

基于大数据的“高铁经济”策略研究

——以江苏淮安为例

李文慧

淮阴师范学院城市与环境学院, 江苏 淮安
Email: liwenhui@hytc.edu.cn

收稿日期: 2021年5月17日; 录用日期: 2021年6月8日; 发布日期: 2021年6月18日

摘要

2019年底淮安市正式接入高铁路网, 高铁为淮安的经济再次腾飞赋能。本文使用大数据技术, 基于互联网中公开信息, 分析挖掘与淮安市相近的城市在开通高铁前后的举措, 分析相关数据, 进而给出建议措施, 为淮安经济发展提供数据支撑和决策参考。

关键词

高铁, 区域经济, 大数据

Research on the Strategy of “High-Speed Railway Economy” Based on Big Data

—Taking Huai’an, Jiangsu Province as an Example

Wenhui Li

College of Urban and Environmental Sciences, Huaiyin Normal University, Huai’an Jiangsu
Email: liwenhui@hytc.edu.cn

Received: May 17th, 2021; accepted: Jun. 8th, 2021; published: Jun. 18th, 2021

Abstract

At the end of 2019, Huai’an City was officially connected to the high-speed railway network, which enabled the economy of Huai’an to take off again. This paper uses big data technology, based on the open information on the Internet, analyzes and excavates the measures of cities similar to Huai’an before and after the opening of high-speed railway, analyzes the relevant data, and then

puts forward suggestions and measures to provide data support and decision-making reference for the economic development of Huai'an.

Keywords

High-Speed Railway, Regional Economy, Big Data

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 背景

高铁是一个相对的概念,由于世界各国(地区)的高铁的技术研发水平和发展历程不尽相同,到目前为止,关于高铁的定义尚未统一。纵观高铁发展史,可以发现高铁的开通对一个城市或地区的发展具有重要影响,我国将高铁的建设归纳入国家建设的重要部分,目前已经初步形成“四纵四横”的快速客运网络,高铁对我国经济社会的发展起着很重要的作用[1]。我国高铁建设虽然起步相对较晚,但是发展迅速[2]。我国《中长期铁路网规划(2016版)》指出[3]:到2020年,一批重大标志性项目建成投产,铁路网规模达到15万公里,其中高速铁路3万公里,覆盖80%以上的大城市。

根据现有数据,高铁沿线城市的GDP指标随着高速铁路建成到运营都有不同程度的上涨。如京沪高铁对沿线城市经济增长的影响显著为正,其直接贡献度达到了17%[4]。高铁的建成到运营对区域经济起着很重要的促进作用,这就侧面体现了高铁投资的收益效果,也是高铁对区域经济发展带来的直接经济成果。

对于接入高铁路网的淮安(图1),如何充分利用高铁的优势来发展地方经济?以前的高铁新通城市如何实现抓住机遇,采取了哪些有效措施,有什么经验教训?为进一步探讨上述问题,本文以“高铁经济”为研究核心,基于Python使用数据挖掘方法找出相关城市的各种措施方法,分析高铁为地方经济带来的各种影响。在上述研究的基础上,结合淮安实际情况,提出因地制宜的策略措施建议,为淮安经济的发展做出参考和借鉴。



Figure 1. Huai'an rice-shaped high-speed railway network

图1. 淮安米字型高铁网络

2. 基于 Python 的数据抓取与分析

本论文主要用于为淮安市迎接高铁发展通道提供支持，在已通高铁的相关城市中，按照经济规模、城市人口、省内地位等因素选择确定了徐州、常州、镇江、淄博、德州、枣庄、沧州、金华、湖州、随州、孝感、芜湖共计 12 个城市，用这些城市的相关数据来分析高铁开通前后的相关数据与措施。

2.1. 基于 Python 的数据获取

Python 是一种解释型脚本语言，自从 20 世纪 90 年代初 Python 语言诞生至今，它已被逐渐广泛应用于系统管理任务的处理和 Web 编程，排名常年位居 TIOBE 排行榜前列，2020 年连续三年保持第三。

本文使用其核心模块 `urllib` 从中国政府网开始，抓取相关外链，通过 `urllib.request.urlopen` 方法获取页面数据后，使用正则表达式识别出非本网站下属的链接，若该网址尚未保存则写入 Access 数据库中。抓取完本页后，使用该方法读取已抓取数据库中未访问的网址，进一步挖掘。为避免无穷抓取，选择抓取 `gov.cn` 结尾的政府网址(18,134 个) [5]。

2.2. 数据获取

该功能模块中，调用 `requests` 库和 `re` 库 `findall` 函数，根据网址正则表达式匹配来挖掘相关页面，挖掘出所有本网站网址的下属链接及其内容，存储到数据库中，挖掘流程如图 2。

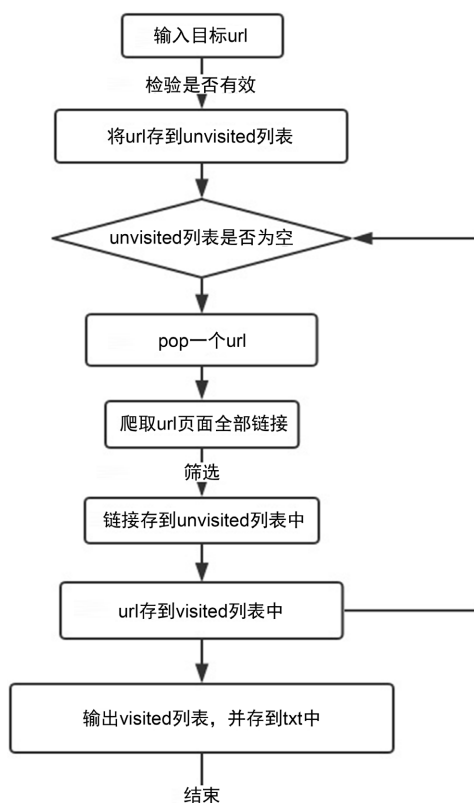


Figure 2. Data acquisition process for all pages of a single website

图 2. 单个网站所有页面数据获取流程

2.3. 语义解译与分词

自然语言分词(Chinese Word Segmentation)指的是将一个汉字序列切分成一个个单独的词，分词就是

将连续的字序列按照一定的规范重新组合成词序列的过程。在英文的行文中，单词之间是以空格作为自然分界符的，而中文只是字、句和段能通过明显的分界符来简单划界，唯独词没有一个形式上的分界符，虽然英文也同样存在短语的划分问题，不过在分词处理中，中文比之英文要复杂、困难得多。

经济词语较多，本文使用 jieba 分词，首先去除上述 14 个城市之外的信息，然后添加自定义的“高铁经济”关键词进行分词，进而将相关词语出现次数之和大于 5 次，则该文与高铁经济紧密相关，从而实现文本聚类。

3. 高铁对现有相关城市的经济影响

选中的 12 个城市在高铁开通前后各两年的 GDP 数据见图 3，从图中看出，德州在接入高铁之后，GDP 有着非常明显的提升。2010 年，德州实现地区生产总值(GDP) 1657.82 亿元，按可比价格计算，比上年增长 12.9%。其中第一产业增加值 210.51 亿元，增长 3.2%；第二产业增加值 899.55 亿元，增长 14.3%；第三产业增加值 547.76 亿元，增长 13.8%。实现旅游总收入 45.44 亿元，增长 29.4% (德州市人民政府 2010 年政府工作报告)。此后两年，GDP 增长分别为 11.7%、12.1%。而安徽省芜湖市，在 2015 年实现地区生产总值 2457.32 亿元，比上年增长 10.3%。其中，第一产业增加值 120.02 亿元，增长 4.3%；第二产业增加值 1540.60 亿元，增长 10.3%；第三产业增加值 796.70 亿元，增长 11.0%。实现旅游业总收入 388.70 亿元，比去年增长 23.88%。

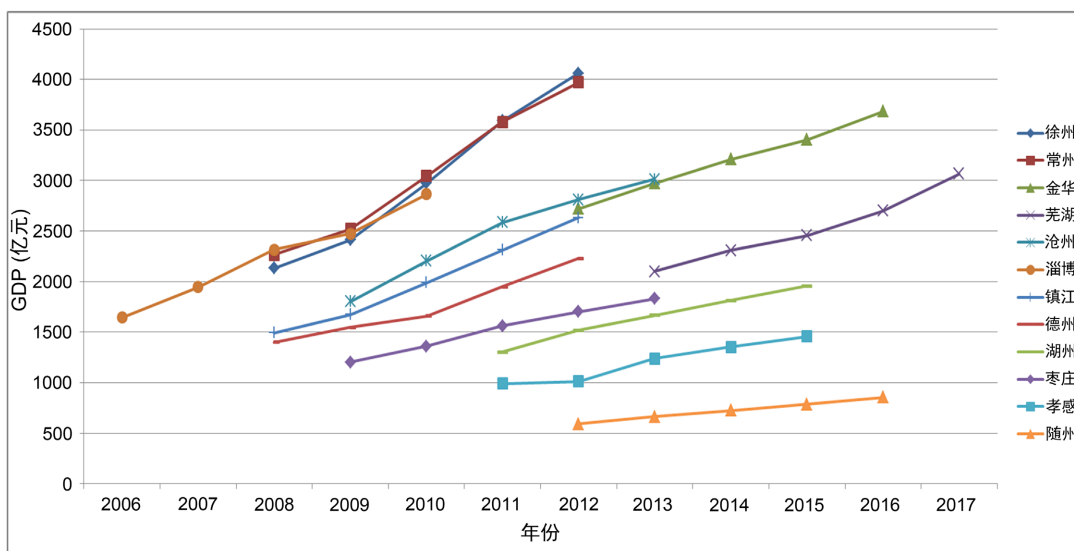


Figure 3. GDP of each city before and after the opening of high-speed railway

图 3. 高铁开通前后两年，各城市 GDP

从图 4 可以发现，我省城市在接入高铁网络之后 GDP 持续保持高速增长，尤其是徐州在接入高铁网络之后的 GDP 增速明显加快，作为华东重要门户城市的徐州，因其地理位置及其重要，是国家综合交通枢纽，徐州东站以每天到发 271 趟高铁车次，位于全国第 10。因此，借着京沪高铁的东风，依靠大量的客流，进一步带动了徐州的发展。高铁开通以来，徐州商贸业快速发展，社会消费品零售总额 3 年增速均居江苏省前列。2014 年，徐州新兴产业实现产值 4300 亿元，规模比高铁开通前翻了两番多，占规模以上工业比重达到 37.7%，提高了 21.1 个百分点。“徐州要以开发区速度对接高铁速度，将着力在旅游业、商贸物流业和新兴产业上下功夫，不断把高铁溢出效应转化为发展动力，推动徐州产业转型升级，积极呼应和全面融入沿线地区对内对外开放的大潮流”。

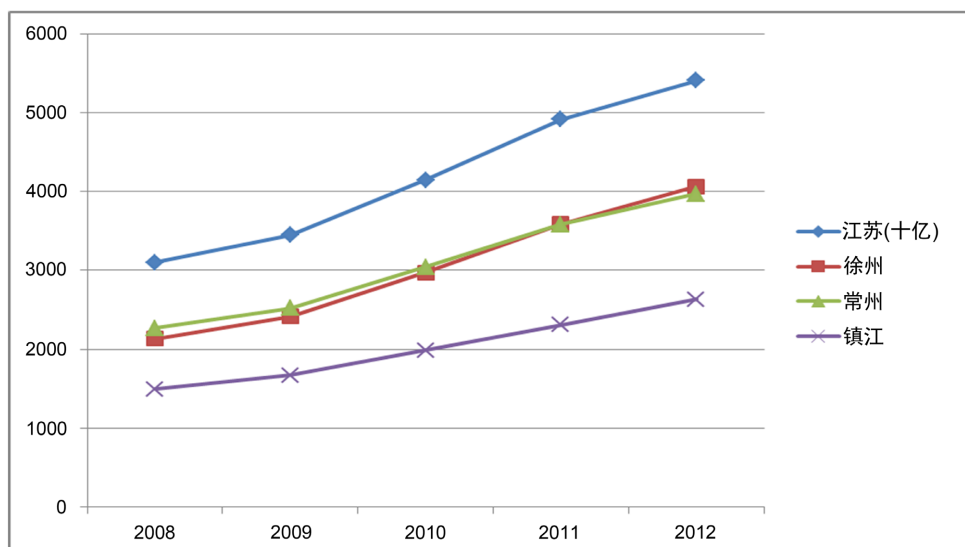


Figure 4. GDP of Jiangsu Province and related cities

图 4. 江苏省与相关城市的 GDP

高铁的通车共享了区域的旅游资源，伴随时间成本的降低，长途游览转变为短途游览，短途游览缩短为同城游览当前旅游者的出行频率得到了急速提高(图 5)，乘坐高铁的商务出行乘客百分比要比普通铁路高很多(例如：长吉线高出 17%，京沪线高出 11%)，并且商务出行者所占比例很高(例如：长吉线 40%，京沪线 67%)，高铁沿线的主要景点和旅游次数快速增长(曲阜的游客量在高铁开通后得到增长)，再以浙江为例，除群岛舟山之外全部通车高铁，利用先天优势以高铁为出行手段形成旅游圈大力发展旅游业，2017 年金华市旅游人口同比 2010 年增长 55%，湖州市旅游人口增长 62%，七年间年均 GDP 增长率分别为 9.2%、8.3%。

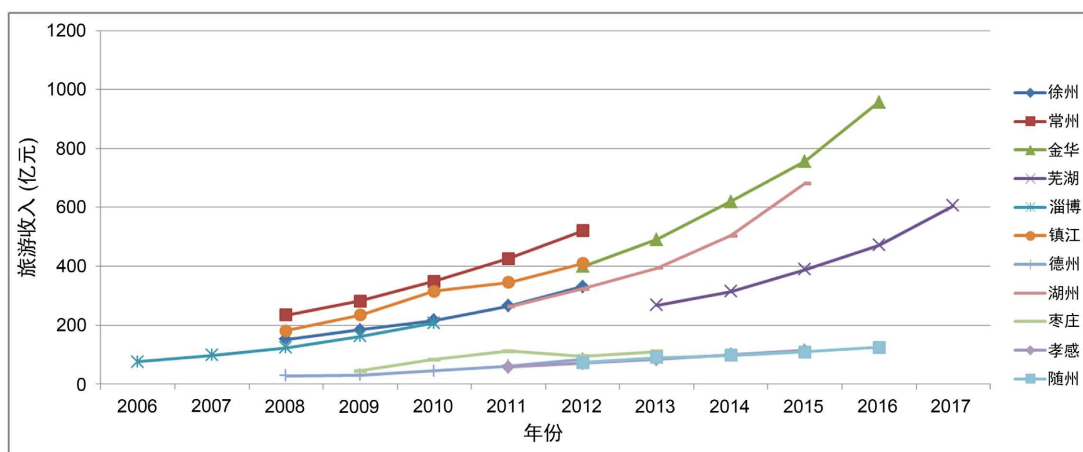


Figure 5. Tourism revenue of each city before and after the opening of high-speed railway

图 5. 高铁开通前后两年的各市旅游收入

从京沪高铁开通来看，引起的交通可达性提升对于沿线城市旅游业发展产生了显著的促进作用。京沪高铁开通所引起的城市可达性每提高 1%，沿线城市旅游客流约增加 0.53% 到 0.62%。

双重差分法，英文名 Differences-in-Differences (DID)，是一种普遍使用的政策效应评估方法。本文中，为满足 DID 模型数据“平行趋势假设”需求，选择样本数据在开通前后 GDP 相差较小的 11 个城市

数据来进行分析，其中：沧州、德州、芜湖、枣庄、徐州、常州、枣庄、湖州、镇江 9 个较早开通高铁的城市作为实验组，淄博、淮安作为控制组，选取实验组开通前两年、开通后一年数据，控制组相近三年数据(表 1)。

Table 1. DID model description statistics

表 1. DID 模型描述统计

	Before	After	汇总
Control (控制组)	10	5	15
Treated (实验组)	16	8	24
汇总	26	13	39

根据双重差分(DID)计量模型结果(表 2)，通高铁前 Diff 效应值为 1300.323 > 0 且呈现出 1%水平的显著性，表明着实验前实验组效应值明显高于控制组效应值。而通高铁后 Diff 效应值为 1485.938 > 0 且呈现出 1%水平的显著性，意味着实验后实验组效应值明显高于控制组效应值。进一步说明了高铁对地方经济短期内具有很好的促进作用，而长期效应中一定程度上拉大了一二线城市的差距[6]。

Table 2. Summary of DID model results

表 2. DID 模型结果汇总

Time	项	效应值 GDP	标准误	<i>t</i>	<i>p</i>
Before	Control (控制组)	43.131			
	Treated (实验组)	1343.454			
	Diff (T-C)	1300.323	297.029	4.378	0.000**
After	Control (控制组)	318.227			
	Treated (实验组)	1804.165			
	Diff (T-C)	1485.938	348.297	4.266	0.000**
	Diff-in-Diff	185.615	315.047	0.589	0.560

R^2 : 0.480, 调整 R^2 : 0.418; * $p < 0.05$, ** $p < 0.01$ 。

4. 机遇与挑战

高铁的开通，对地方经济有着较为明显的提升作用，是地方发展经济的重要机遇。

4.1. 区域经济发展的机遇

高铁的建成缓解了春运的压力，提升运力后大大增加了外来务工家庭的幸福感[7]。高铁的大范围通车带动二线楼市，一小时生活圈让强二线的经济发展起来，二线城市依附一线城市获得其剩余资源使二线的城市的曝光率得以大大提高[4]。

高铁的成功提高了通车城市的科技与就业。以河北为例，唐山、沧州为河北省重工业发展中心，自 2010 年通车起，沧州市 GDP 快速增长年均增速为 10%，重工业比重显著增加，工业增速为 5.2%，这也是得益于高铁的建成，智力人才的等各种生产要素的流入[8]。

随一个城市高铁的通车，必然会造成高铁站附近城区的翻新，高铁新城的建成、高铁新区的发展。以山东枣庄为例，2011 年接入第一条高铁，时至 2018 年 GDP 的 2402.38 亿元，同比 2010 年(1362.04 亿元)增长 76.38%。

4.2. 随之而来的挑战

在国际上,日本新干线带来的负面影响表现为小城市人口稀少、青壮劳动力短缺、消费低迷、社会凋敝。而在国内,学术界、政界、媒体等都对高铁“虹吸效应”进行过关注和讨论。如铜陵市,合福高铁、宁安城铁开通后,受“3小时经济圈”内上海、南京、杭州、武汉、合肥等发达城市影响,一些优势企业将总部迁入周边大城市的可能性和意愿增强,导致铜陵铜产业等优势产业外移。

5. 总结与建议

曾经大运河的上的“九省通衢”之地,“南船北马”之乡,淮安地处江苏长江以北核心地区,坐落于古淮河与京杭大运河交点,为南下北上的交通要道。根据国家《中长期铁路网规划》(2016~2030),京沪高铁“第二通道”的线路确定为北京-天津-东营-潍坊-临沂-淮安-扬州-南通-上海高速铁路,宁淮铁路与沂淮铁路建设已列入江苏省“十三五”铁路发展规划。实现与京沪高铁在南京的交汇,并联通宁杭高铁、宁合高铁等,进一步建成南京都市圈与京津冀连接的“中轴式”纵向通道,大大缩短辽东半岛乃至整个东北地区南下东南沿海的铁路里程,必将成为中国东部地区又一条新的重要的经济大动脉。

相对于三四线城市来说,高铁的建设与通车无疑是一次机会也是一种挑战,这是时代赋予的一柄双刃剑。部分城市抓住机遇直面挑战,实现城市的迅猛发展,经济迅速跻身二线城市的行列,也不乏对部分城市造成了人才流失城市人口老龄化的结果。而当今时代是互联网的时代,是大数据的时代,互联网为我们研究制定城市发展提供了便利的条件,多角度多方向准确的数据让政府的决策更精准更切合城市未来的发展。

基金项目

江苏高校哲学社会科学研究一般项目资助(编号:2019SJA1620)。

参考文献

- [1] 陈海友. 高铁在经济社会发展中的重要作用[J]. 财经界, 2020(19): 19-21.
- [2] 邹德慈. “高铁时代”城市与交通规划聚焦[J]. 城市交通, 2010(3): 1.
- [3] 伊然. 国务院通过中长期铁路网规划 五大举措促进铁路交通建设[J]. 工程机械, 2016(8): 61.
- [4] 杨晓智. 交通基础设施的增长效应和溢出效应分析——以京沪高铁为例[J]. 天津商务职业学院学报, 2017(1): 3-11.
- [5] 杨斌, 李文慧. 基于 Python 的涉台大数据获取与处理[J]. 计算机科学与应用, 2019, 9(1): 63-69.
<https://doi.org/10.12677/CSA.2019.91008>
- [6] 方大春, 孙明月. 高速铁路对长三角城市群经济发展影响评估——基于 DID 模型的实证研究[J]. 华东经济管理, 2016, 30(2): 42-47.
- [7] 国资小新. 大数据告诉你: 中国最强高铁站原来在这里[Z/OL].
<https://3g.china.com/act/news/945/20170526/30575604.html>, 2017-05-26.
- [8] 周伟. 京津冀协同发展视角下产业转型升级探讨——以河北沧州市为例[J]. 商业经济研究, 2016(9): 195-196.