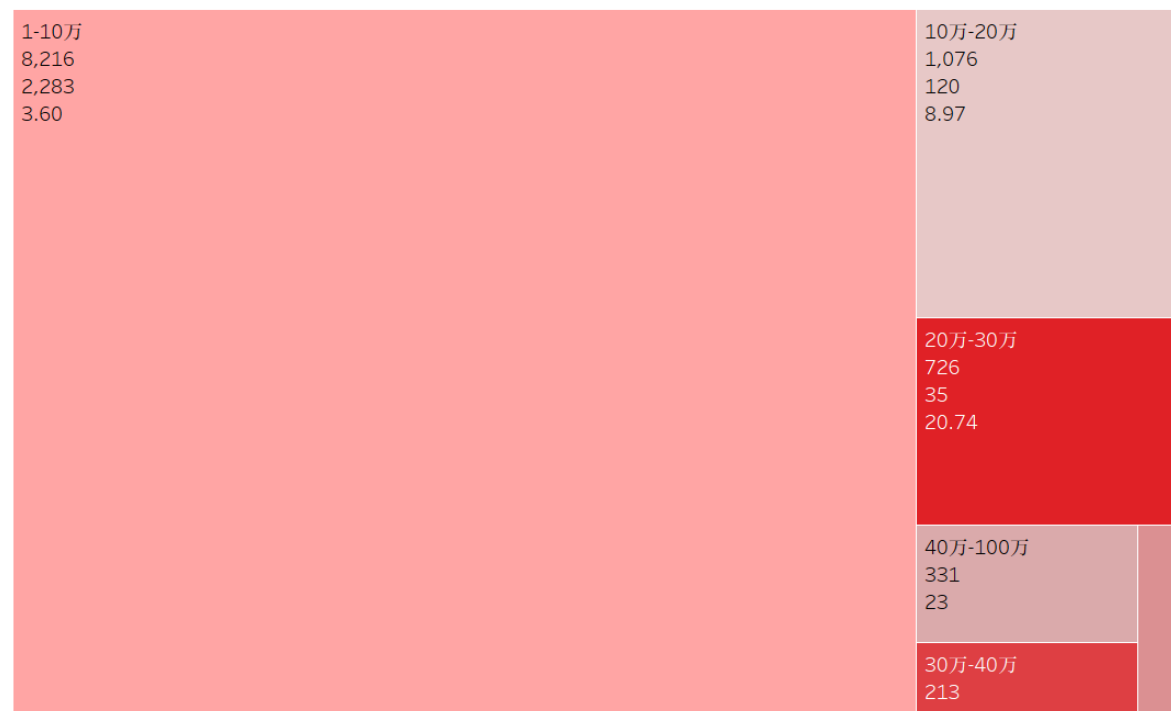


店铺粉丝数与订单量

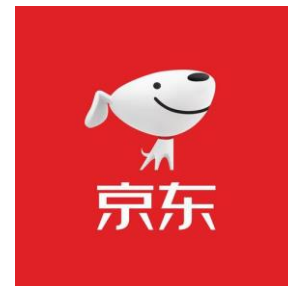


- 77%的订单来自于拥有1-10万粉丝的店铺，平均每家店的订单量为3.6
- 粉丝数20万-30万的店铺平均订单量最高，为20.74单/店铺

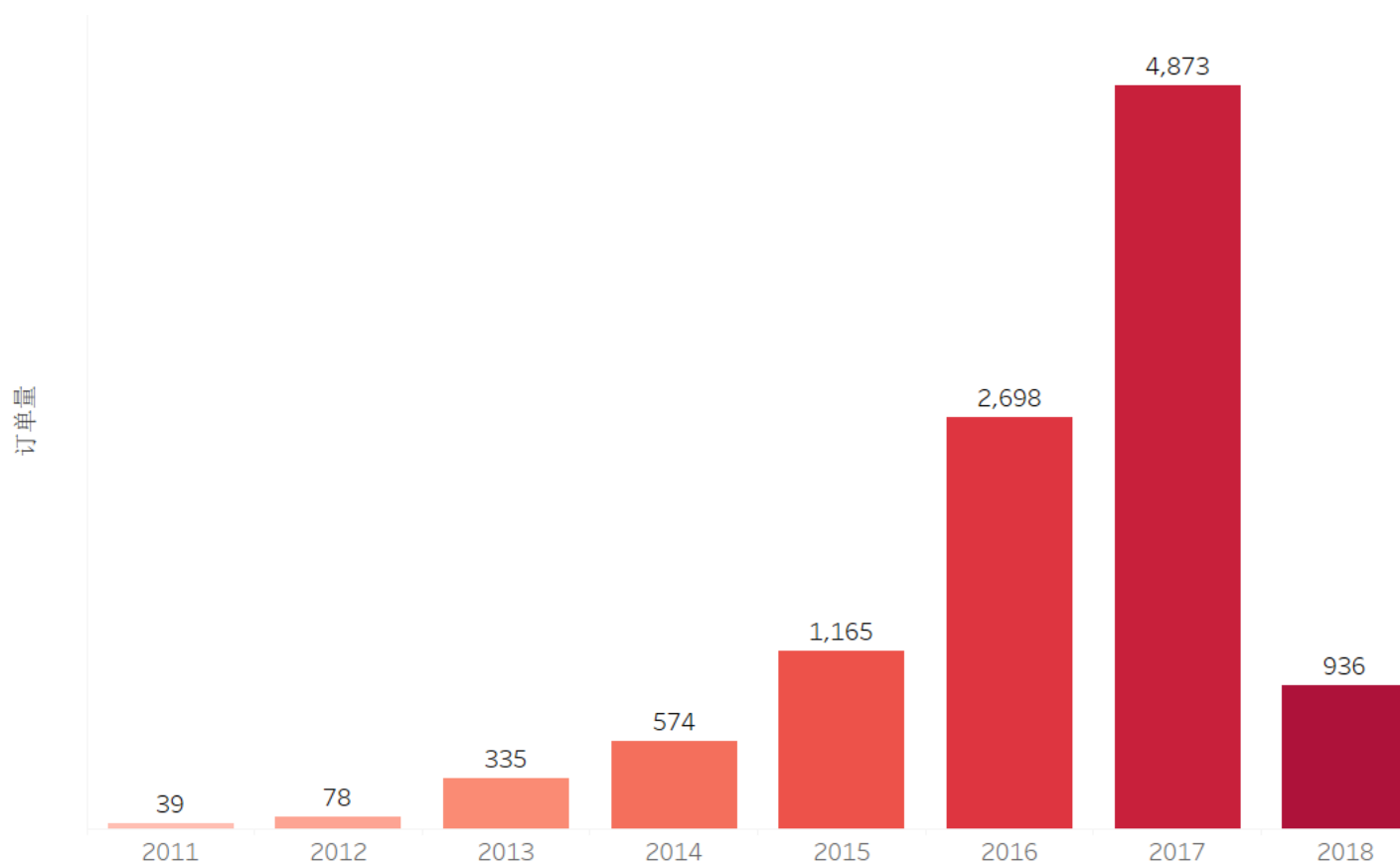
店铺粉丝数与订单量



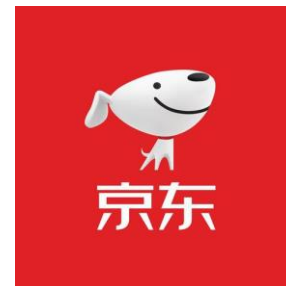
产品上市时间与订单量



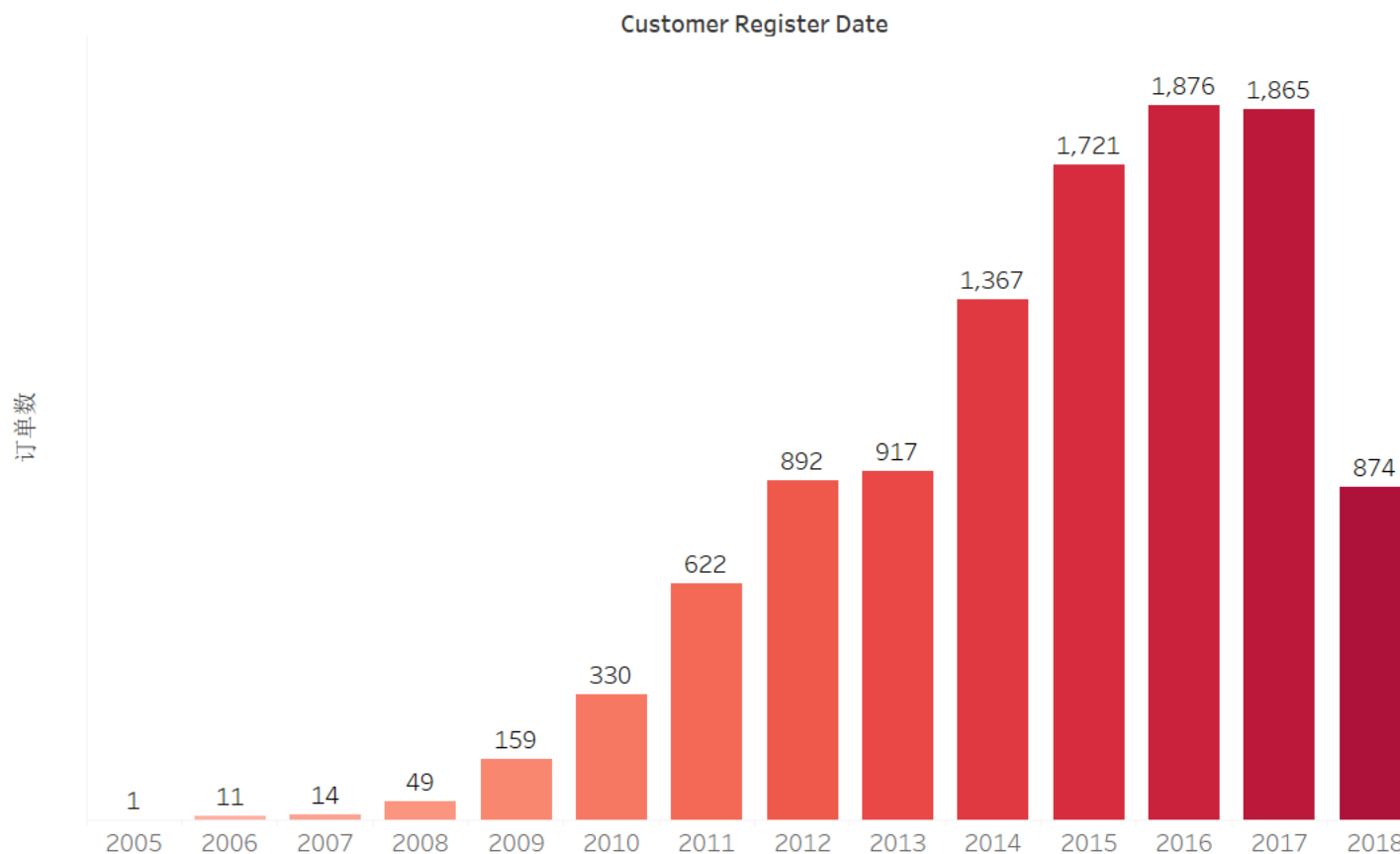
- 2017年上市的产品订单量最高，上市时间越早的产品订单量越低



消费者注册时间与订单量



- 注册时间在最近5年内的消费者订单量最高，注册时间越久，订单量越低

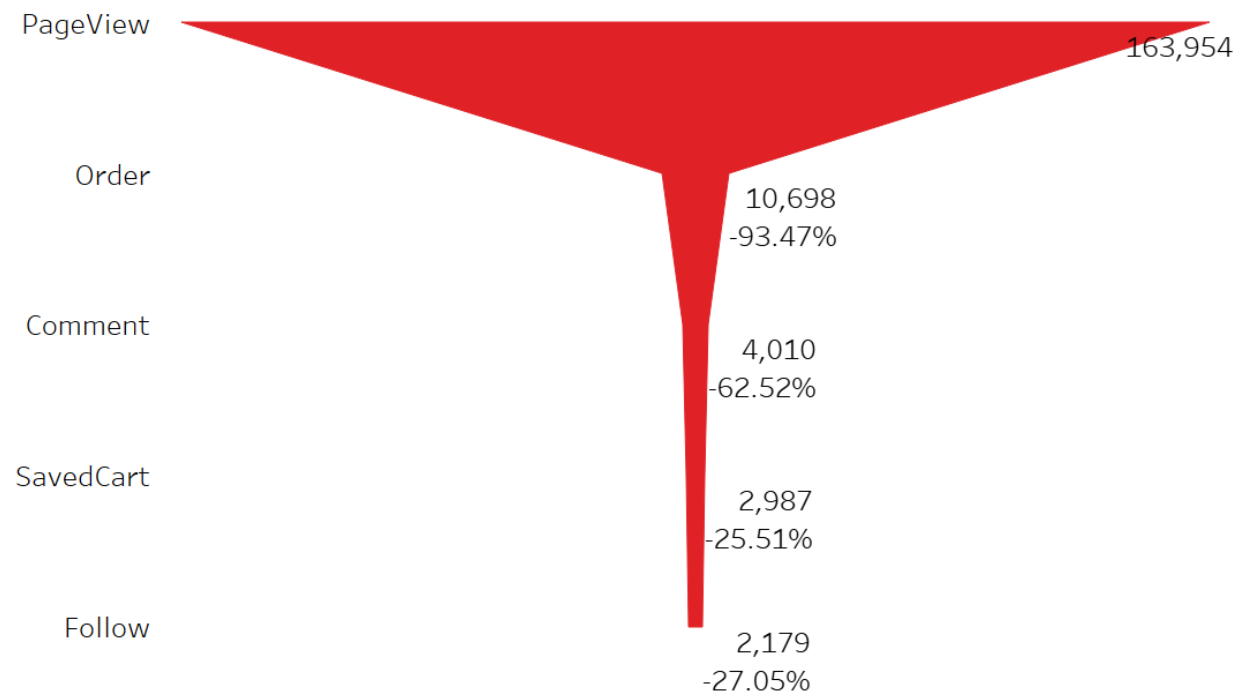


消费行为漏斗图



- 页面浏览至下单的转化率为6.5%
- 大部分消费者会直接在商品页面下单，极小部分消费者会使用购物车

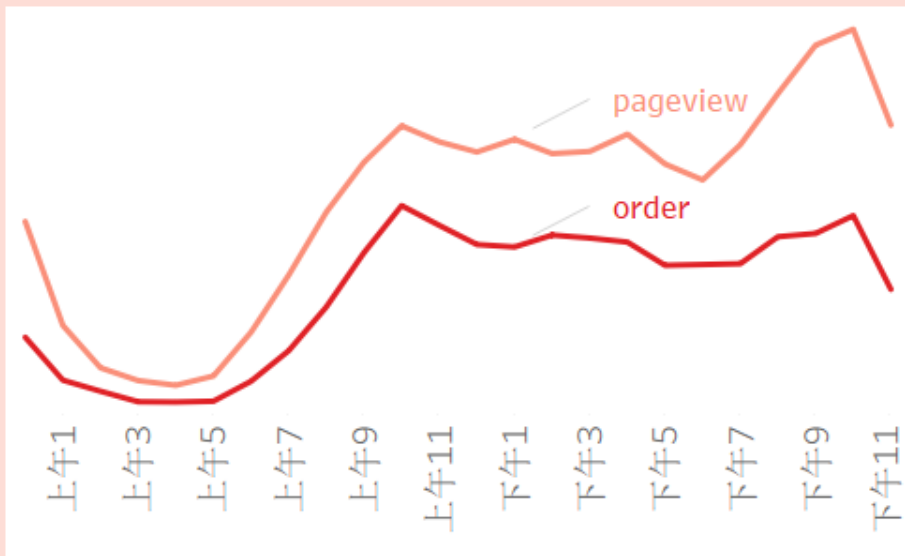
消费行为漏斗图



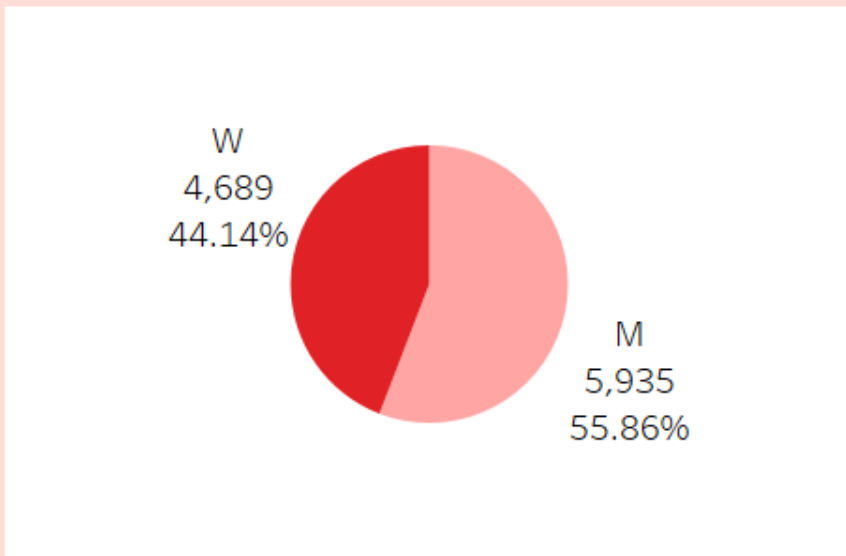


京东消费者行为分析

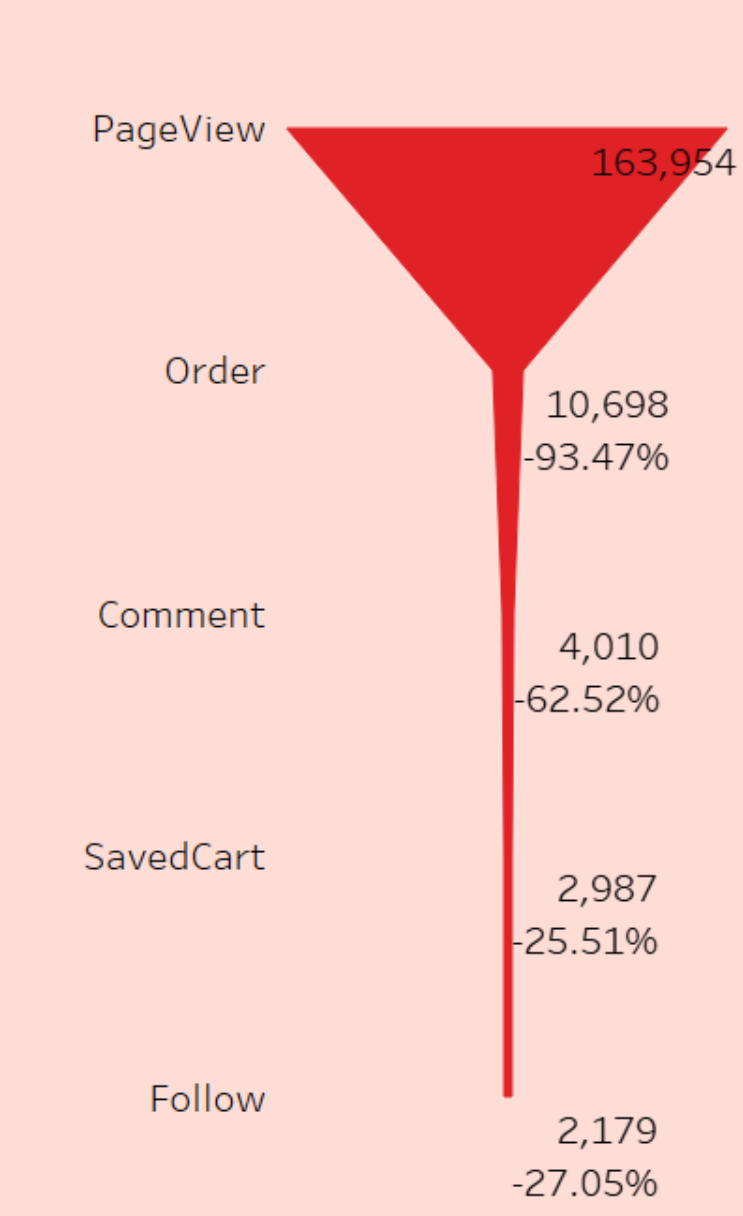
用户下单趋势（小时）



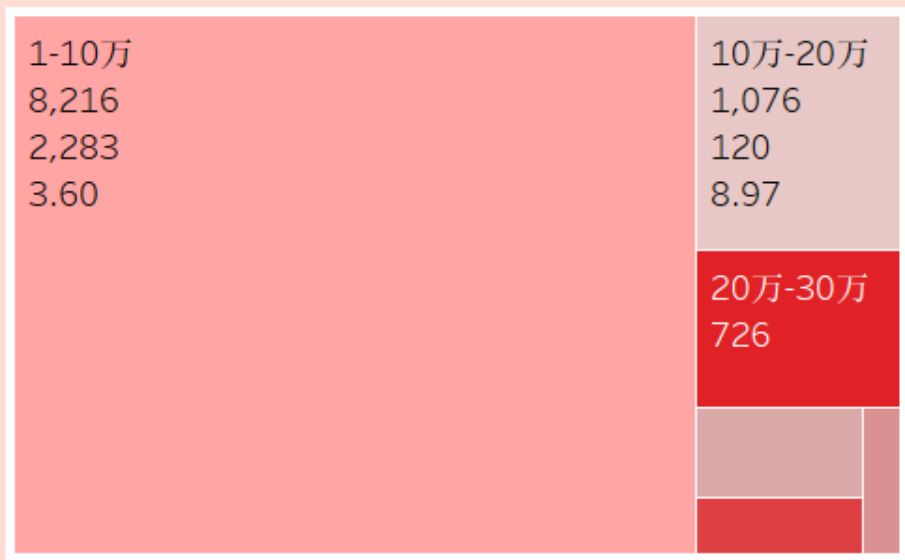
消费者性别



消费行为漏斗图



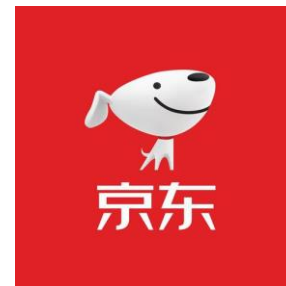
店铺粉丝数与订单量



店铺种类分布



提升销量 – 聚类分析



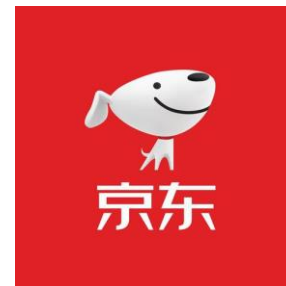
- 通过聚类分析 (k-means) 将用户细分为组内相似的类别，并进行针对性的营销策划
- 变量选择：客户信息（年龄，性别，地区，学历），店铺信息（店铺种类，产品种类，店铺评分，店铺粉丝数，品牌，店铺近一个月销量），购买记录（消费金额，购买数量，是否使用折扣券，消费时间段）
- 类别分析：组之间差别应较大，组之内的用户应相似，根据每组特征制定有效的营销方案，如针对爱用折扣券的消费者组应该多发放优惠券，针对年轻妈妈组应多推送婴幼儿产品信息，针对爱护肤的年轻学生组应推送性价比较高的美妆护肤品，针对爱科技的男性用户组应推送新上市的产品信息

提升销量 – 分类分析



- 通过分类模型来判断消费者是否会购买产品
- 模型选择：决策树，分层限制为5-7级
- 变量选择：客户信息（年龄，性别，地区，学历），店铺信息（店铺种类，产品种类，店铺评分，店铺粉丝数，是否品牌，店铺近一个月销量是否领先同类），购买记录（消费金额，购买数量，是否使用折扣券，消费时间段）
- 结果分析：从决策树可以直观的观察出会进行购买的用户集合以及影响其购买行为的关键因素，从而可以制定针对性的营销计划

总结



- 京东男性用户占比55%，略高于女性；男性用户最爱购买食物，美妆护肤用品和衣物；女性用户最爱购买美妆护肤产品，家用电器，食物。均为日用快消产品。此类产品品类多，上新速度快，应保证新品及时上市并在晚间（8-10点）向粉丝推送
- 尽管大部分订单都来自粉丝较少的店铺，但是粉丝数高的店铺订单量高于粉丝数低的店铺，可能是由于粉丝能及时得知店铺上新和优惠，店铺应给与激励增加粉丝数。同时应深挖小店铺订单量低的原因，如果是流量不够，则应帮助其引流；如果是转化率不高，则应升级店铺产品展示，客服质量等
- 注册时间5年内的用户订单量远高于注册5年以上的用户，管理者应与消费者生命周期相对比，如果是自然流失，则应注重拉新；如果是产品未能满足消费者需求，则应加紧店铺建设；如果是app用户体验不佳，则应挖掘具体环节并优化
- 大部分消费者喜欢直接在商品详情页下单，跳过加购物车环节，需要深挖原因。