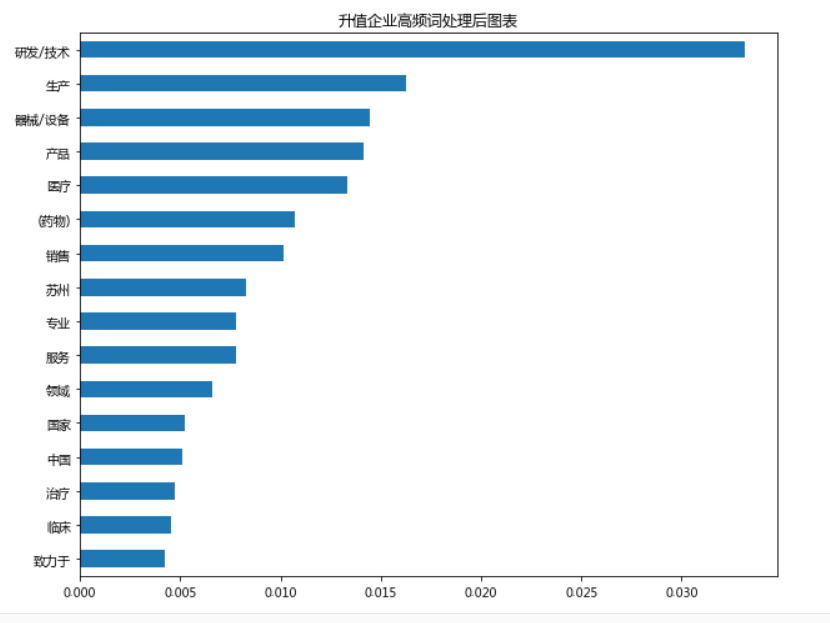
**第一部分 指标确定**  
创新活力指数=企业估值/注册资本  
创新活力指数是为了探索哪些因素会使企业更加富有竞争力、生命力、发展力，它们响着企业的兴衰；为得到这些影响企业兴衰的因素，必须知道哪些企业是“兴”，哪些企业是“衰”。本课题从创新创业绩效来衡量企业的创新创业活力。  
从数学角度上看，我们团队自然地联想到了有相似划分和功能的“survival-model”（生存模型），尝试把已有数据与模型结合起来进行数据挖掘工作。然而，从所提供的数据内容上发现，企业介绍表格的“经营情况”这一个属性，能为模型建立提供基础。但是在模型建立过程中，又发现了很多不合理性：生存模型过于依赖于死亡时间、死亡数据量、死亡因素特征，而数据中没有死亡时间，死亡因素特征不明显，死亡数据量远远不足（在“经营情况”中，注销(即死亡数据)所占比重不到5%）；另外，用“生”与“死”来描述一个企业的发展历程过于极端，因为企业的发展过程是一个复杂过程，并不是非生即死所能代表的。  
但从生存模型中得到启发，利用（升值-贬值）代替非生即死的特征，创造突破口。以“投资回报率”（投资回报率=企业估值/注册资本）代表企业的（升值-贬值）情况，来对创新活力指数进行衡量；投资回报率>1为升值、投资回报率<1为贬值，以此对企业进行分类进行数据挖掘。

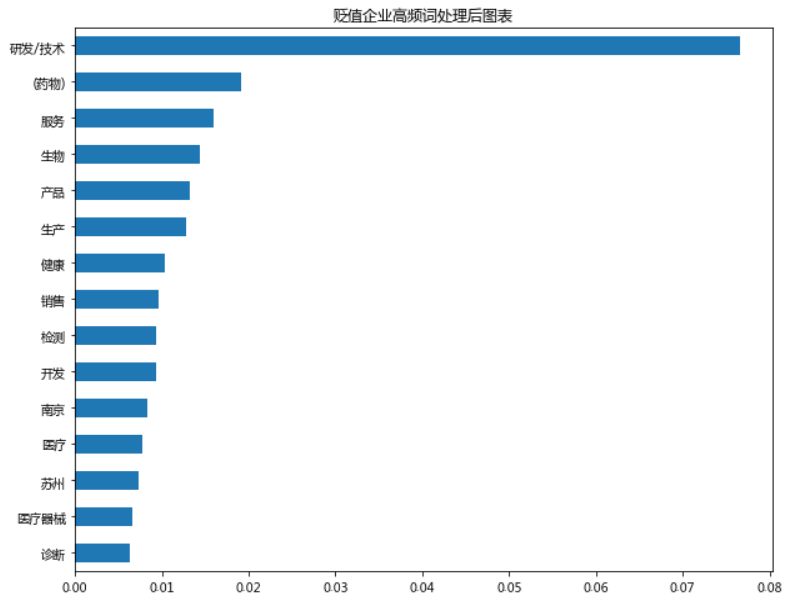
**第二部分 数据预处理**  
（1）数据预处理：  
在所有的3000多家企业中，一共有242家企业能直接求出投资回报率（即这242家企业的企业估值、注册资本都不为空），升值企业占比为0.7355。只利用242家企业的信息来分析是远远不够的，为了合理地进行数据扩充，提高数据利用率，本项目利用机器学习SVM支持向量机的文本信息分类方法。具体步骤如下：  
对“企业简介”进行分类（选择“企业简介”有两个原因：一方面数据表格中企业简介数据量大，另一方面企业简介能代表企业的重要特征）。以242家已求出回报率的企业作为训练集训练SVM模型，不断调整参数得到合适的模型，利用该模型对所有的企业进行预测，拟得到2301家企业的投资回报率（模型参数设置和训练过程详细说明在可视化报告中）。预测中升值企业占比为0.7701，与已有数据的升值企业占比0.7355相近，可以认为模型合理，效果较好。

（2）外部数据源：  
生物医药行业受政策的影响较大，新药的开发、药品的上市周期、财政补贴等都会对生物医药行业产生影响。鉴于本次大赛没有提供政策类数据，本课题进行了数据补充。  
数据来源：中国知网数据库政府文件子库。  
数据获取方式：研究成员在咨询从事信息检索与生物医药研究的专家后，确定了数据获取策略。首先通过阅读部分生物医药相关的政策文件，总结出关键词；然后通过关键词和地区名称约束，来检索江苏省以及各地级市的近20年发布的所有相关政策；利用爬虫软件爬取所有政策。

**第三部分 数据挖掘**

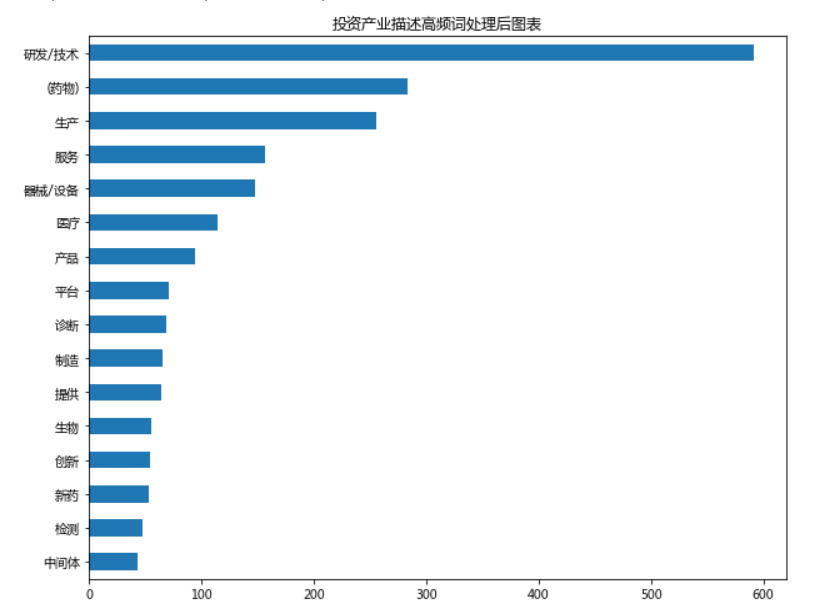
1. 企业简介的高频词分析  
   研究思路：企业创新活力除了与此前提到的指标相关外，还与企业的定位不可分割。如果某些词语在升值企业中占比高，但是在贬值企业中占比低，就可以看出什么因素是两类企业的差距所在。  
   研究方法：通过机器学习将企业分为升值企业/贬值企业后，提取出升值公司、贬值公司的企业简介高频词汇（如下图），观察分布规律。  
   研究结论：

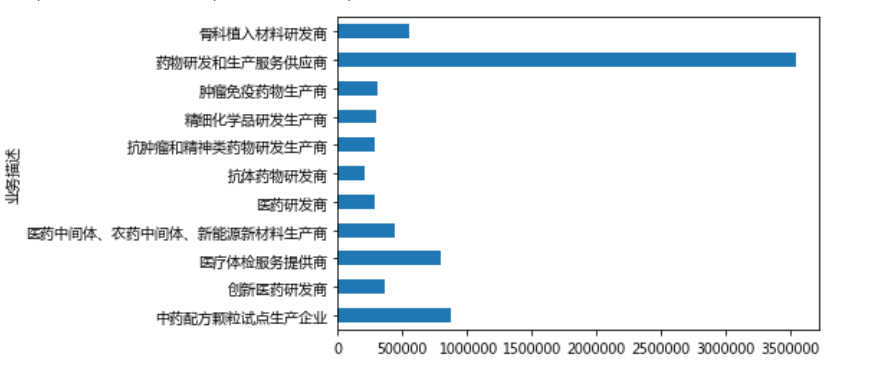




A. 无论升值还是贬值企业，“研发/技术”的占比都位居第一，并且贬值企业中“研发/技术”的占比要高于升值企业。  
原因探究：这一方面说明“研发/技术”在生物医药行业的重要性，这也与“研发/技术”类业务最吸引投资相吻合；另一方面说明生物医疗行业作为一个“高技术、高投入、高风险、高回报、产业周期长”的高精产业，“研发/技术”的成果转化是拉开企业差距的重要因素。如果倾向于做“研发/技术”的企业却出现成果转化不到位的问题，会极大地大地影响到企业的生存和发展能力。  
政策启示：  
a. 以市场为导向，建立企业为主体的技术转化模式。缺乏市场信息，只注重理论和技术水平的先进性，忽视市场需求，会导致研究开发的项目成果不被市场所接纳，生物医疗科技成果转化率偏低。  
b. 加大对生物医药技术转化资金的投入，拓宽技术转移的资金投入渠道，完善风险投资机制。国际上研究开发、中试、成果的商品化的经费投入一般比例是1:10:100，而我国的投入比例仅为1:1.1:1.5。我国科技成果转化资金企业自筹的占56％，国家科技拨款占26.8％，风险投资仅占2.3％。生物医药技术转化资金严重不足，导致多数生物医药技术成果都停留于实验室，被“束之高阁”。  
c. 建立产、学、研联合技术转移模式。  
d. 完善中介服务体系，搭建科技成果转化的桥梁。“产业双资平台”是提升生物医疗科技成果转化率的高效平台。通过整合吸纳产业领域内技术类企业、金融类企业入驻到平台，使企业间互通“资本+资源”有无，助力产业良性发展。

B. 升值企业中，“器械/设备”占比高；贬值企业中，“服务”占比高。  
原因探究：因为医疗器械涵盖了可直接或间接用于人体的仪器、设备、器具、体外诊断试剂及校准物、材料以及配套的计算机软件等，是集合了医药、机械、电子、塑料等多个行业的资金、技术密集型产业，现代医疗设备产品技术含量高，行业壁垒高，研发投入大，利润丰厚；另一方面，由前文可知，“服务”比“器械/设备”更吸引投资，但是投入产出却落后于“器械/设备”。  
政策启示：  
a.再次说明要重视研发的成果转化。器械/设备是技术研发的成果转化之一，是企业的技术壁垒与长远发展的重要条件。此外，对被投资次数多和投资数额高的企业的简介进行分词，统计结果如下图，投资频率最高的TOP5企业业务是：研发/技术、药物、生产、服务、器械/设备。投资数额的断层TOP1企业业务是：药物研究和生产服务供应商。可以发现“服务”高于“器械/设备”，这说明“服务”业务的成果转化率还需要加强。





b.资本投资与政策扶持可以适当向投资回报率更高的“器械/设备”倾斜。  
c.重视“基于器械的第三方服务”的发展。单纯的“医疗服务”没有技术门槛，可替代性高，发展竞争力不足。从卫计委对第三方诊断、影像与病理等的发文支持，以及大型医疗企业在这方面的布局与并购可以看出，基于医疗器械与设备的第三方诊断、影像、病理、物流、等新兴的集约化服务模式将有力地推动医疗行业的发展。由于这些领域同时也是资本密集型行业，未来将会有大量的投资及并购事件发生。企业需要意识到：为医院服务、与医院共赢的“互利模式”，会比抢医院生意的“颠覆模式”更符合目前的医疗商业环境。

C.升值企业的高频词后出现“苏州”，贬值企业的高频词中出现“南京”。这说明苏州地区的企业发展势头强于南京地区。在后文中企业活力的空间分布中再具体说明。

2. 五指标的k-s检验

研究思路：我们拟认为注册资本、商标数量、著作权数量、主要人员数、核心团队人数五个指标对企业的创新活力有影响。如果这五项数据和投资回报率的分类密切相关，可以说明它们与企业的活力有关，也反过来进一步说明我们的模型建立是合理有效的。  
研究方法：利用统计分析的k-s检验方法对这两类企业的五项数据分布情况分别进行差异性检验（如升值企业的注册资本与贬值企业的注册资本所有数据进行比较检验）；如果存在差异性，那么这个差异性是由分类结果导致的，说明分类因素（投资回报率）与所检验的数据是紧密相关的  
k-s检验的原假设H0：两类数据分布无差异，p<0.05时，拒绝原假设，即认为分布有差异，p<<5 认为分布差异性十分显著。

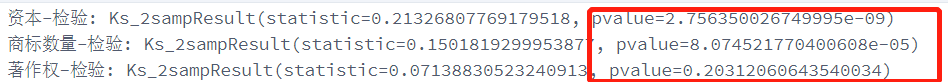


Image Name

检验结果如下图：除了著作权-检验的p值大于0.05,其他四个检验的p值都是10的负5次方以上，说明这四项数据分布

结果一：商标数量与回报率成正比。

原因探究：Bosworth和Rogers认为企业所拥有的独特资产是造成企业之间竞争优势和业绩差异的重要原因，这些独特资产主要包括技术诀窍、声誉、商标、专利等垄断性资源。作为重要的知识产权形式之一，商标权是企业防止其资产贬值的必要手段之一。  
关于商标价值的基本论点是，由于企业可以通过使用商标向消费者发出信号，表明某一产品有特定并恒定的质量，以此来减少消费者的“搜索成本”，因此商标有助于解决买方和卖方之间的信息不对称，从而弥补市场失灵。作为企业开拓市场的标志，商标是企业信誉和竞争优势的载体，而商品或服务的质量是商标信誉的基础。因此，商标在引导消费者认牌购物的同时，又鞭策、促进生产者或经营者为维护自己的商标信誉而努力提高产品或服务的质量。因此，商标作为企业知识产权的重要组成部分，对商标投入的力度必然会对企业的市场价值与经营绩效有一定的影响。  
此外，商标与高科技是带动企业发展的双轮，是企业发展进程中互为依托，相互促进、共同发展的不可分离的重要组成部分。首先，高科技是强大的生产力，当高科技转化为商品以后，只有及时注册商标，高科技的市场成果才能得到有效的转化。因此，这支持了我们的结论：商标数量与回报率成正比。  
政策建议：一方面，政府需要考虑如何完善商标权注册制度，完善注册流程，按照严格执行法律规定，切实保护好企业的商标权。另一方面，政府决策时需要考虑鼓励企业注册商标，培育更多的品牌产业，通过政策引导强化企业的知识产权观念。

结果二：注册资本与回报率成正比。  
原因探究：注册资本越大，信托公司的业务范围和业务规模也越大，资本回报率也越高。中小高新技术企业融资现状——企业规模较小、竞争力弱、管理水平不高；中小高新技术企业注册资本不多，企业规模偏小，自有资金不足以应对企业发展壮大的需要。与大型企业相比，市场竞争力明显不足，融资能力不强.此类企业在初创期对资金的需求量尤其大，回报率较低，难以吸引投资者进行投资。政策建议：加大对初创公司以及中小企业的财政支持力度，降息减税，降低融资难度；放宽招商引资限制，吸引更多投资进入。多举措高企业注册资本，改善资金流，提高盈利能力和回报率。

结果三：著作权数量对企业发展影响不显著。  
在k-s检验中，检验结果认为著作权数量与投资回报率不是紧密相关地，这是与实际违背的，可能是因为没有发挥著作权的作用，也即存在成果转化不到位的问题，这与上文分析的结果是一致的。以下给出对著作权的研究与分析。  
原因探究：首先看《著作权法》第十条规定：“著作权的内容包括人身权和财产权，二者又细分为：发表权、署名权、修改权和保护作品完整权；复制权、发行权、出租权、展览权、表演权、放映权、广播权、信息网络传播权、摄制权、改编权、翻译权和其他著作财产权”。著作权的对象是作者所创的文学、艺术和科学作品，是有形资产。其作品包括以下形式：文学作品；口述作品；音乐、戏剧、曲艺、舞蹈作品；美术、摄影作品；工程设计、产品设计图纸及其说明；地图、示意图等图形作品；计算机软件；法律、行政法规规定的其他作品。  
这说明著作权资产是无形资产，其价值评估受到多方因素的影响。结合江苏省的实际情况分析如下：  
A.经济和市场环境。一般而言，如果著作权交易发生在经济比较发达的地区，其价值会高于经济落后的地区。江苏省内各个城市的经济发展水平不一，受地域限制，一些较为落后的地区著作权的价值难以发挥。  
B.社会环境：社会中个体的著作权意识以及政府制定的政策是否对个体进行著作权意识的培养和引导都会影响著作权价值。通过分析江苏省针对生物医药行业知识产权的政策可知，知识产权保护侧重于药物和医疗器械研发专利，并没有专门针对工程设计图等形式的著作权。  
C.2019年2月，江苏省政府印发《关于推动生物医药产业高质量发展的意见》。《意见》明确要加强知识产权保护，营造良好创新环境。实行知识产权标准化管理，实现全省生物医药企业贯标全覆盖。建立生物医药产业知识产权维权援助专家组，支持企业积极应对涉外知识产权争端。加强抗肿瘤药、高端医疗器械、基因治疗药物等重点领域知识产权预警研究，指导企业和科研机构有效防范应对知识产权风险。  
D.作品的质量：不同类型作品的著作权其价值可能存在很大差别。创作投入、质量和寿命周期差异都会影响著作价值。因此，不考虑质量的的著作权数量对企业发展影响不大。  
E.作者的知名度：典型的例子就是名作家的知名度远高于新进作家，其作品的被期待程度也更高。著作价值需要企业的声誉背书，单纯的数量考量难以影响企业的回报率。  
F.收益难以量化衡量：著作权通常是通过销售或使用来获得收益。前者的收益情况比较直观，一般市场上有惯用的计算方法；而后者则通过使用作品实现收益，其过程中难以量化的因素比较多，因此可能使价值评估的过程更为复杂。因此，单纯的数量考量难以影企业的回报率。  
政策建议：确立科学完善的著作权资产价值评估的标准体系，通过政策引导企业对著作权保护的重视，并支持企业的创新成果转落地的定位不可分割。如果某些词语在升值企业中占比高，但是在贬值业中占比低，就可以看出什么因素是两类企业的差距所在。

结果四：主要员数、核心团队人数都对企业的创新活力有影响。升值企业的主要人员数、核心团队人数都少于贬值企业。

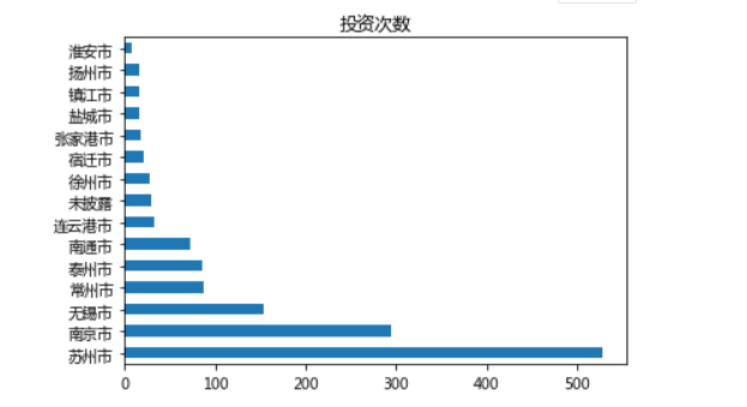


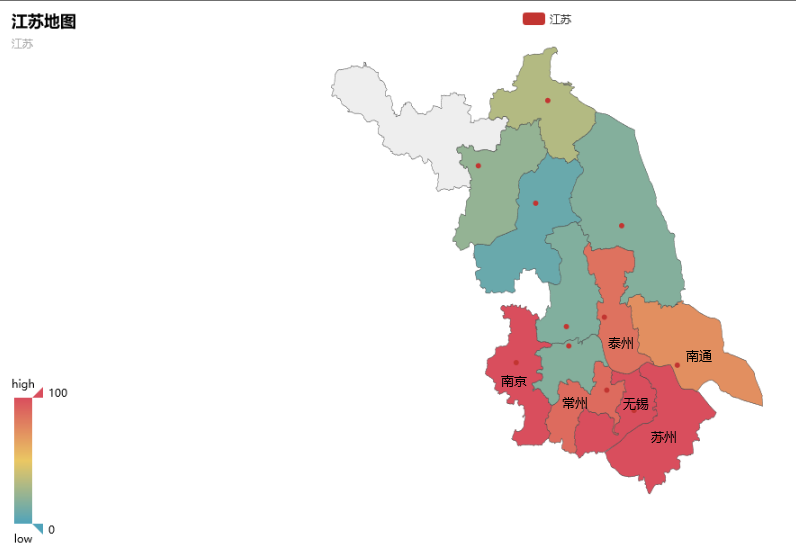
  
（第一张图是升值企业 第二张图是贬值企业）

原因探究：一个好的核心团队应当包含四类人：

A. Product Visionary：能洞悉市场需求的人；   
B. Designer：能设计出好用户体验的人；   
C. Hacker：能快速解决程式问题的人；   
D. Hustler：能快速解决一切其他问题的人。   
拥有这四种属性的团队可以是少于四个人的(一人多能)。因此，核心团队成员应是扁平化的，去层级的，互补的，精简的.此外，企业持续高增长的关键是组建核心团队，而整个公司会形成好多个团队，每个团队人数不能太多，核心团队也是如此。精简的团队利于高效运转。所以核心团队人数都少于贬值企业。  
企业战略建议：  
精简的团队利于高效运转。

1. 创新活力地域分布  
   研究思路：不同地区的企业发展现状不一。本部分是分析江苏省内企业创新活力的分布。  
   研究方法：分析被投资次数和升值贬值企业的地域分布，并通过相关“创新创业类政策”的政策变化来侧面解释不通地区企业的发展现状。  
   研究结果：  
   A. 投资次数地域分布

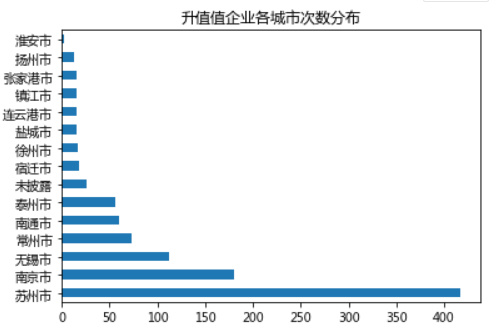


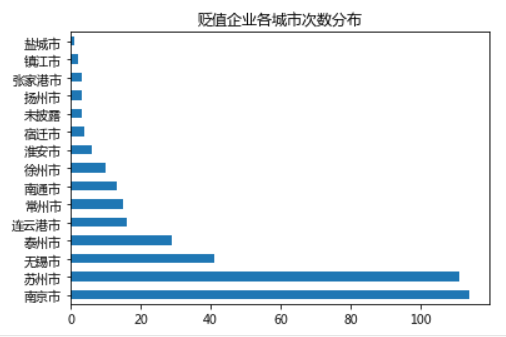


数据分析：仅从医药类企业投资次数来看，江苏省的各大城市间还是存在较大差异，结合江苏省地图做出的投资次数热力图，可以发现投资次数排在前列的苏州、南京、无锡、常州市、泰州市皆位于苏南地区，而苏北的城市投资明显落后于苏南。虽然江苏省已经出台了推动苏南经济发展的政策，但是两个区域的差异不是可以简单缩小的，从历史上来看，苏北的工业基础相对薄弱，多以第一产业为主，且由于深处腹地，公路不发达，河运又不太能跟上现代经济发展的节奏，交通的不便利导致较少有机会接收大城市的辐射，所以相对来说企业发展不到位，投资机会较少。而苏南经济发展体系完备，位。于黄金水道——长江三角洲，交通发达，且背靠上海等重要城市，联系紧密一旦有较为优势的医药类自然会吸引大量投资。对于医药企业这类技术密集型企业来说，投资对于创新研发至关重要，在保持苏南医药类企业飞速发展，且加快苏北同类企业发展势头的前提下进行如下措施。

政策建议：  
a. 加强省内合作  
加强苏南、苏北两个区域之间或省内各城市之间合作，可以利用苏南比较强的创新能力和开发能力针对性地帮助苏南和苏北相关医药类企业学习到更加先进的管理的方式或技术。同时苏北第一产业相对发达一些，具有更丰富的劳动力或生产资源，合作的过程，苏南可以引入劳动力及生产资源，这通过双向往来的方式促进江苏医药类企业稳步发展。  
b. 加强对外合作  
目前来说，苏南地理位置优越，交通便利城市建设，旅游环境、基础设施建设都比较优越，得到了外省企业或海外企业的青睐，在外资引进方面确实取得了非常丰硕的成果，苏北若在投资方面有一定发展的话，可以通过加强与苏南企业的联系，取得和省外及外资企业的联系，帮助海外优秀企业看到苏北企业的发展潜力来加强投资，促进苏北些医药企业的发展。  
c. 提升软硬环境建设  
投资方投资时，除了企业自身的实力之外，还会考虑当地的交通情况、城市建设情况、基础设施建设情况，以及政府办事效率，公共服务情况，苏北应加大相关行业的基础设施建设投入，促进城市文明建设和城市的其他方面的建设，同时，在公共服务和正式办政 府办事效率方面，政府也应当采取一定的措施来改善。提升政府的效率，可以更加倾向于引入数字化等等方面的技术和办法。  
d. 利用好区位优势  
苏北并不是完全缺少区位优势，连云港作为沿海的14个开放城市，具有非常好的政策便利条件和地理位置，可以以连云港作为一个中心带动周围苏北城市的一个发展，通过航运优势吸引省外的投资力量。

B.升值贬值地域分布

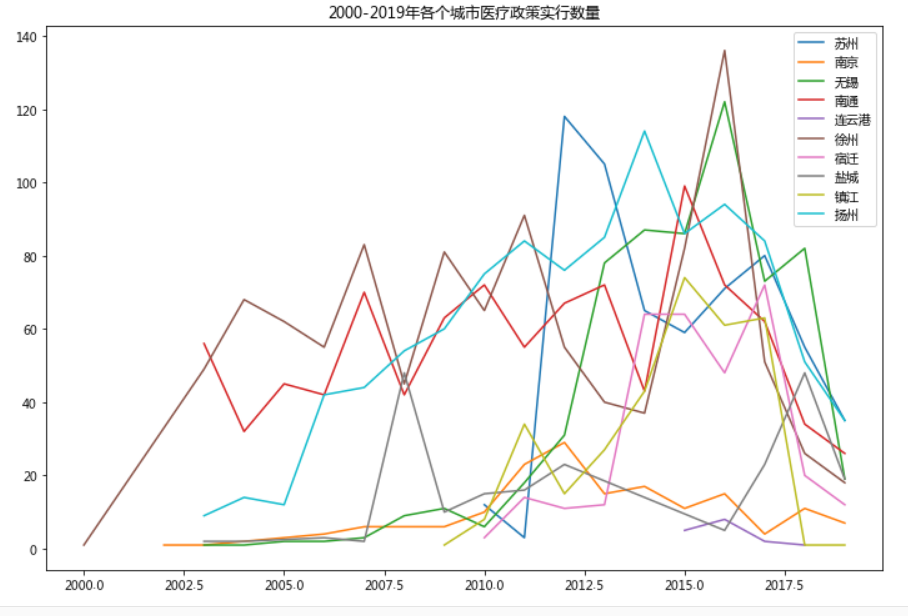


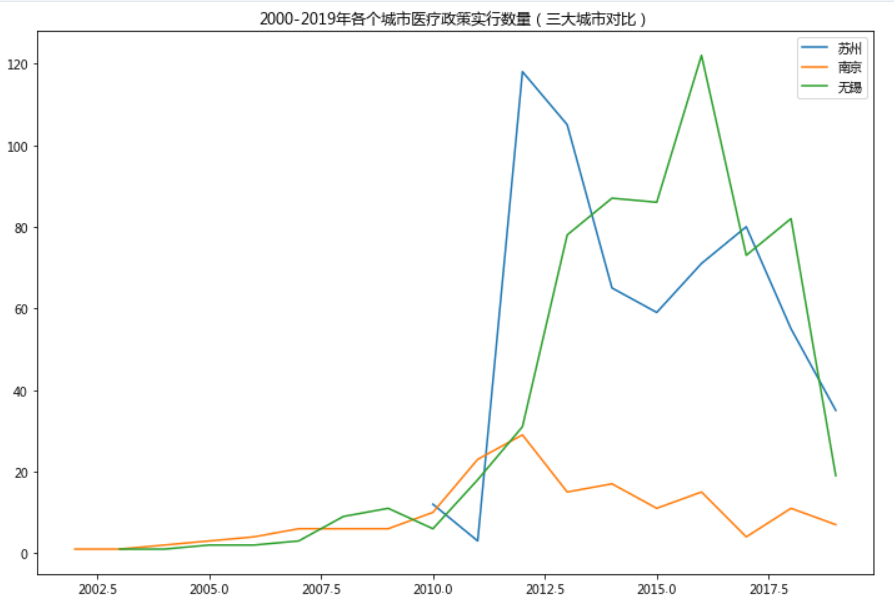


江苏省十四个城市每个城市企业增贬值次数在两个图中所处位置没有太大的差异，比如说苏州、南京、无锡、泰州等城市，他们的增值次数分布靠前，在贬值的企业分布里面，也同样靠前。这些城市的企业在增值数量较大的情况下，贬值数量也不少，而苏北的城市，无论是在升值和贬值的情况下，表现都很平稳，可以说明苏南的一些城市，对于医药类企业创新活力关注度更高，也更愿意去尝试，虽然目前可能还存在医药行业体系不够完善，波动、成果转化度不够等问题但总体是发展的状态，然而苏南的城市活力较低，关注度不够。

目前江苏省已形成“四大医药产业基地”:泰州医药城、连云港新医药产业基地、南京“药谷”、苏锡常医药产业群。 而南京苏州无锡常州泰州和连云港六个城市已经成为江苏医药产业重要聚集区，形成鲜明特色的几个区域医药板块，但仍存在一些问题阻碍发展。  
政策建议：  
a.促进医药产业集聚发展  
江苏目前有各种各样的医药产业园，也形成了一定的规模和产业链，但是各园区内的企业、创新平台、服务机构等之间并没有进行 充分协调合作，孤立大于合作，没有产生协同效应，缺少有效的内部交流与合作， 无法真正发挥创新的效用。因此继续突出个产业园区发展重点，加强园区和周边的小园区的融合和合作，每个大园区基地间形成差异化发展。  
b.加强校企合作，提升成果转化能力  
企业目前的创新转化能力不够，仍然受到投资、资金、人才等的影响难以自食其力，因此依托高校科技和人才优势以及大学周边地域势，逐步形成有特色的技术创新和技术转移平台，推动企业、高校、科研院所共同建设研究与开发机构、实验基地等，由企业提供研究经费，高校和科研院所提供人才和技术，并吸收企业高级技术人才共同参与研发工作。

C. 政策地域分布



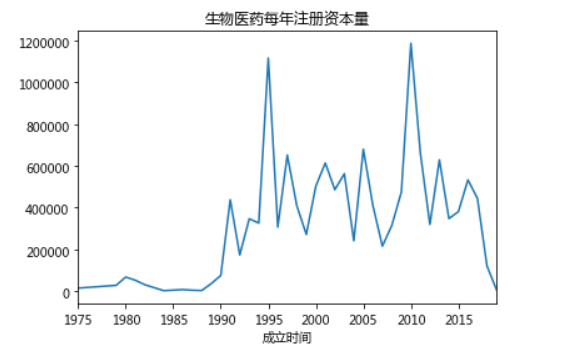


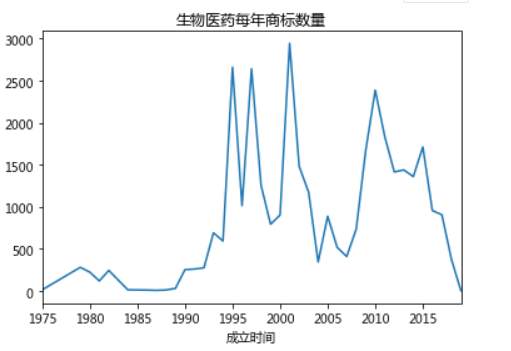
数据分析：结合江苏省的投资次数热力图，选取苏州、无锡、南京三个城市，对比发现近20年实施的生物医疗类政策，苏州远多于南京。这与前文分析的升值企业的高频词为“苏州”，贬值企业的高频词为“南京”相呼应。

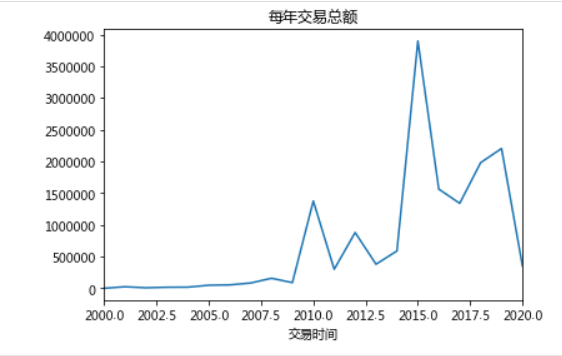
政策建议：南京市政府可以在能力范围内加大对生物医疗行业的政策支持力度。

（第一张图为全省的政策分布 第二张图为前三名的政策分布）

D. 时间走势及政策预测









数据分析：生物医药每年注册资本量自1990年起开始呈现上升状态，且在之后多出现起伏状态，这和我国对于生物医药的关注有必然联系。我国生物医药产业起始于20世纪90年代，正是注册资本开始上升的时候，国家对生物医药的产业政策支持则发端于 2006年颁布的《国家中长期科学和技术发展规划纲要》 之后，国家科技 重大专项“重大新药创制”正式启动，2012 年颁布《生物产业发展规划》 正式将生物医药产业纳入国家发展战略，标志生物医药产业进入政策受惠期。2005年后无论是注册资本量还是商标数量都有一个大幅度上升，2012年前后皆有一个小幅度的回升，在这个时期国家政策利好，促进了生物医药行业的发展。而在这段时期内，江苏省也出台了《江苏省医药产业调整振兴规划纲要 (2009-2012)年)》来优化医药产业布局，集中建设特色园区。

江苏省生物医药行业于2015年交易额达到顶峰，而那一年正式生物医药界并购很强劲的一年，但之后确实一路下降，自2018年以来无论是交易时间还是交易次数更是再次大幅下降，生物技术股高估值提高了大规模并购的风险、具有较高不确定性等等都是影响并购下降的重要因素。  
政策预测：  
在2020年，我国医药制造业收到新冠疫情爆发影响，医药企业不同程度延迟开工，停工停产、运输受阻、物流管制与人员交通限制等情况对药企的经营带来了明显影响。尤其对于一些实力不强的企业来说，短期内甚至将面临资金链断裂和倒闭的风险，所以在短期内会面临注册资本、注册商标数量以及交易额等大幅下降的趋势。但是长期来看，新冠疫情引起了国家以及不同行业对生物医药的重视和兴趣，未来生物有巨大的发展空间。包括国家有关部门要求积极支持企业复工复产和新项目开工建设，对复工复产企业和新建项目涉及防治新冠肺炎的专利、商标申请注册开辟“绿色通道”，加快办理效率。所以长远来看生物医药是具有极大发展潜力的行业。