

#### **Business Data Analytics Report**

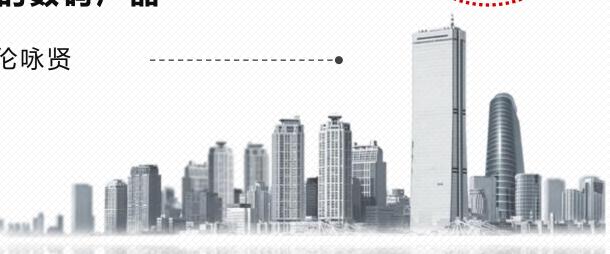
## 京东数码商品

### 销售特征及流行趋势研究

探寻最具价值的数码产品

汇报人:伦咏贤









目录 CONTENTS 第一章

项目简介

第二章

数据统计可视化分析

第三章

影响数码产品销量的因素

第四章

消费者的评价分析

第五章

总结



# (制) 以同的介

- □ 项目背景
- □ 项目目标

- □ 京东简介
- ■数据准备

### 





#### 项目背景:

科技的迅速发展带来了数码产品的日新月异,极大的丰富和便利了人们的日常 生活。在网购成为日常的今天,人们购买数码产品多数会从网上购买。

#### 京东简介

京东3C数码起家,主打电脑、数码、通讯、家用电器产品销售,产品种类繁 多,近年来发展的产品领域涉及多个,不仅限于3C产品。

而京东一直在数码产品销售方面有这着亮眼的销售业绩,这一方面得益于低线 市场爆发的强劲消费力,另外一方面得益于C2M对的"爆品"催化效果。

#### 项目目标

本项目在京东商城上收集了娱乐影音、摄影摄像、智能数码等各类数码产品信息 息,经过数据清洗过滤之后,抽取数据特征,通过比较、构成、分布分析等方式,解 读数码产品的行业现状和销售特征、挖掘影响数码产品销量和受欢迎程度的主 要因素、探究最具价值的数码产品。

### ■ 1.2-数据准备



数据爬取

用Python中的 Request库爬 取京东商城的数 码商品相关数据 并将数据转换成 csv格式

数据规模

爬取京东 10923条所有 商品相关数据和 19037条商品 评论数据

数据处理

运用python使用字符串处理方法,删除、创建商品评价、好评率、重量等对应的数据列

相关字段

商品ID, 现价, 品牌, 评价数量, 是否自营, 重量 类别, 商品评价 追评, 好评率等

本项目收集了京东商城的10类数码产品数据,包括娱乐影音、摄影摄像、智能数码等各类数码产品信息,经过数据清洗过滤之后,抽取各类数码产品的销量、价格、评价、品牌等参数,通过比较、构成、分布分析等方式,探索数码产品的销售特征,并对比京东自营和非自营的产品,将相同品类的不同品牌之间进行对比,找出各类产品的优势品牌、优势店铺;通过线性回归方法和主成分分析挖掘影响数码产品销量的因素,用文本分析寻找受欢迎程度的主要因素,探究最具价值的数码产品。

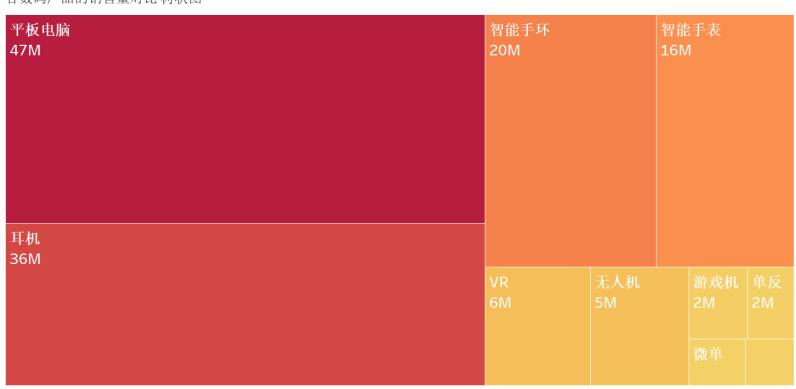


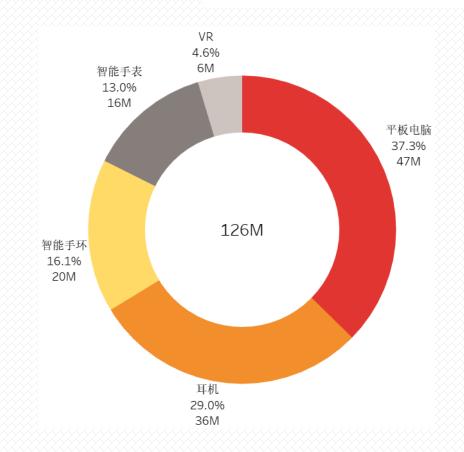


### ■ 2.1-各数码产品市场占比







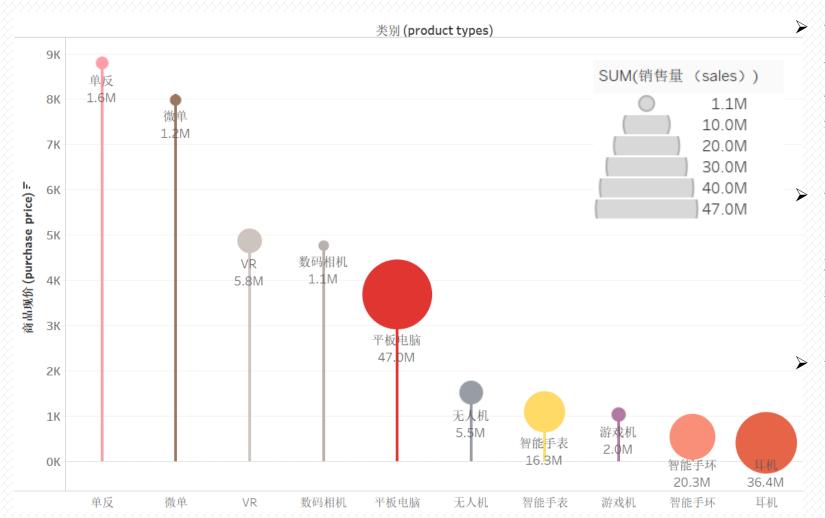


通过树状图显示10类数码产品的市场占比,平板电脑以47万的销售量领先,耳 机以36万的销售量排名第二,智能手环和智能手表也颇受欢迎,游戏机、单反、 微单、数码相机市场份额较少;

基于销售量排名前5位的产品,画出环形图,得出平板电脑 的市场占比37.3%最高,耳机占了29%,智能手环占了 16.1%,智能手表占了13%, VR以4.6%占比最小;

### ■ 2.2-各数码产品价格与销量分析





▶ 气泡图中看出,单反、微单等相机的平均价格较高, 在¥5000-¥9000范围,而平板电脑、无人机、智能 手表、游戏机、智能手环、耳机等价格处于中低价格 范围¥100-¥5000;

气泡大小显示出销售量的多少,对比后发现,中低价的数码产品更受欢迎,特别是平板电脑以47万的销售量领先其他产品,均价最低的耳机销售量排名第二,智能手环和智能手表也有20万的销售;

分析得出,中低价格的数码产品更吸引顾客;

### ■ 2.3-各数码产品的好评率、差评率分析

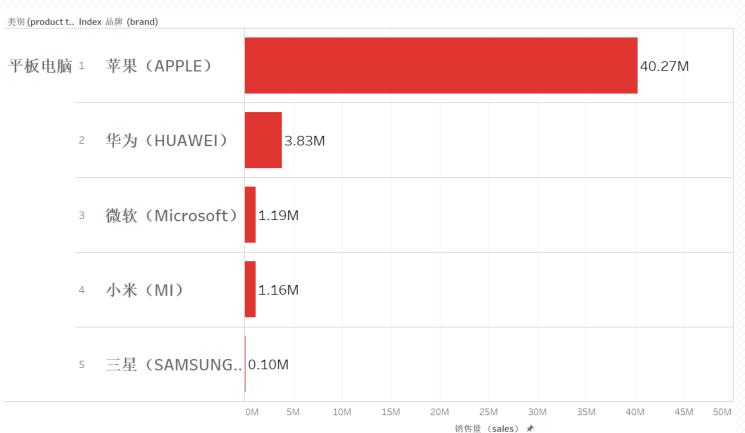




- 从左图的数码产品好评率和差评率对比来看,单反的好评率最高,达到91.5%,其他的摄影摄像设备, 比如数码相机、微单,和娱乐影音设备的好评率也是属于前列;
- 智能设备的好评率相对低,比如智能手表、智能手 环的好评率在50%以下,比较靠后;
- ➤ 智能设备VR的差评率达到1.32%最高;其他数码产品的差评率较低;

### ■ 2.4.1-各数码产品Top5的优势品牌——电脑

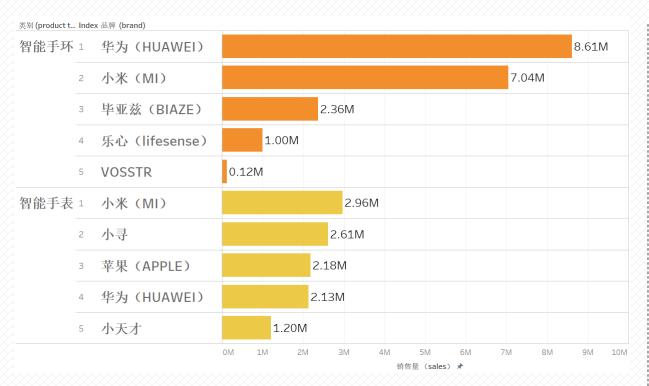




➤ 基于多种品牌平板电脑的销量,分析出优势品牌;从左 图看出,苹果平板电脑占据相对优势,以40.27M的销 量领先第二名华为品牌37M左右的销量;微软、小米、 三星的平板电脑的销量分别占据后3名;

### ■2.4.2-各数码产品Top5的优势品牌—智能设备 JD. 京东





#### 智能手环:

• 华为品牌的销量排名第一,小米排名第二;

#### 智能手表:

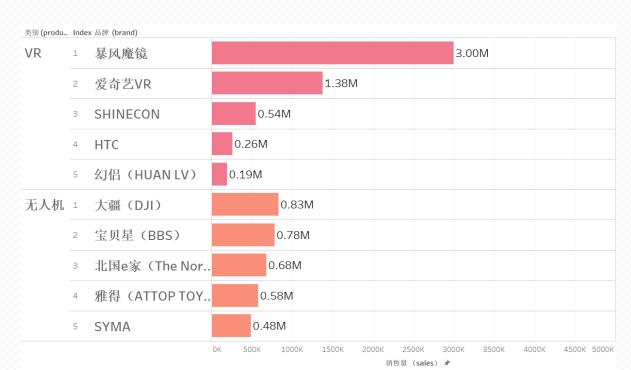
• 小米位列第一,小寻、苹果、华为、小天才分别是第2、3、 4、5名;

#### VR:

暴风魔镜品牌备受欢迎,销量排名第一,爱奇艺排名第二;

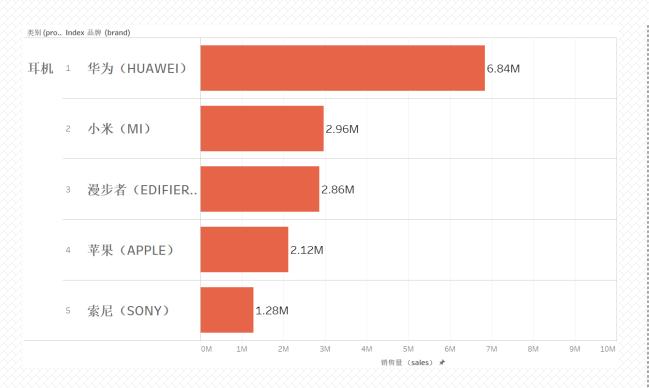
#### 无人机:

大疆无人机品牌销量第一;



### ■2.4.3-各数码产品Top5优势品牌—娱乐影音



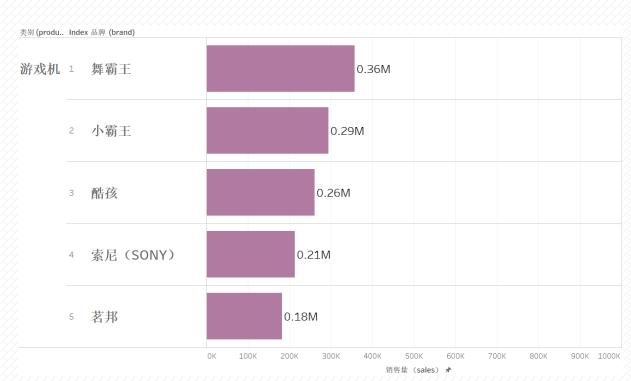


#### 耳机:

- ▶ 华为品牌的耳机以6.84M销量占据相对优势;
- 一些大品牌比如小米、漫步者、苹果、索尼也是前五的优势品牌

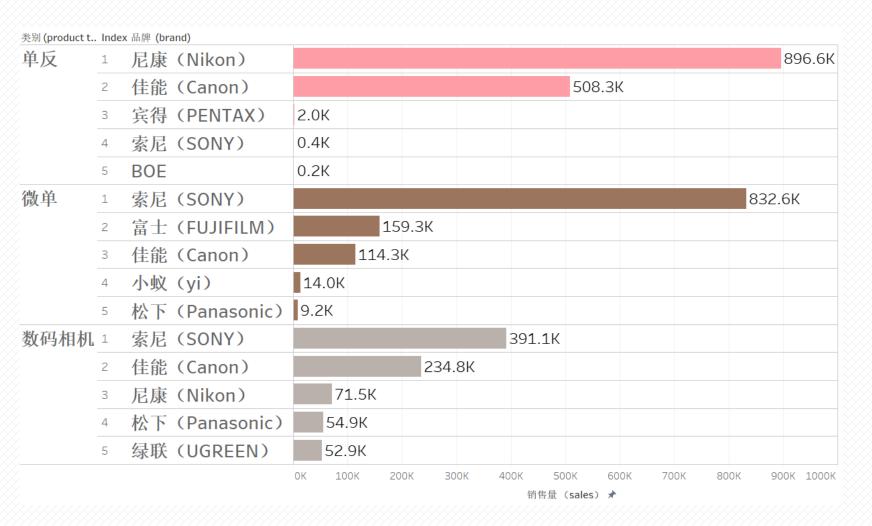
#### 游戏机:

舞霸王品牌的游戏机销量最多,其次是小霸王、酷孩、 索尼、茗邦;



### ■2.4.4-各数码产品Top5的优势品牌—摄影摄像 JD. COM





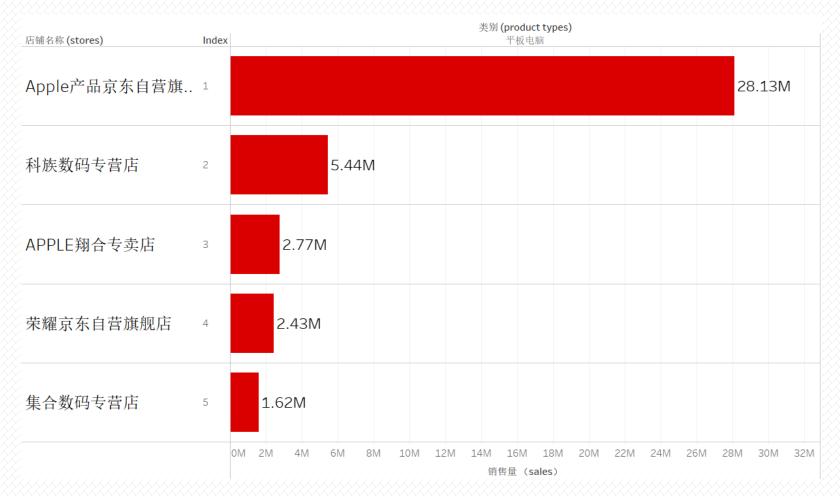
▶ 摄影摄像设备中,尼康的单反以896.6K 销量排名第一,领先第二名的佳能有 400K左右的销量,其次宾得、索尼、 BOE依次排名;

微单中,索尼的销量遥遥领先其他品牌, 备受大众欢迎,随后是富士、佳能、小 蚁、松下的品牌;

数码相机中,同样是索尼的销量第 一,其他老品牌比如佳能、尼康紧 跟其后;

### ■ 2.5.1-各数码产品Top5的优势店铺——电脑

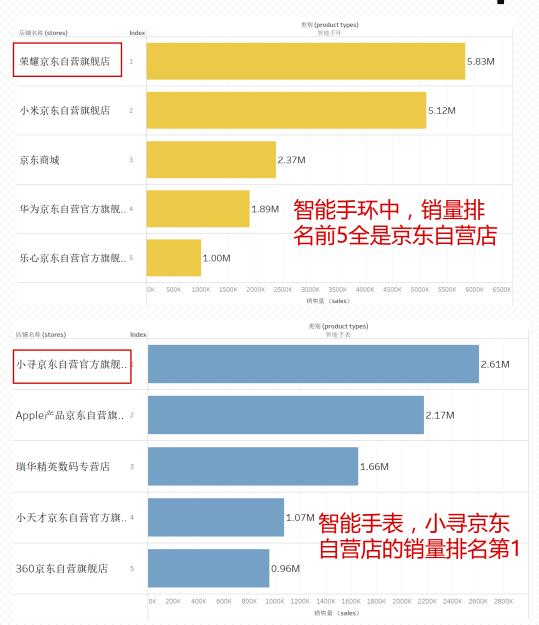


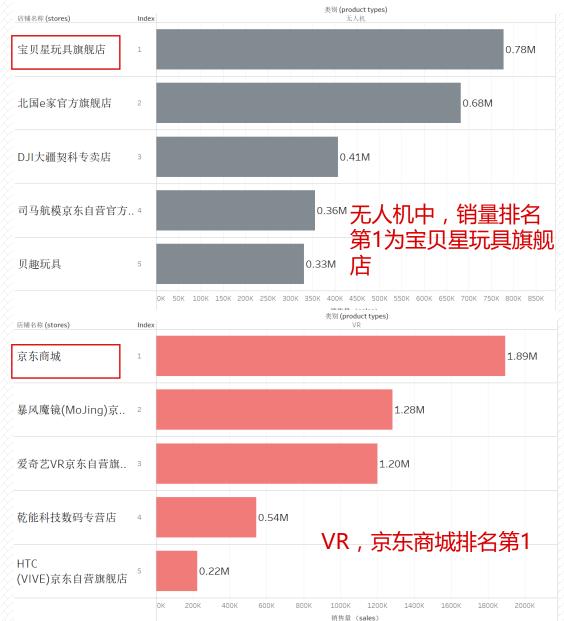


- 排名前五的优势店铺包括京东自营和非自营的;
- 苹果产品的京东自营旗舰店销量排名第一,比第二名多出23百万左右的销量,有突出的优势

### ■2.5.2-各数码产品Top5的优势店铺一智能设备 JD. 京东

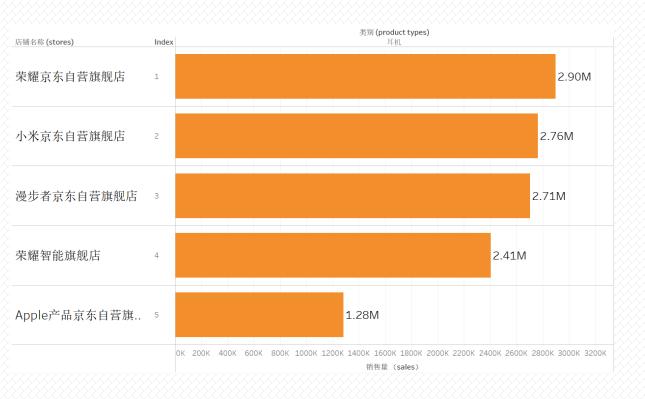






### ■2.5.3-各数码产品Top5的优势店铺—娱乐影像 JD. COM



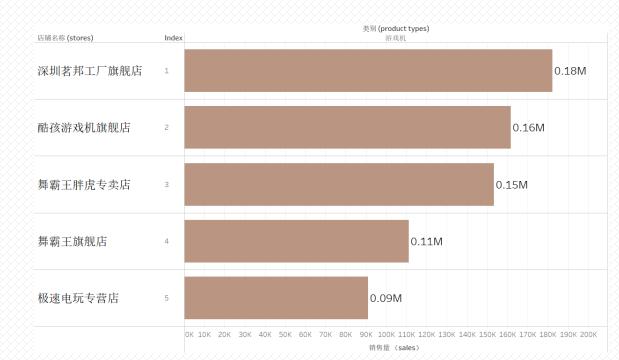


#### 耳机:

▶ 前5名的优势店铺,京东自营店荣耀旗舰店排名第一,小 米第二,漫步者、荣耀智能、苹果的产品紧跟其后;

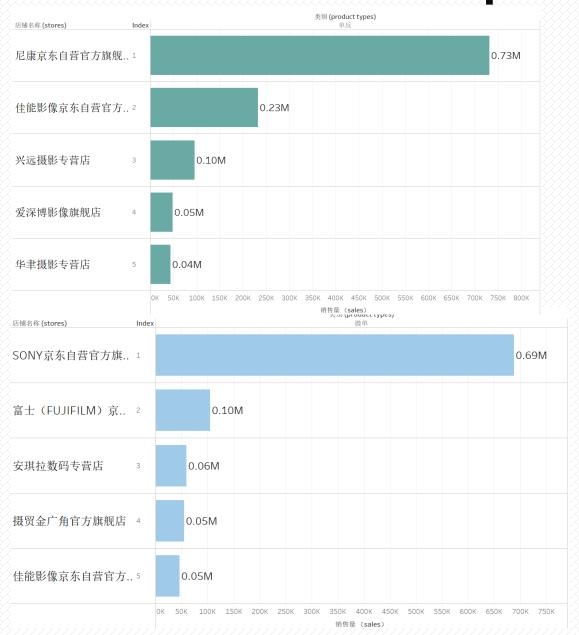
#### 游戏机:

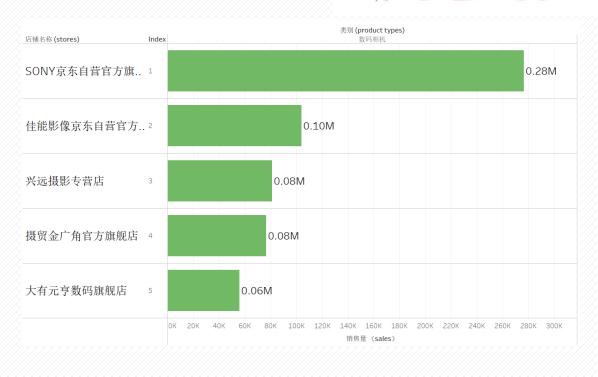
▶ 前5名的优势店铺主要是第三方的非自营店铺,以深 圳茗邦工厂店为首,酷孩游戏机、舞霸王、极速电玩 店铺分别是后4名;



### ■2.5.4-各数码产品Top5的优势店铺—摄影摄像 JD. COM





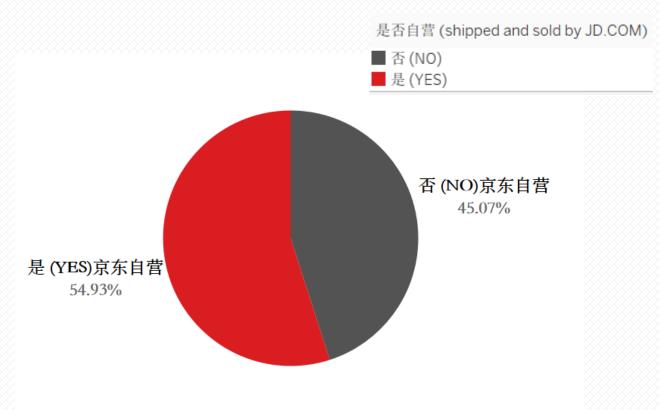


- ▶ 图中分别显示3种摄影摄像的优势店铺
- ▶ 单反销量中,尼康京东自营店以0.73M销量遥遥领先其他店铺,佳能 以0.23M排名第二。
- ▶ 微单销量中,索尼SONY京东自营店以0.69M销量领先其他销售微单的 店铺,富士排名第二
- > 数码相机销量中,索尼SONY京东自营店仍然占据位于第一,佳能影像 排名第二

### ■2.6.1-京东自营 vs 非自营 销售和评价对比



#### 京东自营销售占比饼图



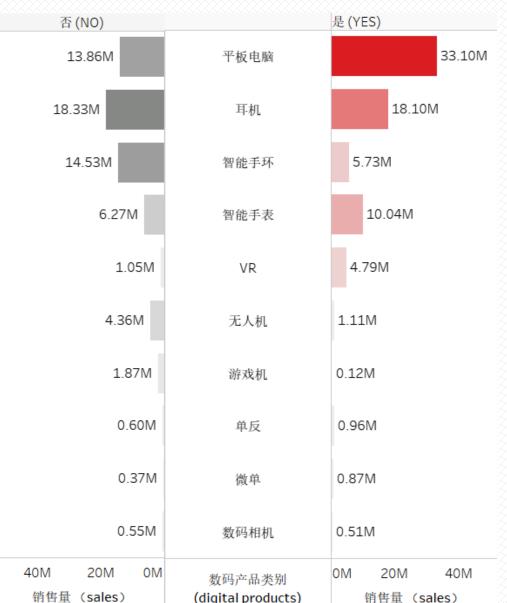
是否京东自营	评价?(销 售量)数量	好评率		
是	75 M	81.5%		
否	68M	72.5%		

- 通过发现京东自营的优势店铺,我们进一步对比京东自营和非自营的数码产品在评价上的差别;
- 从右边的饼图看出,京东自营的销售量以54.93%的占比领先非自营的45.07%;
- ▶ 从表格上看出,京东自营的好评率比非自营的明显高出9%;

总结:京东自营的数码产品对比非自营的占据了很大的优势,在

评价数量、销售量、好评率上都是位列第一

### ■2.6.2-京东自营 vs 非自营的各类数码产品对比 JD.京东



- 通过发现京东自营的优势,进一步分析京东自营和非自营的各类 数码产品在销售量上的差别;
- 从旋风图看出,京东自营的平板电脑,智能手表、单反、微单更受欢迎;而非自营的耳机销量比京东自营的更胜一筹,同时非自营的智能和娱乐设备比自营的也更受欢迎,比如手环、无人机、游戏机;
- 京东商城中的商品多为各大旗舰店或京东自营销售,在保证产品多样化的基础上也注重品质化、专业化;

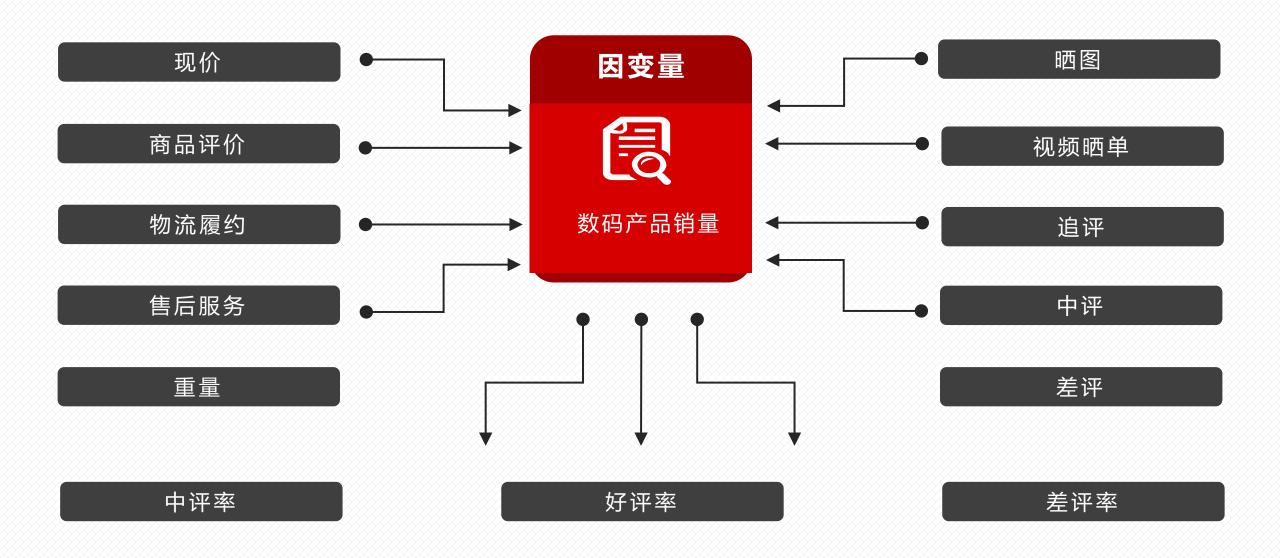
对比京东自营的销售量和非自营的旋风图 是否自营 (shipped and sold by JD.COM)





### ■ 3.1-影响销量的因素





### ■ 3.1.2-影响销量的因素





▶ 左边热力图,颜色条表明变量间的相关性强度,显示自变量与因变量以及自变量之间的相关系数,初步探索相关性。图中分析出,销量(评价数量<sup>①</sup>)和视频晒单、追评、好评、等评价有较强的相关性;

- 0.8

- 0.2

发现全部评价与好评是1的相关性,所以舍去"全部评价"、"好评"的数据列;

注释:①用评价数量表明销量

### ■ 3.3-建立多元线性回归模型



构建多元回归模型,销量为因变量,自变量为现价,商品评价,物流履约,售后服务,晒图,视频晒单,追评,中评,差评,重量,好评率,中评率, 差评率;



初步建立的多元线性回归模型存在缺陷: 存在多重共线性

经标准化和多次修正,从13个变量里提取3个主成分 $X_1$ ,  $X_2$ ,  $X_3$ , 即降维到3维,然后建立修正多元线性回归模型,尽量让信息损失最小化和消除多重共线性影响。分析发现X2的P值>0.05,所以 $X_2$  舍去。最后进行主成分回归分析,得出2个主成分, $X_1$ 和 $X_3$ ;

			王成分  egressi		5果====== esults 		
Dep. Vari Model: Method: Date: Time: No. Obser Df Residu Df Model: Covariance	Fr rvations: uals: :		2020 0:12 7721 7717 3	Adj. F-sta Prob	uared: R-squared: atistic: (F-statistic Likelihood:	):	0. 811 0. 811 1. 103e+04 0. 00 -81721. 1. 635e+05 1. 635e+05
	coef	std err		t	P> t	[0. 025	0. 975]
const x1 x2 x3	5651. 6557 1. 103e+04 107. 8531 757. 6985	108. 828 60. 684 84. 055 101. 108	181. 1.	932 724 283 494	0. 000 0. 000 0. 199 0. 000	5438. 323 1. 09e+04 -56. 918 559. 499	1.11e+04 272.624
Omnibus: Prob(Omni Skew: Kurtosis:		11.	281 000 834 536	Jarq	in-Watson: ue-Bera (JB): (JB): . No.		1. 838 29044498. 245 0. 00 1. 79

~ <del>~</del> //  $\square$  /  $\square$  /  $\square$  /  $\square$ 

### ■ 3.3.2-建立主成分回归模型



	现价	商品 评价	物流 履约	售后 服务	晒图	视频 晒单	追评	中评	差评	重量	好评 率	中评 率	差评 率
X1	-0.07	0.02	0.12	-0.06	0.27	0.45	0.49	0.45	0.5	-0.03	0.09	0.01	0.01
Х3	-0.49	0.16	0.3	0.23	-0.17	-0.03	-0.02	80.0	0.06	-0.12	-0.68	0.26	0.05

 $X_i = k_1 \times 现价 + k_2 \times 商品评价 + k_3 \times 物流履约 + k_4 \times 售后服务 + k_5 \times 晒图 + k_6 \times 视频晒单 + k_7 \times 追评 + k_8 \times 中频 + k_9 \times 差评 + k_9 \times 重量 + k_{10} \times 好评率 + k_{11} \times 中评率 + k_{12} \times 差评率$ 

#### 主成分载荷图

- ✓ 当返回主成分回归模型的各个特征向量,可以发现,系数的绝对值越大,对应变量在计算主分量时就越重要。
- ✓ 第一个主分量X1与视频晒单、追评、中评、差评具 有较大的正关联,因此X1分量主要度量数码产品的评价。
- ✓ 第二个分量PC3与<mark>好评率和现价</mark>具有较大的负关联,因此PC3分量主要度量数码产品的评价和现价的影响。



 $X_3$ 

视频晒单、追评、中评、差评

好评率、现价





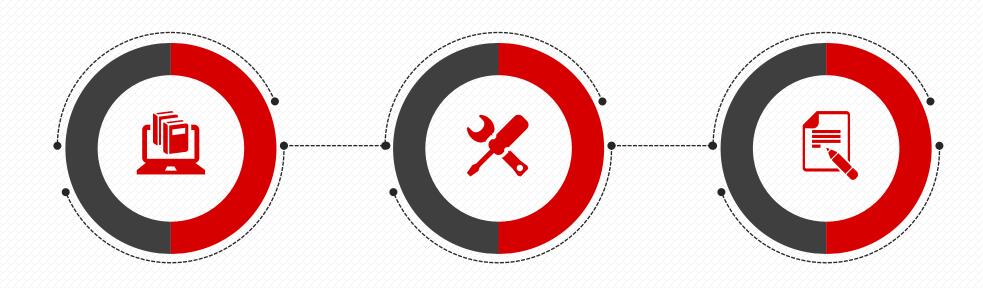
基于上述模型预测出,数码产品的评价是影响销量的最主要因素,其中包括视频晒单、追评、中评、差评、好评率对数码产品销量影响最大,再者,现价也在一定程度上影响





### ■ 4.1-消费者评价的分析





▶ 消费者评价词云图。针 对12类数码产品的消费 者评价进行分词处理, 画出词云图;

- ▶ 针对12类数码产品进行 情感分析,得出3种情 绪的占比;
- ➤ LDA主题模型建立:基于销售额 均值排名前3的商品的顾客评价 建立LDA主题模型,根据各关键 词提取了4个核心主题并总结概 括

### ■ 4.2-消费者评价的词云图





- 》 基于10类数码产品的评价内容,包括单反,平板电脑,微单,数码相机,无人机,VR,智能手环,智能手表,游戏机,耳机,进行分词处理,画出词云图。
- 消费者的评价关键词集中在"功能"、"不错"、"很快"、"效果"、"满意"、"京东"、"真的"、"喜欢"、"手机"、"相机"、"物流";
- 消费者对京东商城的数码产品总体评价不错,特别是功能好、质量高、物流快;



基于消费者对京东商城12类数码产品的评价数据,从"正品"、"满意"、"好"、"不错"等关键词里反映出顾客对京东数码产品的多数好评和认可,体现京东在数码产品销售上的优势和好口碑,从而为京东商城对数码产品的未来发展上提出决策参考。

### ■ 4.3-情绪分析



	积极情绪	消极情绪	平和情绪
游戏机	80%	14%	5%
无线耳机	81%	14%	4%
有线耳机	91%	5%	3%
头戴式耳机	83%	12%	4%
智能手环	86%	10%	3%
智能手表	87%	9%	3%
VR	91%	5%	3%
无人机	89%	7%	3%
平板电脑	92%	5%	1%
单反	89%	7%	3%
微单	90%	6%	2%
数码相机	91%	5%	3%

- ➤ 基于12类数码产品的消费者评论,使用python的 SnowNLP分析3种情绪(积极情绪,平和情绪,消极情绪)的占比。
- ▶ 平板电脑的积极情绪的占92%,是12类商品中占比最高
- > VR、数码相机、有线耳机积极情绪较高
- > 游戏机和无线耳机的消极情绪最高

### ■ 4.4.1-LDA 主题建模——平板电脑



▶ 基于销售量总额排名第1的数码产品的顾客评论建立LDA主题模型, 根据各关键词提取了4个核心主题并总结概括。



#### 电脑质量一流

主题 1:好 不错 流畅 屏幕 高 系统 很快 够用 性价比 电



#### 物流优势

主题 3 :速度 物流 快 好 很快 满意 购物 喜欢 真的 包装





#### 商品用途适合学习

主题 2:学习 学生党 视频 好 真的 iPad 不用 平板 用来



#### 品牌优势

主题 4:买 正品 苹果 不错 京东 平板 喜欢 iPad 手机 系统

### ■ 4.4.2-LDA 主题建模——耳机



➤ 基于销售量总额排名第2的数码产品的顾客评论建立LDA主题模型,根据各关键词提取了4个核心主题并总结概括。



### **111**

#### 耳机质量好

主题1:耳机 音质 好 不错 声音 舒服 听 效果 清晰 质量



#### 物流优势

主题 6:京东 购物 快 好 满意 好评 宝贝 速度 物流 服务



#### 用途好

主题 2:环境 降噪 耳机 音质 模式 功能 音 外观 高 听歌



#### 实惠优势

主题 4:一个 买 好 质量 收到 不错实惠 京东 送 东西

### ■ 4.4.3-LDA 主题建模——智能手环



➤ 基于销售量总额排名第3的数码产品的顾客评论,建立LDA主题模型,根据各关键词提取了4个核心主题并总结概括。



#### 手环质量好

主题 1:功能 手环 彩屏 显示 感觉 屏屏幕 小米 好 亮



#### 物流优势

主题 3 :京东 物流 收到 很快 下单 好速度 快递 第二天 下午



#### 功能好

主题2:买心率 手环 好 功能 续航 时间 运动 充电 睡眠



#### 用途广

主题 4:公交 小米 手环 刷 卡 NFC 地铁 功能 买 价格

### 4.4.4-LDA主题建模——总结分析





京东早期以正牌3C电子数码产品起家,通过高品质的电子产品做起了一定的量和品牌效应,主打电脑、数码、通讯等产品。通过大规模采购和统一质量管理配送而达到消费者认可满意;



京东的优势也在于可以通过大数据技术分析用户画像、精准推送,洞察用户需求及未来消费走势,推进厂商创造出更贴合消费者需求的产品,比如针对学生群体推出学习用途的电脑、方便出行的智能手环等;

京东有布局全国的仓配物流网络,是完全自己建立仓储、物流配送系统,采取分布式库存管理,提前把各供应商库存汇集到各区域,订单产生后快速配送;而且京东搭建起了"物流+互联网+大数据"相融合的一体化产业生态平台,为平台的购物体验打好了地基,因此一条龙式的服务受到了大多数消费者的认同。



京东商城以自营模式为主,同时向第三方卖家招商,发展出旗舰店等合法商家,吸引了受欢迎的品牌,比如华为、苹果等;这种B2C商务模式需要企业对商品质量严格把控并且熟知商品信息;因此自营的产品质量和售后服务上面,也都有一定的保障。







### ■ 5.1-总结



1

#### 数据统计可视化分析

通过各个维度数据比较、构成、分布的可视化分析,平板电脑的销售量和市场占比最高,耳机排名第二,智能设备如手环、手表、VR销售量也有可观的销量,而摄影摄像设备如单反、微单的市场占比较低;通过价格与销量分析出,中低格价的数码产品更受欢迎;

至于各类的优势品牌,苹果的平板电脑占据了优势,销售量遥遥领先其他品牌;而属于智能设备、娱乐影像设备的,华为、小米、索尼等大众品牌更受信赖和欢迎;摄影设备中,尼康、佳能、富士等经典品牌有明显优势;最后在分析店铺和自营对比中发现消费者更钟爱京东自营的商品;

#### 影响数码产品销量的因素分析

基于主成分回归分析表明,评价和现价是影响数码产品销量的最主要因素,其中评价包括视频晒单、追评、中评、差评、好评率对数码产品销量影响大。顾客的评价影响潜在消费者的购买欲望,当好评率越高,产品得到的好回馈越多,越受消费者欢迎;而产品的价格也影响消费行为,顾客考虑数码产品是否有性价比,能否承受其价格;

#### 消费者的评价分析

基于消费者的评论数据进行了情绪分析,显示平板电脑受到最高比例的积极反响,而游戏机最低; 从LDA主题模型间接反映出,消费者的关注重点一般在数码产品的质量、用途、物流、自营的方面;京东近年 来开发低线级市场,利用全球范围的智能物流基础网络、供应链平台等为低线城市用户提供高品质的产品和 24小时到达的物流服务,而且采用的各大旗舰店或京东自营销售使得质量有保证;

2

3

### ■5.2-最具价值的数码产品——平板电脑





- 平板电脑在10类数据产品中最为出众,销售量以47M占据第一, 市场占比在前5名产品里占了37.3%;
- ✓ 消费者的评价里,积极情绪以92%位列第一;消费者认可平板电脑的高质量、高性价比、用途广和优质物流服务,
- ✓ 京东自营模式和旗舰店保证了平板电脑的高品质;
- ✓ 近年来京东发展低线级市场,此市场累计消费增速最快的有电脑产品,得益于低线市场爆发的消费力,从而使平板电脑产品拥有巨大的销售潜力;同时,由于C2M模式的采用(即制造业与消费者的衔接),京东利用了智能引擎、大数据分析等,挖掘用户数据、联合品牌商、整合资源,共同打造用户需要的产品而平板电脑因学习用途的优势受到学生党的欢迎;
- ✓ 平板电脑在各个方面都是佼佼者,是最具价值的数码产品;

平板电脑是最具价值的数码产品

### ■ 5.3-局限与展望





#### ● 局限

本报告仅有不同数 码产品的横向对比分析, 缺乏各类商品的时间序 列数据,不能反映各类 商品的发展趋势,分析 有欠缺。



- 探索消费者关注的其他因素;
- 制造多样化的促销策略;
- 保证数码产品的价格和质量相符;
- 注意发展潜在性客户群体, 比如大学生;





### THE BUSENESS PROJECT

# THANKS

谢谢聆听

