

# 广东省各区/县生鲜电商表现分析

队伍编号：1

队伍名：老爷保号

成员：姚瀚普 - 宾夕法尼亚大学

陈钺 - 埃塞克商学院

吴祥茵 - 杜克大学

## 背景介绍

过去二十多年，中国电商行业经历了惊人的发展。2003 年的非典爆发是中国电商行业的一个重要转折点。当时由于对疫情的担心，人们被迫开始接受网上购物，需求端的大幅增长催化了当时正在萌芽期的中国电商，在其后近二十年里，网络购物已经逐渐成为大众不可或缺的生活方式之一。中国电子商务行业市场规模已从 2002 年的 1809 亿元，发展到 2020 年的接近 35 万亿元。

在“互联网+”的大背景下，稳定增长的电商行业也在寻找新的市场，而生鲜就是许多创业者的兴趣所在。得益于庞大的土地面积和人口基数，我国是一个传统农业大国，民间素有“民以食为天”的俗语。巨大的市场潜力和对互联网颠覆传统行业的自信显得生鲜电商市场是个新风口。然而自 2005 年首家生鲜电商易果网上线以来，该行业并未形成气候：跟据中国电子商务研究中心的统计数据，国内生鲜电商行业有 4000 余家入局者，其中仅有 4%收支平衡，88%陷入亏损，仅有 1%的企业能实现盈利。早期从业者总结出生鲜产品的保质期短、季节性强、储藏运输要求高、消费弹性系数小等特点制约了生鲜线上化的发展，于是，这二十年生鲜电商并未搭上电商行业顺风车。

这不堪直视的成绩不仅是因为供应端的品控、供应链、损耗率的不足，需求端的市场也并未打开。彼时人们购买果蔬肉蛋等生鲜产品时更加习惯在菜市场 and 超商等线下的购物场景，因为人们对生鲜有新鲜和安全的需求，需要亲自挑选，同时生鲜作为高消费频率的刚需特点也反向增强了用户对线下消费场景的信任和依赖，而不像电器等商品已经培养了用户网购习惯。生鲜电商除了本身基础设施等硬实力的增强外，亟待一个需求端破局的契机。

2020 年突如其来的疫情打乱了世界的节奏，为了避免感染和尽快控制疫情，人们纷纷响应号召减少外出，于是线上生鲜配送成为许多家庭的选择，生鲜电商的需求暴增。近年发展迅猛的冷链物流体系支撑起生鲜电商的供应链，对生鲜配送更高品质和质量的需求也倒逼供应链环节的优化：新出现了前置仓模式，即在区域大仓库的下游将商品前置到更接近消费者的场景（覆盖 1-3 公里的社区），提高了配送效率。而冷链网络和前置仓模式的推广需要较重的资本投入和较高的运营效率，使资本有深度介入的空间，加以众多其他因素的催化，生鲜电商赛道已成为资本市场的新宠。

在新零售的风口和疫情培养了用户习惯的双重刺激下，生鲜这一原本与群众生活密切相关，但又存在感低的产业，已经摇身一变成为话题焦点。根据中国市场信息调查业协会《2020 线上生鲜行业报告》，生鲜网购保持高速增长的趋势，2020 年市场规模高达 2475.7 亿元，同比增长达 48.9%，增速远超往年。诚然，生鲜电商的各种商业模式仍在探索中，这行业究竟是像非典后的电商一样蓬勃发展成为生活方式，亦或只是疫情下的昙花一现，结果需要市场和时间的检验。

## 研究问题及研究方法

本文通过研究广东省各区/县（或同级行政单位，下文同）内的生鲜电商网点数量，与当地的经济、人口等指标，探究广东省各地区生鲜电商发展状况，并通过聚类分析将各区/县分为特点鲜明的四类地区，并给出针对这几类地区的生鲜电商发展建议。

## 数据说明与描述

### 数据来源

1. 广东省各区/县的人口及经济数据：取自各市 2020 年统计年鉴、政府网站，统计于 2019 年末；
2. 广东省各区/县的生鲜电商数据：通过八爪鱼采集器于 2021 年 4 月采集自百度地图；所获数据包括各生鲜电商名称、经纬度、地址；

### 数据说明

研究分析广东省内区/县 185 个，爬取广东省内生鲜电商 POI 数据 922 条。数据集字段包括：城市、区/县、土地面积（K m<sup>2</sup>）、常住人口（万人）、人口密度（人/K m<sup>2</sup>）、常住人口增长率（%）、地区生产总值（万元）、地区生产总值增长率（%）、人均地区生产总值（元）、消费价格指数（CPI）、食品消费价格指数、生鲜电商数量、生鲜电商密度（家/K m<sup>2</sup>）、每万人生鲜电商数量。需要说明的字段描述如下：

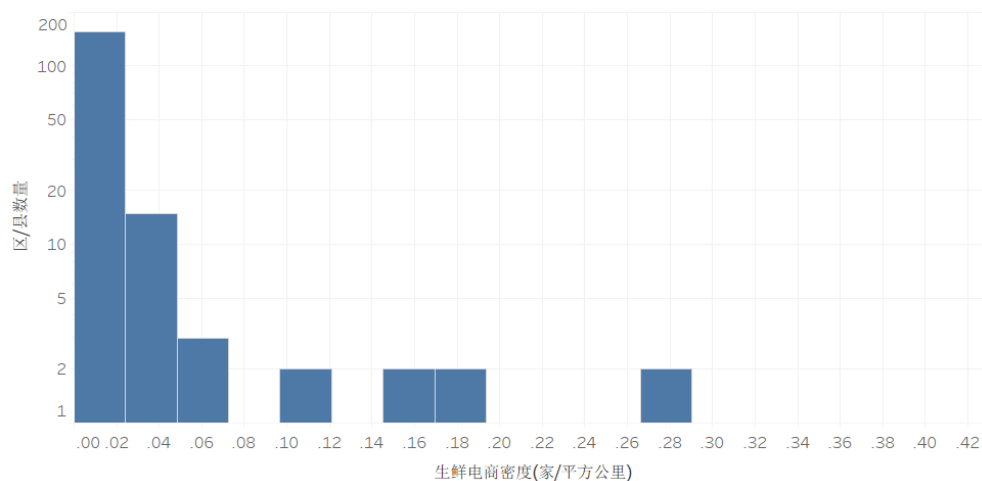
表 1 数据集字段说明

字段	说明
地区生产总值增长率（%）	其中广州、江门、东莞在统计年鉴中无分区/分县/分镇数据。
人均地区生产总值（元）	其中深圳、广州、中山、东莞在统计年鉴中无分区/分县/分镇数据，故按如下公式计算：人均地区生产总值 = 地区生产总值/常住人口。
消费价格指数（CPI）	表现地区物价水平，以该区所在市的指数为该区指数。
食品消费价格指数	表现地区食品物价水平，以该区所在市的指数为该区指数。
生鲜电商数量	范围内生鲜电商网点数量，以市场占有率较高、且能在百度地图中查询到生鲜网点的品牌：美团买菜（美团优选）、叮咚买菜、朴朴超市、每日优鲜、盒马鲜生。
生鲜电商密度（家/K m <sup>2</sup> ）	范围内生鲜电商数量/土地面积。
每万人生鲜电商数量	范围内生鲜电商数量/常住人口*10000。

### 数据描述分析

各区人口经济数据：各区/县常住人口平均值为 62 万人，地区生产总值（GDP）平均值为 576 亿元，不计东莞中山各镇，各区/县常住人口平均值为 80 万人，GDP 平均值为 736 亿元。

各区/县生鲜电商：各区/县生鲜电商数量、每万人生鲜电商数量、生鲜电商密度平均值分别为 7.2 家、0.045 家、0.018 家/平方公里；不计东莞中山各镇，分别为 8.2 家、0.056 家、0.024 家/平方公里。各区/县生鲜电商密度多小于 0.02 家/平方公里（图 1）。

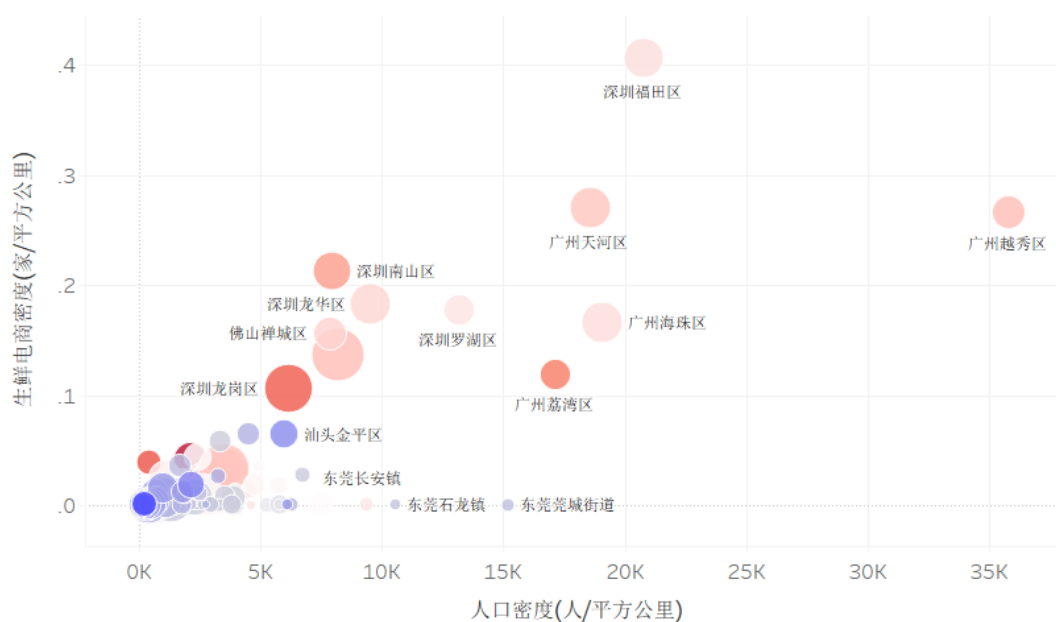


注：纵坐标已做对数处理

图1 区/县生鲜电商数密度频率分布

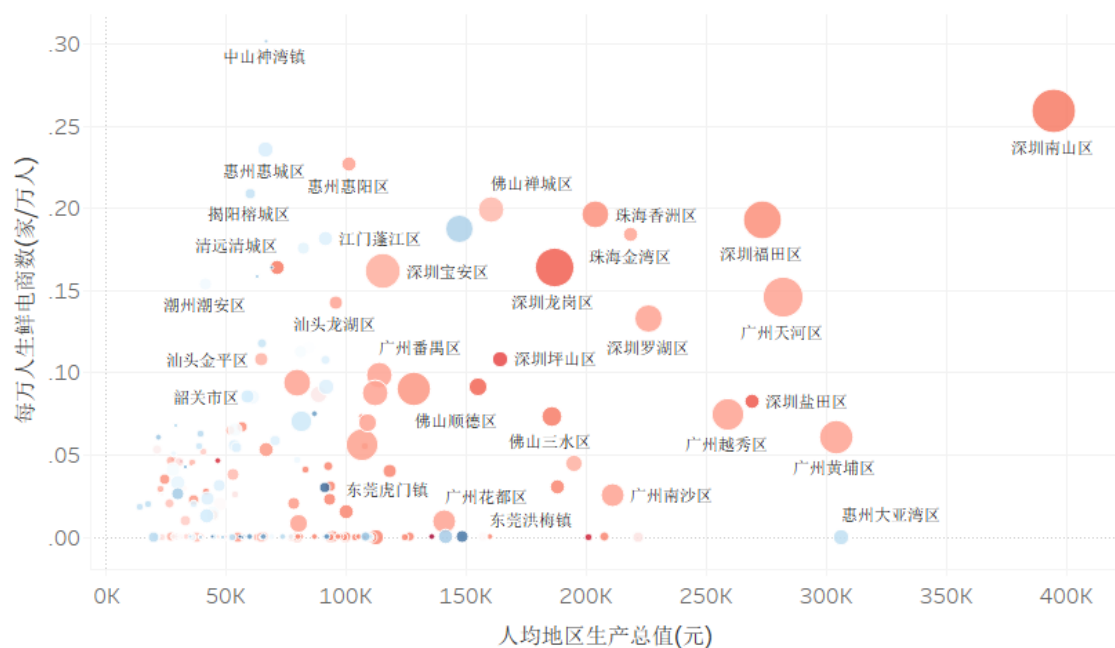
如图2、图3，**生鲜电商密度随人口密度增大而增大**：人口密度高的地区，消费需求大，更能吸引生鲜电商进驻，故这些地区的生鲜电商密度较高。**每万人生鲜电商数量随人均GDP增大而增大**：人均GDP高的地区，其消费水平更高，更能吸引生鲜电商进驻，故人均生鲜电商数量较多。

如图2、图3，**广深可作为生鲜电商持续发展地区**：广州、深圳多个区，每万人生鲜电商数量多密度大，较好的满足消费需求；且经济上规模大增速快，市场规模、市场潜力较大。**深圳龙岗可作为优先布局地区**：深圳龙岗的人口密度及生鲜电商密度都处于中等水平，但其经济人口规模均较大，对于生鲜电商的需求较大；且其经济人口增速均极大，市场潜力大。



注：圆圈大小表示常住人口，颜色表示人口增长率，红色增长率高

图2 人口密度与生鲜电商密度关系图



注：圆圈大小表示 GDP，颜色表示人均 GDP 增长率，红色增长率高

图 3 人均地区生产总值与生鲜电商密度关系图

## 数据分析与解读

为了更好的探究各区/县特点，本文采用聚类方法（属于无监督机器学习）对区/县进行类别划分，考察其在经济水平、人口情况、物价变动水平、生鲜电商水平等 4 个层面 10 个指标变量的聚类中心情况。10 个指标变量均为数值变量，所以用 K-Means 模型进行聚类，并且在聚类前对各变量进行标准化处理。

确定聚类类别个数为 4 ( $K=4$ )。确定使用 K-Means 算法后，需要确定 K 值即簇个数。本文综合手肘法和轮廓系数法进行簇个数的选择：1. 手肘法，通过拟合模型来选择最佳聚类数 K。纵坐标为误差平方和，最佳 K 值出现在曲线的拐点处（图 4）；2. 轮廓系数法，纵坐标为轮廓系数，以高取值处的 K 值作为最佳 K 值（图 5）。综合两种方法的结果，选择 4 个类别进行聚类，经实际测试， $K=4$  时聚类效果好于其他值，无极小样本的地区类别，各类特征也较明显。

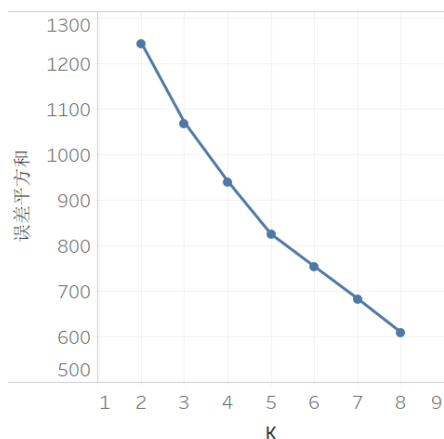


图 4 手肘法结果图

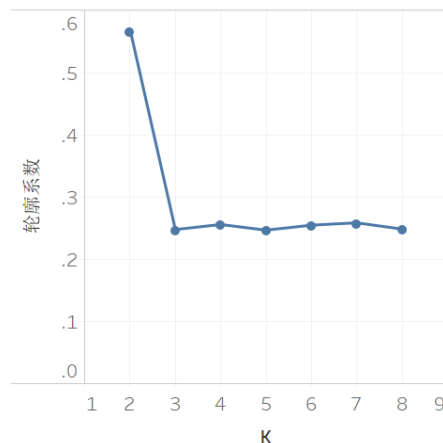


图 5 轮廓系数法结果图

聚类所得四类地区包括热战区、优化区、潜力区、下沉区：

### 1. 热战区：如火如荼的主战场

如图 6，此类地区有 10 个，包括深圳南山区，广州越秀区等。

此类地区的特点：经济人口规模极大，GDP 增长速度中等，人口密度极高，人均地区生产总值高，生鲜电商密度高。生鲜电商属于市场集中度高，头部效应明显的行业，这样的地区正是市场巨头的必争之地，故将其定位为热战区。

### 2. 优化区：精益求精的新贵族

如图 7，这类地区有 24 个，包括惠州惠阳区，潮州潮安区等。

这类地区的特点，是经济人口规模较大，GDP 增长速度中等，人口密度中等，人均地区生产总值较高，生鲜电商密度较高。这说明其市场规模较大，但发展潜力中等；其已有一定的生鲜电商布局，竞争较大；同时，中等的人口密度，使其不能像热战区一样方便地配送，故对单个网点的配送范围及仓储能力要求高。面对中等的市场潜力，较大的竞争及较高的运营管理要求，这类地区应注重经营而非盲目扩张，故将其定位为优化区。



图 6 热战区聚类中心特征



图 7 优化区聚类中心特征

### 3. 潜力区：精致的潜力股

如图 8，此类地区有 54 个，包括佛山高明区，东莞虎门镇等。

这类地区的特点：经济人口规模较小，GDP 增长速度较高，人口密度较高，人均地区生产总值较高，CPI 高，生鲜电商密度低。其经济人口规模虽然较小，但这类地区有 2/3 为东莞、中山两市的镇级行政区（两市无区级行政区），较细的分割不能较好展现这类地区的经济规模；其较高的人均生产总值及人口密度，也反应了该区高质量的经济人口水平，比较“精致”。同时，GDP 增速大，说明该类地区的发展潜力大；当前生鲜电商密度低，很可能无法满足该类地区未来的发展需求。这类地区经济人口水平质量高，发展潜力大，如精致的潜力股，属于潜力区。

### 4. 下沉区：沉默的大多数

如图 9，此类地区有 97 个，包括韶关乐昌市，梅州梅江区等。

这类地区特点是经济人口规模小，GDP 增长率低，人口密度低，人均地区生产总值低，生鲜电商密度低。这说明其市场规模及发展潜力小；该类地区多为各市的**农业地区，靠近生鲜产业链的源头**，传统菜市场较多；同时这类地区收入较低，下沉人群较多。生鲜电商市场及潜力较小，应对该区采取针对下沉市场的差异化经营策略。**故称该区为下沉区。**



图 8 潜力区聚类中心特征

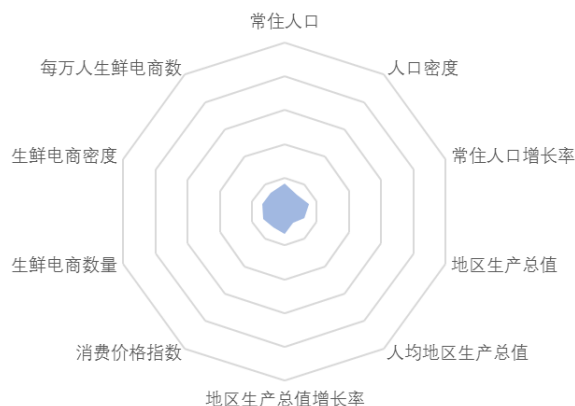
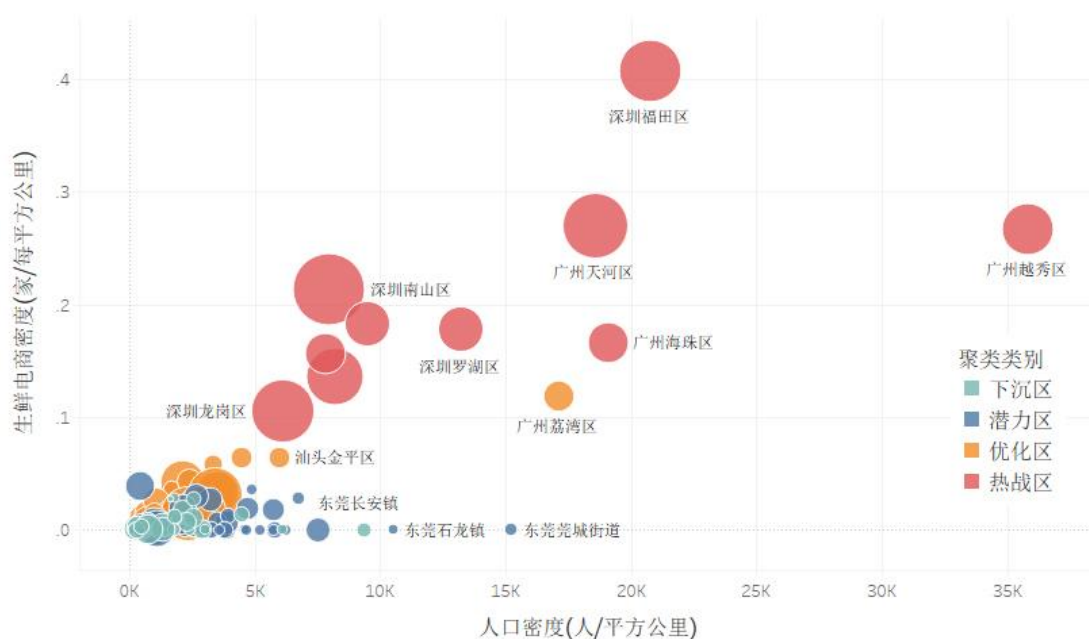


图 9 潜力区聚类中心特征

## 商业应用与总结

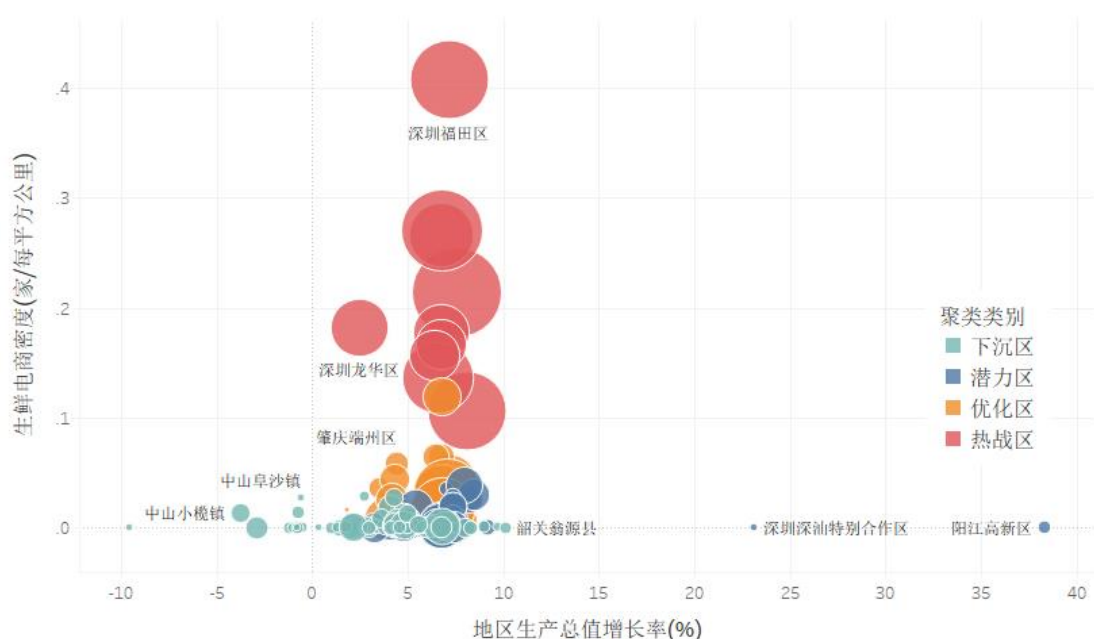
### 商业应用——因地制宜，发现机遇

热战区经济规模大，人口密度高，生鲜电商密度高；优化区经济规模较大，人口密度中等，GDP 增长率中等，生鲜电商密度较高；潜力区经济规模中等，人口密度较大，GDP 增长率较高，生鲜电商密度低；下沉区经济规模小，人口密度小，GDP 增长率低，生鲜电商密度低。



注：圆圈大小表示地区生产总值

图 10 生鲜电商密度与人口密度聚类结果



注：圆圈大小表示地区生产总值

图 11 生鲜电商密度与地区生产总值增长率聚类结果

### 1. 热战区：如火如荼的主战场

商家应在热战区合理对网点选址并积极扩张。热战区特点说明其市场规模极大，网点大量增加不会因需求有限而发展受限；极高的人口密度及高电商密度，使单个网点的配送范围更聚焦，且无须耗费过多运力进行长距离配送。在资金较为充裕的情况下，将资本投入到市场需求大，运营较顺畅的大城市，既能更好满足消费者需求，也能稳定保证引流获客渠道，且网点密度达到一定水平时可以形成规模效应，摊薄运营成本。

### 2. 优化区：精益求精的新贵族

生鲜电商行业复购率高，好的运营管理能够提升消费者粘性，带来区内生鲜电商的持续盈利。新发展起来的优化区商家须更注重已有站点的优化，精益求精。这包括供应链优化及精细化运营，着重提升区内的供应网络效率，提升单个网点的配送仓储及周转能力。基于较强的供应链、运营水平，当资金充裕、选址合适时，可适当增加网点数量。

### 3. 潜力区：精致的潜力股

商家可通过先发优势在潜力区积极扩张。在潜力区，较高人口密度、较高经济增速、低生鲜电商密度条件下，建议：1. 适当地加速铺开小型网点，抢先占领一些市场份额的同时保留进退的空间，可以把握住该地区经济增长的机会；2. 适当加大营销纳新的力度，提前培养用户黏性；3. 最后，随着经济逐步发展，在后期调整预先建立的小型网点：若网点承载能力不够则扩建扩容，若供过于求则调整上游供货量，或通过精细化投放等方式优化供应端。



#### 4. 下沉区：沉默的大多数

该区多为农村地区，人口密度和消费水平低，经济总体量和增速差，从数据上看似乎对生鲜电商价值不高。但**该区离生鲜供应端最近**，所以可以充分利用该区的农业资源，整合邻近区域的本地生鲜产品优势，**延续自产自销、发展本地农场直销等不同于其他生鲜电商的经营模式，盘活本地生鲜电商市场，发展地区内小循环。**

#### 总结

本文爬取汇总广东生鲜电商数据，对省内生鲜电商网点数据和各区/县内人口经济指标等数据进行 K-Means 聚类，从而达到识别各区生鲜电商状况，并依据其特性分为热战区、优化区、潜力区、下沉区，从而达到精准制定商业建议的目的。