# **我国中央政策发布对国内创业趋势的影响分析**

## **一、研究方法**

通过审慎的筛选和梳理，确定研究采用的数据，数据包含125规划全文、135规划全文、145规划全文及2010-2020中央创业政策，其中，中央创业政策为自建的政策库。

在国务院官网搜索标题中包含“创业”词条关键词的文件，时间范围为2010年至2020年。经过整理，形成了所使用的政策库初稿。为增强梳理工作的严谨性，删除与创业相关性较弱的政策文件，最终形成了本次探究的资料库，共277条政策。

作为本次探究的重点，对“青年”相关的政策进行单独分析。以政策库为基础，筛选出和青年密切相关的创业扶持政策。筛选标准如下：（1）在梳理政策库阶段，整理了关键词一列，在此列中查找和青年相关的内容；（2）查找政策标题，看标题里所指明的对象是否大量包含了青年这一群体，关键词例如“大学生”、“毕业生”、“农村青年”；（3）看政策的内容，在网页中打开该政策文件，在网页上查找青年相关关键词。通过筛选，进一步整理出与“青年”相关的创业政策65条。

另外，单独选取了云南地区的各项数据进行分析，验证资料库及各规划推断出来的结论。

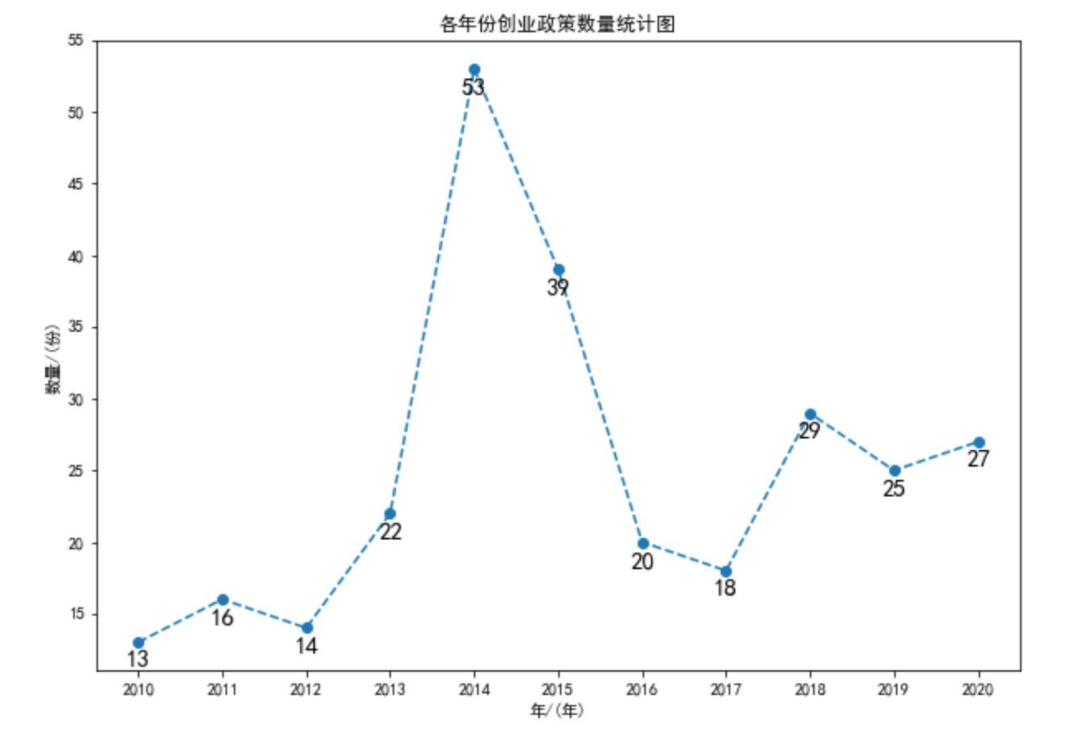
本次研究在文本分析阶段主要以中央层面的部门政策文件为研究对象，没有覆盖地方政府层面的创业扶持政策。地方层面主要分析数据变化印证政策发布对地方经济产生的影响。

## **二、创业政策数据分析**

在完成对整体创业政策库的收集汇总后，从出台时间及政策内容等方面进行一系列数据统计及文本分析。

### 1.总体呈两阶段特征**【该部分主要是基于政策库中数据数量的宏观分析】**

使用python的pandas库和matplotlib库将“创业政策数据集.xlsx”中的政策以发布年为标准做出各年份创业政策的数量统计图。



**（图一）**

由图可见，创业促进政策的时间分布明显的呈现出两段式特征，以2014年为界，在此之前，创业相关政策发布频率较低，2014-2020年七年间共出台212条政策，两年政策出台数占七年政策出台数的近44%，2014年起政策呈现井喷式爆发。

经过各规划文本及政策库分析可得政策峰值前的阶段主要关注创业教育，以创业促进就业、推动创业教育为主，创业政策尚处于起步阶段。在2014年9月夏季达沃斯论坛上，李克强总理发出“大众创业、万众创新”的号召,在“大众创业万众创新”思想指导下，2014年政策频出，之后的政策体系逐渐趋于完善。（关于各阶段的政策侧重点会在后文详细分析）

**2.以125、135、145规划发布时间为大致节点的创业政策数据分析【该部分主要是基于125规划、135规划、145规划文本本身的分析】**

（1）125规划期间前后（双创前）以创业促进就业为政策导向，探索和初步建立体系（2010 - 2015）【创业教育及基础建设】

使用python的jieba库先对125、135、145规划文本进行预处理，将各规划进行分词、统计词频、排序，并将词频为20以上、三个文本中均有出现且不具备特殊领域特征的词、量词、动词提取出来作为停用词表（停用词.txt在文件夹中可见），然后对125规划文本分词、删除停用词并生成多个新文本文件，类型分别为分词+词频版本和仅分词版本（处理文本的代码源程序及文本文件均在文件夹中）。

125规划前后，政策名包含关键词“大学生”“青年”“少年”“教育”的政策有54条，2010-2020包含此类关键词的政策有65条，占近83%。因此可以推断，125规划期间创业相关政策的主要关注点在于“创业教育”，主要面向青年，以创业促进就业，推动创业教育为主，但政策覆盖不够全面，在当时创业政策体系初步建立、尚未健全的情况下，以青年为受众、以就业为导向的政策体系探索仍能够为创业政策体系的健全完善奠定坚实的基础。此前，125规划的文本分析处理中（在前面生成各规划仅分词的文本文件的基础上，将仅在125规划中出现的词语提取出来并将这些新词按词频从大到小排列（该代码源程序见文件夹））关于新词词频前十位依次是中小企业、事业单位、职能、五年、创新型、保障性、逐步提高、经济区、城市化、产业结构，基本没有特殊的具有特定领域代表性的词条的出现，符合推断。

（2）135规划期间前后政策完善（2016 – 2020.1.7） 【农村创业及政策完善】

135规划期间，包含关键词“农村”“农业”“下乡”的政策有21条，2011-2020包含此类关键词的政策有39条，占比超过50%；因此可以推断，125规划期间创业相关政策的主要关注点在于“农村创业”，主要面向农村地区，以创业促进就业，改善农村经济情况为主，政策体系已经逐渐趋于健全。此前，135规划的文本分析处理中（在前面生成各规划仅分词的文本文件的基础上，将仅在135规划中出现的词语提取出来并将这些新词按词频从大到小排列（该代码源程序见文件夹））关于新词词频前十位依次是贫困地区、行动计划、权力、海上、免费、技术装备、贫困人口、动力、毕业生、沿线，“贫困地区”“贫困人口”等新词的出现，基本符合在此期间国家政策关注点主要聚焦农村创业就业的推断。

1. 疫情促进政策升级（2020.1.7- 2020.12.24）【小微企业复工复产及毕业生创业就业】

2020年初，一场突如其来的疫情给中国经济造成冲击，小微企业生存、高校毕业生和农民工就业创业等面临严峻的考验，如何应对疫情带来的政治经济等各方面影响逐步成为国家、政府、人民关注的焦点，疫情期间的创业政策关注点主要有以下特征：

一是，加大了对一些行业复工复产的支持力度，该期间包含关键词“中小企业”的政策有5条，2010-2020包含此类关键词的政策有8条，占近62% 。

二是，疫情对产业发展既是挑战也是机遇。一些传统行业受冲击较大，而智能制造、数字化等新兴产业展现出强大成长潜力。

（4）145规划展望主要发展产业【科技创新产业】

由于疫情引发的智能制造数字化热潮持续上升，根据145规划新词词频统计，（在前面生成各规划仅分词的文本文件的基础上，将仅在125规划中出现的词语提取出来并将这些新词按词频从大到小排列（该代码源程序见文件夹）），排名前11的关键词依次是高质量、数字化、时代、万公顷、供应链、群体、粤港澳、民营企业、电子、医院、量子。因此可见，科技创新驱动会成为未来五年的主要发展指标，国家全力迈向创新驱动，建设数字中国。

### 3.小结**【该部分主要是基于政策库中具体政策文本本身的分析】**

**（**1）通过python的wordcloud库和matplotlib库，基于各规划的分词+词频文本，并提取出词频大于5的词语，可绘制125规划、135规划、145规划的词云图（词云图代码源程序见文件夹）。从各规划词云图像来看，我国在2010年-2020年期间三大规划重点基本相同，说明这10年间，我国各方面状况相对稳定，其次，经济词条始终最大，说明我国以经济建设为中心，但推测各时期侧重点可能有所不同。

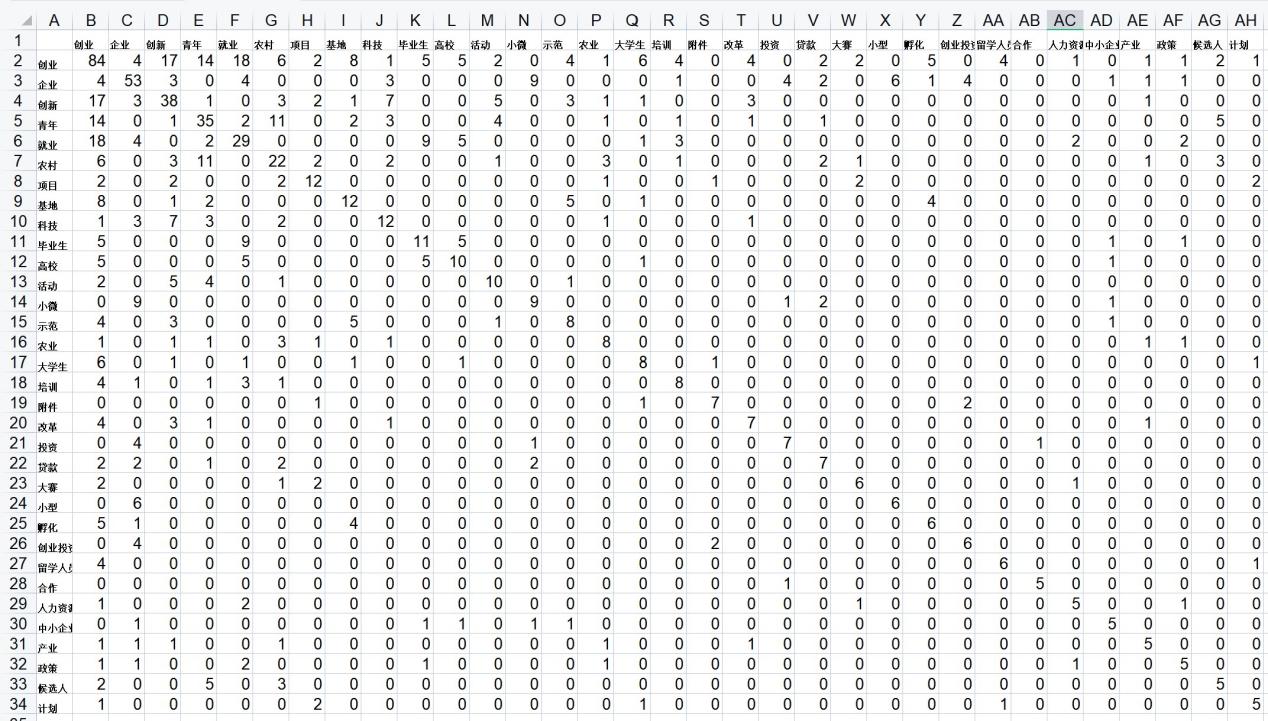
  

**125规划词云 135规划词云 145规划词云**

在2010-2020期间创业、创新、青年（大学生）一直是创业政策的焦点。从各部委不同时期政策文件的关键词来看，（利用python中的pandas库、jieba库、requests库、bs4库、re库、json库、os库，陆续访问政策库中的链接并分别对文本进行分词和词频统计的处理，每个政策文本选择词频最高的三个词语，最后按2010年-2015年、2016年-2020年、2019年的方式划分成三大块统计不同时期政策中词频高的词语，发现各时期侧重点有所不同，且与三大规划文本分析的结果高度契合，125规划期间政策关注点排名前10的依次是青年、创业、企业、农村、创新、就业、科技、大学生、小微、全国；135规划前后（双创提出后）政策关注点排名前10的依次是创业、创新、企业、就业、毕业生、农村、农业、项目、示范、合作；疫情期间则为创业、企业、创新、就业、中小企业、科技、培训、农业、中国证券监督管理委员会、示范。可以看出农村地区和毕业生群体一直是国家创业政策关注的主要点，但近年来的政策关注点也有所变化，125规划期间（双创前）政策的重点是以创业促进就业，关注点的实质在于创业教育；135规划期间，随着“大众创业、万众创新”的提出和写入政府工作报告，政策关注点也更多地落实到创业创新和企业相关政策，且侧重点在于农村创业，2019年由于疫情原因，关注点转向中小企业的生存和发展；145规划中，则强调了发展高新产业的重要性。

值得注意的是，疫情期间，面对政治经济等各方面的困境，创业政策也呈现出了鲜明的时期特征。复工复产与中小企业的生存成为国家和政策的主要关注点，各部委陆续推出了相关政策。

同时，为进一步分析创业政策文本中高频词汇间的联系，采用了构建共现矩阵，绘制基于创业政策文本分析的网络图谱来考察政策内容间的相关性。在政策库277份政策文本中，每篇文章摘取3个高频词汇作为关键词，继而筛选出出现频次高于5的关键词，统计得出34\*34的共词矩阵，并将该矩阵使用python的Counter库、collections库、numpy库、xlrd库、xlwt库将34个高频词构成的共词矩阵导入excel表格中方便观测结果。可以看到，以“创业”为核心节点, 与其联系最为紧密的是就业、创新、青年，关联频率分别为18、17、14;以“就业”为核心点, 与其联系最为紧密的是创业、毕业生、高校, 关联频率分别为18、9、5;以“创新”为核心点, 与其联系最为紧密的是创业、科技、活动, 关联频率分别为17、7、5；由创业政策的矩阵绘制的网络图谱主要以创业、创新为核心，以青年、农村、就业等为主要节点。



## **地方数据分析（以云南为例）**

云南省47209277的人口数，只占全国的3.34%，

位列全国第11位。

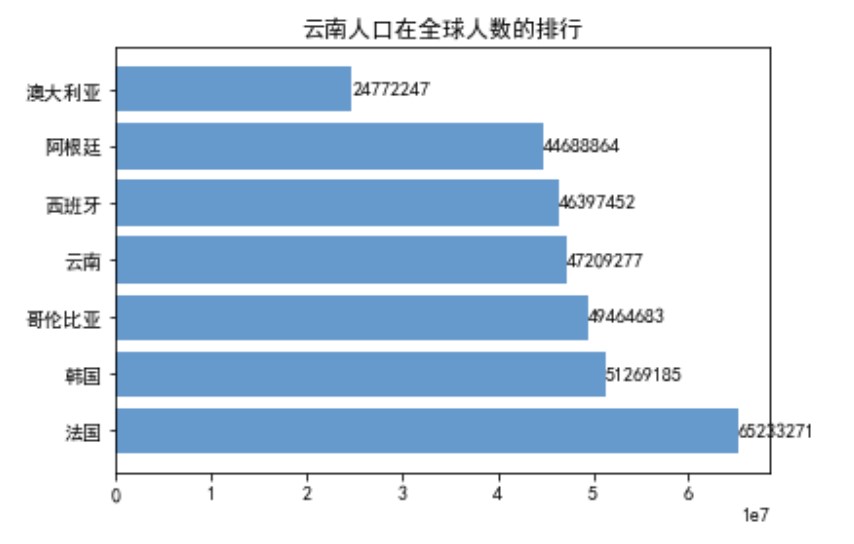
放眼世界，

单单比人数的话，

云南的人口数与西班牙相当，

可以排到全球人口排行榜的30位左右，

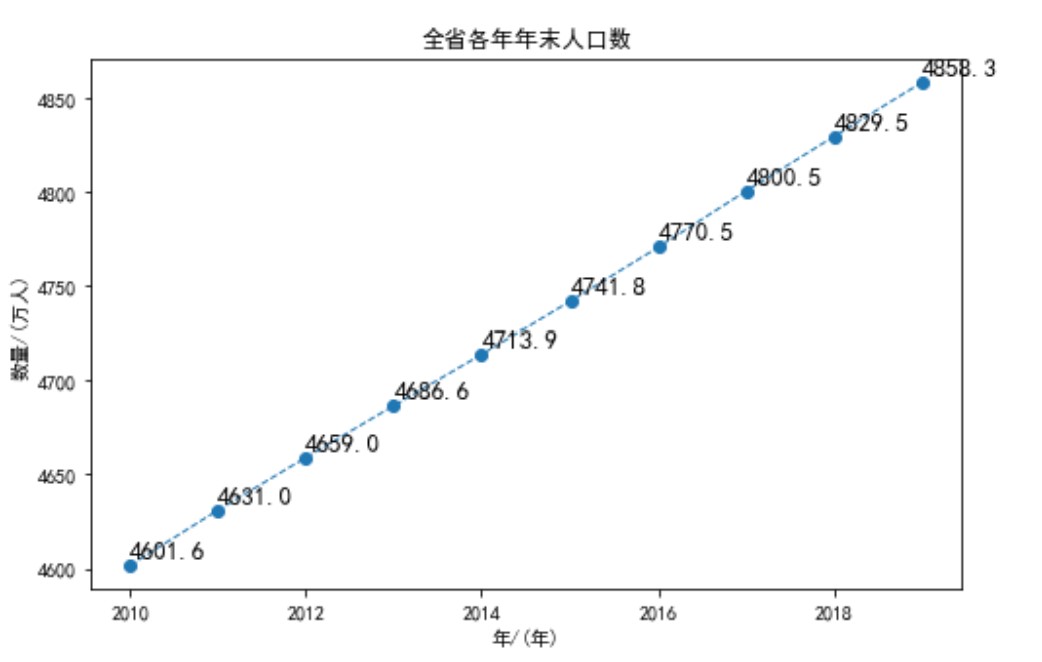
超过阿根廷、澳大利亚等85%的国家。



从2010年到2020年，

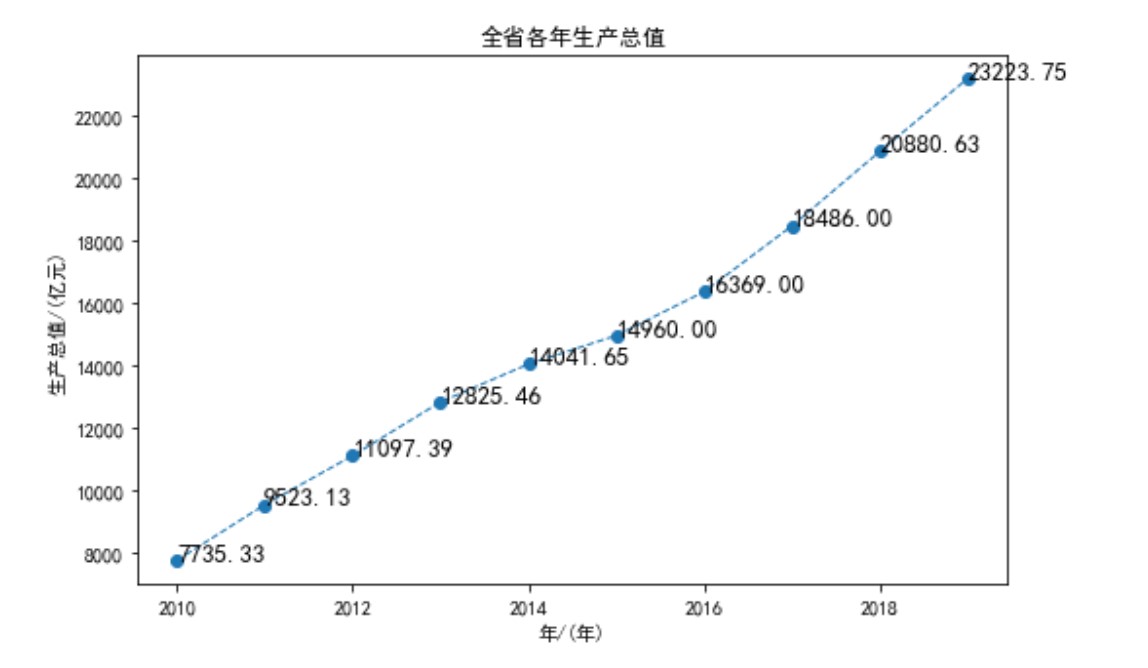
云南人口数不断攀升，

年均增长率达到了0.85，



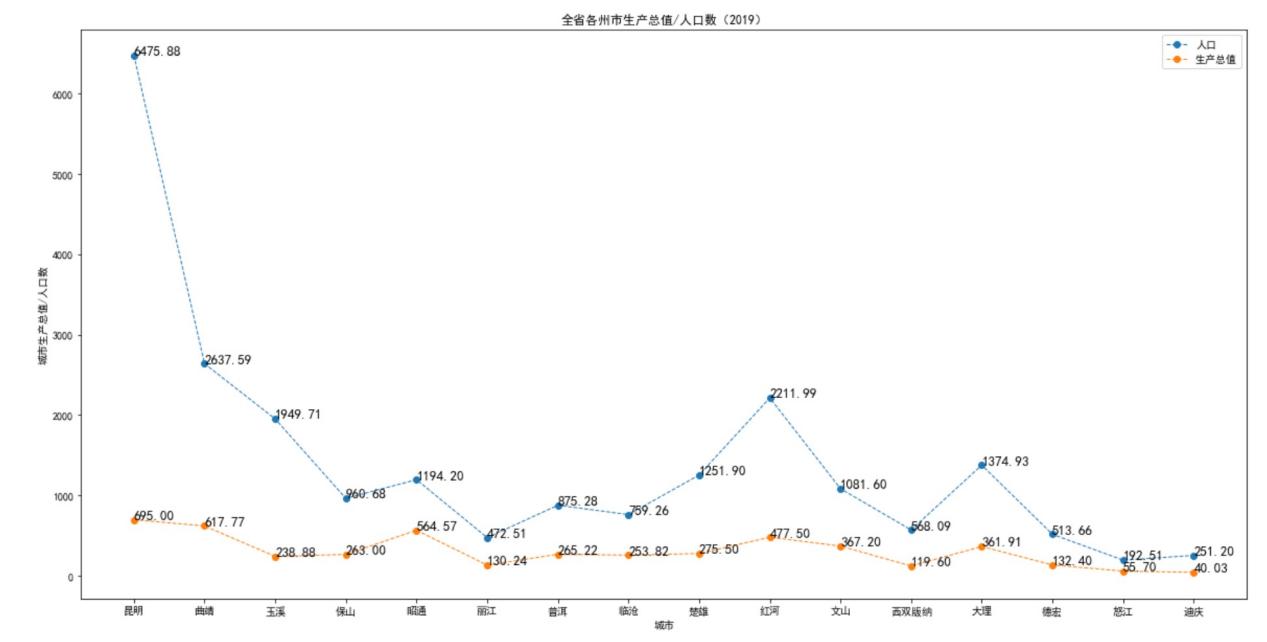
通过观察全省各年生产总值，

全省生产总值也一路攀升，



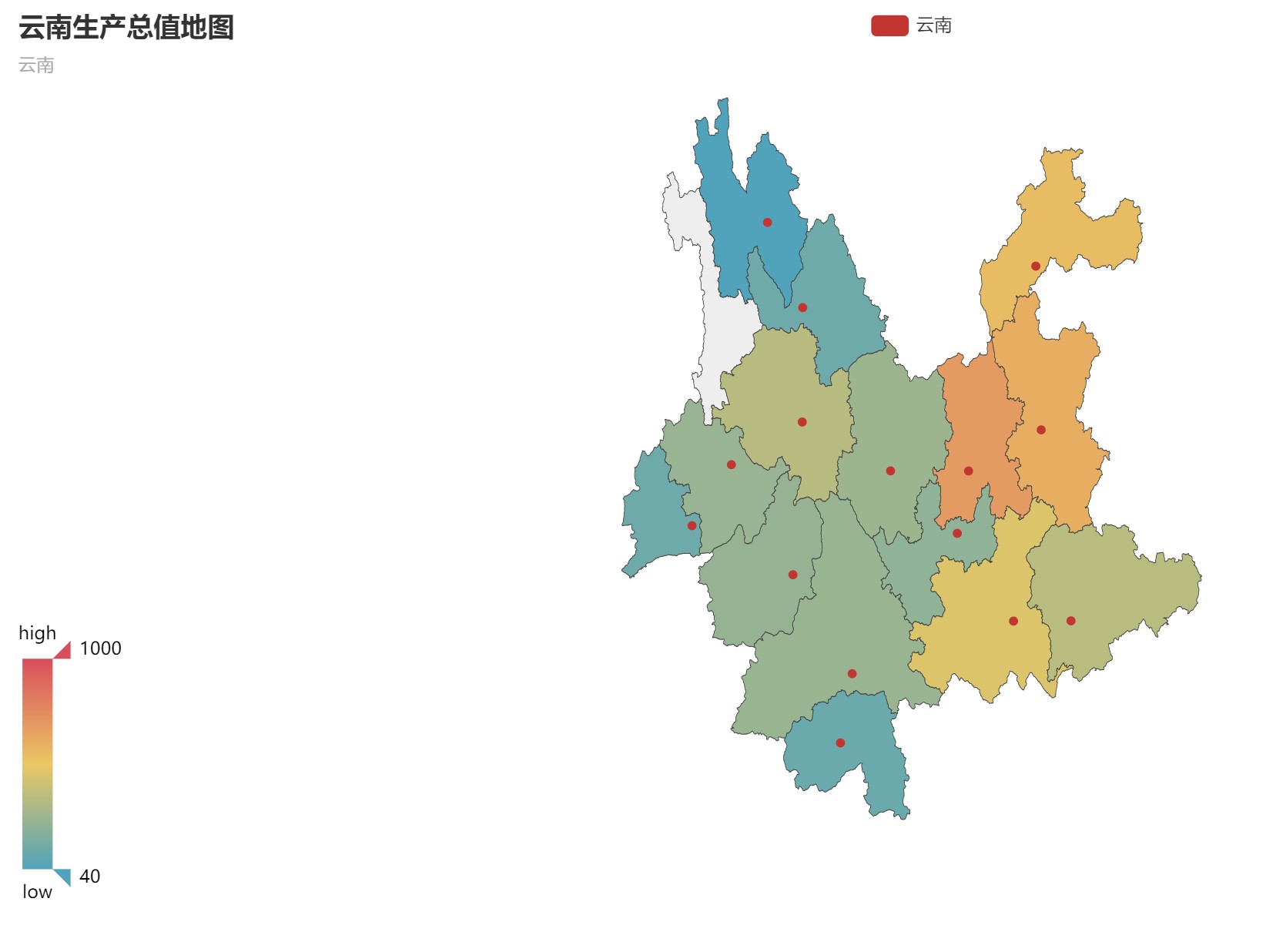
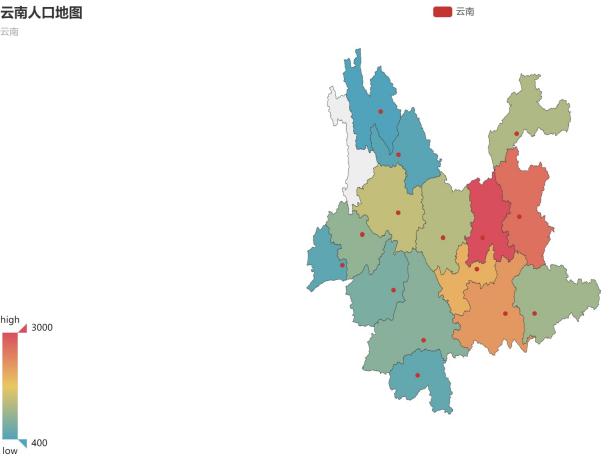
可粗略得到，人口增加有效推动了经济社会的蓬勃发展。

从2019年云南省各州市的人口数和生产总值来看，



整体呈现”人口数越高，生产总值越高“的趋势，

从各州市人口地图和各州市生产总值地图可以看到，



人口主要聚集在以云南省省会昆明为中心的附近州市，

生产总值的高低也基本和人口聚集趋势相同。

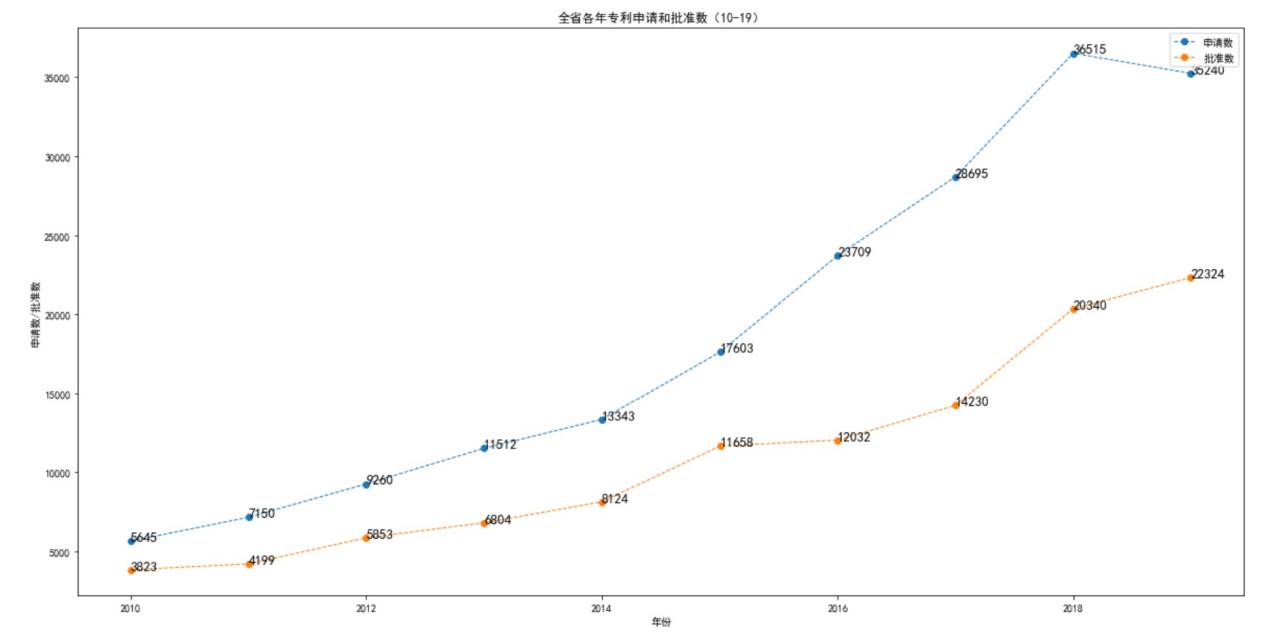
昆明不负省会之名，

人口和生产总值均比其他州市遥遥领先，

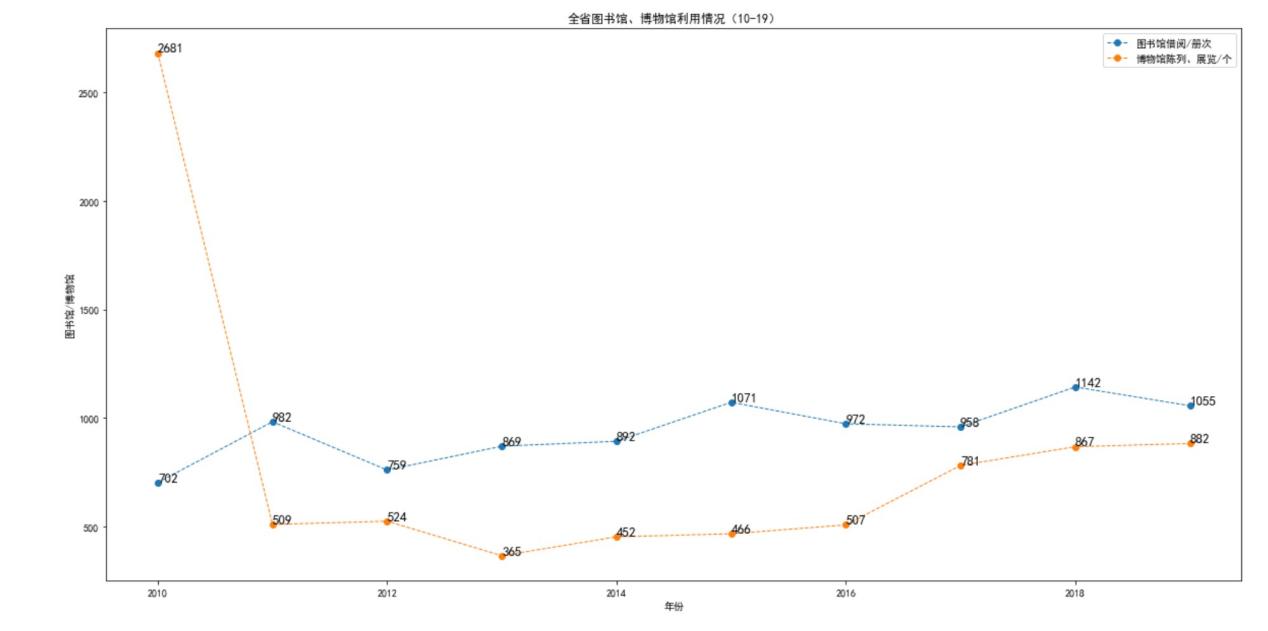
完美的诠释了人口发展的聚集效应和区位优势。

在第二部分基于各规划词频统计的分析中，我们的推测是125期间侧重主要是教育和基础建设，135期间侧重解决农村就业创业问题，并首次提出了网络空间、电子商务等概念，145规划则强调了高新产业的发展。

针对云南省统计局给出的各项数据，我们以专利申请数/批准数，作为衡量云南省的科技发展的指标。



由图可得，云南省专利申请/批准数逐年上升，比较明显的是2014-2018年期间斜率骤升，与此同时，根据图书馆/博物馆使用情况来看：



除去2010年博物馆的异常数据，其他数据趋势与云南省专利申请/批准数类似，也是2014-2015年，2016-2017年期间斜率骤升。

根据云南省国民生产总值和第一/第二/第三产业的年增加值，经过大体估算第一产业（农业）在135规划期间增长率高于其他时期；因而预测，第三产业（高新科技）145规划期间增长率将会大幅提升。

以上，可验证，由125、135、145规划中得出的推测，基本符合现实情况。