

基于多元回归分析对新冠肺炎疫情传播因素的研究

钟思琪 赵梓希 杨鑫 韩佳成
莫那什大学信息技术学院 墨尔本维多利亚州 3168

摘要: 2019 年开始出现一种新型冠状病毒肺炎席卷全球, 由于公众认识不足, 也缺乏追踪感染链, 对此我们以数据为基础, 通过多元回归分析, 得出政策、疫苗、经济和人口密度等因素对新型冠状病毒肺炎疫情影响, 使政府能够更好地控制和管理 COVID-19 疫情, 减少感染人数, 最终降低由于 COVID-19 的传播带来的损失。

关键词: 多元回归分析; 新冠肺炎疫情; 传播; 影响因素

Research on transmission factors of the novel coronavirus outbreak based on multiple regression analysis

Siqi Zhong Zixi Zhao Xin Yang Jiacheng Han
School of Information Technology, Monash University, Melbourne Victoria 3168

Abstract: Since 2019, a novel coronavirus pneumonia has swept the world. Due to the lack of public awareness and the lack of tracking of infection chains, we have based on data and conducted multiple regression analysis to determine the impact of policies, vaccines, economy and population density on the novel coronavirus pneumonia epidemic, enabling the government to better control and manage the COVID-19 epidemic. Reducing the number of people infected and ultimately reducing the cost of COVID-19 transmission.

Keywords: Multiple regression analysis; The COVID-19 epidemic; Spread; Influencing factor

世界正面临着一种前所未有的传染病传播挑战, 这就是新型冠状病毒肺炎 (COVID-19), 这种疾病由新型冠状病毒 SARS-CoV-2 引起, 已被世界卫生组织确定为全球大流行疾病, 造成全球约 7.6 亿人感染, 687 万人死亡。然而, 目前公众对 COVID-19 的认识不足, 也缺乏追踪感染链的方法。对此我们基于多元回归分析政策、疫苗、经济和人口密度等因素与确诊病例数的相关系数得出该因素对新冠肺炎疫情的影响, 了解 COVID-19 的传播特征, 控制传播速度, 使政府能够更好地控制和管理 COVID-19 疫情。

我们以印度新冠肺炎疫情数据为基础, 使用 python 的 pandas 进行数据预处理, 移除缺失值并将离散数据进行分组, 利用 tableau 和 matplotlib 进行数据可视化, 由此我们得出以下宏观因素对新冠肺炎疫情的影响较为显著

一、封锁政策对新冠肺炎疫情的影响。

我们以印度的比哈尔邦、卡纳塔克邦和西孟加拉邦三个邦作为代表, 展示封锁政策对新冠肺炎疫情的影响。

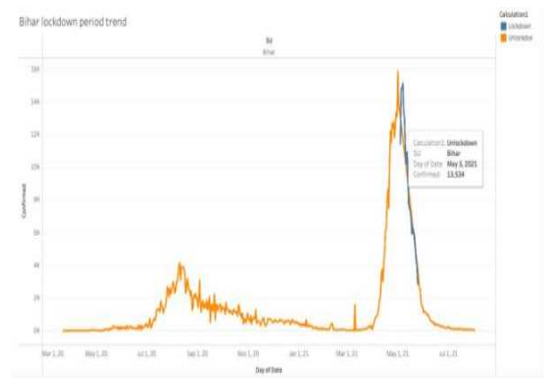


图 1 比哈尔邦 20 日 3 月 1 日至 21 日 7 月 1 日的每日病例变化情况

从图 2 上看出第二波新冠肺炎疫情始于 4 月份的印度, 在接近峰值时, 政府决定执行策略。我们可以看到一个明显的趋势, 几天后, 我们可以看到急剧下降, 最终在封锁后下降到相对稳定的水平。

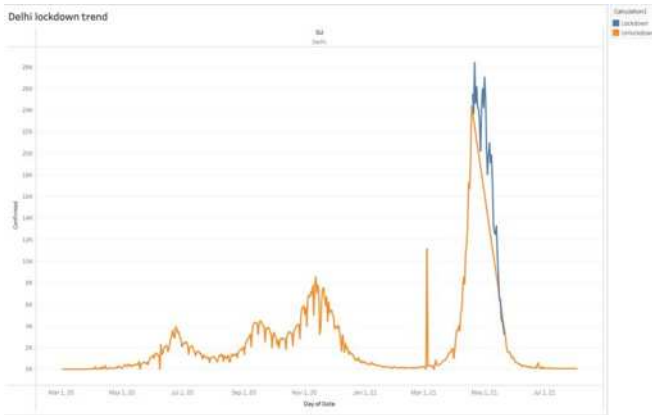


图2 卡纳塔克邦从20日3月1日至21日7月1日的每日病例变化

从图3是上可以看出西孟加拉邦政府宣布从5月16日至5月30日全面封锁，从图上可以看出在封锁期间出现了一个抛物线趋势，这意味着封锁政策并没有立即发挥作用。

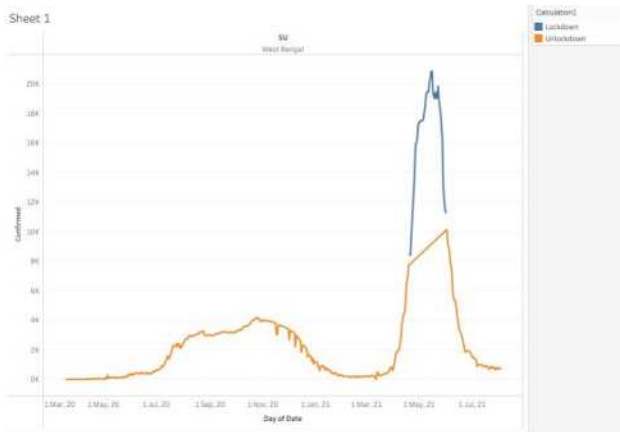


图3 西孟加拉邦20日3月1日至21日7月1日的每日病例变化情况

结果表明，封锁政策在 covid-19 病例预防中发挥了重要作用。它会在趋势上表现为，锁定政策启动后，可能不会立即下降，每天的病例数趋势会持续上升一段时间，然后开始下降，这是因为在封锁之后，一些已经被感染的人仍然会住在那里，需要被发现，但在达到峰值后，我们可以看到每天的病例数大幅下降，因此，封锁政策是一项有助于改善局势、减缓病毒传播的政策。

二、疫苗接种对 Covid-19 病例的影响

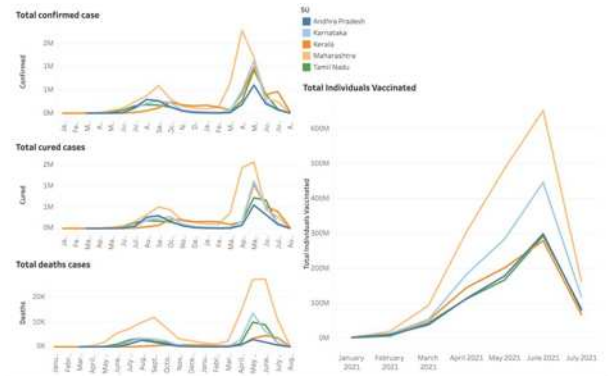


图4 印度确诊总病例数、治愈总病例数和死亡总病例数以及接种次数的趋势图

从图4可以看出，印度确诊总病例数、治愈总病例数和死亡总病例数的趋势相似。从2020年1月到2021年1月，印度确诊病例和治愈病例总数增长到5万多，然后缓慢下降。总死亡人数增加到1万人，然后慢慢下降。然而，从2021年1月开始，印度的确诊病例总数、治愈病例总数和死亡病例总数急剧增加，直到2021年5月达到峰值，然后开始急剧下降。继续观察印度的疫苗接种变化曲线图，我们可以看到，从2021年1月开始，每月的疫苗接种次数开始稳步增加，并在2021年6月达到峰值，之后印度每月的疫苗接种次数开始稳步下降。因此我们得出接种疫苗与确诊病例总数的减少呈正相关，疫苗可以有效减少病毒的传播，减少死亡人数，对 Covid-19 治愈人数有积极影响。

三、经济对 Covid-19 病例的影响

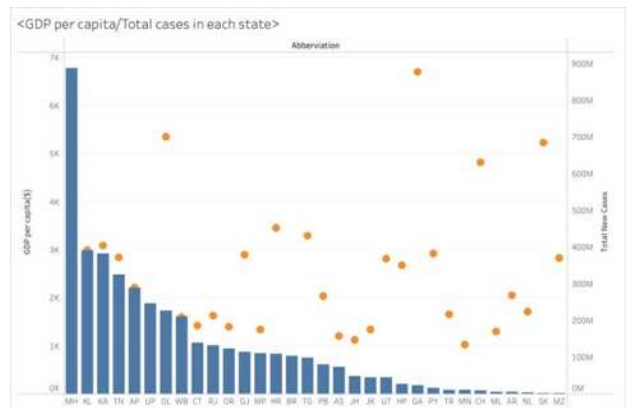


图5 印度各州人均 GDP 和病例总数图

图5的x轴上所有的州(缩写)，在y轴上，有两个变量，左边的一个是人均GDP(以美元计);右边的是这段时间内的病例总数。从图上，虽然我们年不出两个变量之间的关系或趋势，但在控制其他变量，然后再做分析。

根据数据集, 我们知道: 恰蒂斯加尔邦的人口密度为 189, 疫苗 7307 678, 病例总数为 138 447 101。哈里亚纳邦人口密度为 573 人, 疫苗 7948436 人, 病例总数为 109,108,040 人。从数据来看, 疫苗数量接近, 而哈里亚纳邦的密度要高得多, 而结果是恰蒂斯加尔邦的病例比哈里亚纳邦多。但从经济因素来看, 恰蒂斯加尔邦的人均 GDP 为 1416 美元, 哈里亚纳邦的人均 GDP 为 3451 美元。哈里亚纳邦的人均 GDP 比恰蒂斯加尔邦高得多, 这意味着哈里亚纳邦比恰蒂斯加尔邦富裕得多, 因此我们认为较富裕或富裕地区可能有较好的自我保护条件, 感染人数较低。

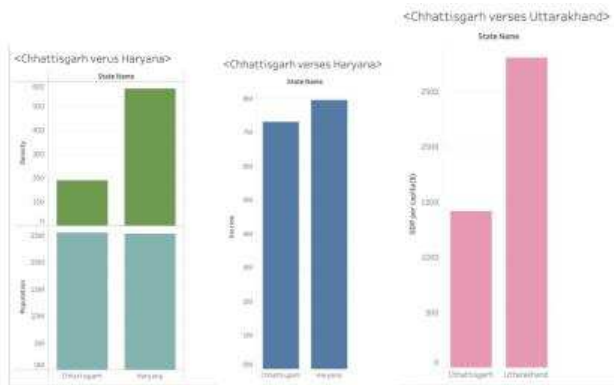


图 6 恰蒂斯加尔邦与哈里亚纳邦数据对比图

为了证明此观点, 我们比较北阿坎德邦与恰蒂斯加尔邦的数据, 北阿坎德邦人口密度为 189, 疫苗 3,821,058, 病例总数为 43,906,703, 这两个地区的人口密度是相同的, 恰蒂斯加尔邦的疫苗数量大约是北阿坎德邦的两倍, 但 covid-19 病例总数确是北阿坎德邦的 3 倍多, 而恰蒂斯加尔邦的人均 GDP 为 1416 美元, 北阿坎德邦的人均 GDP 为 2808 美元, 北阿坎德邦人均 GDP 是恰蒂斯加尔邦的 2 倍。因此进一步证明了经济水平是影响新冠病毒传播的因素之一。

四、人口密度对 Covid-19 病例的影响

在其他影响因素相似的情况下, 我们经过对印度数据分析了人口密度与 Covid-19 总病例数的关系。

从表 1 曼尼普尔邦和贾坎德邦在经济、第二剂疫苗接种率、国际机场和港口方面都很相似, 但贾坎德邦的病例数是曼尼普尔邦的三倍多。仔细看看他们的数据就会发现, 恰尔坎德邦的人口密度是每平方公里 414 人, 曼尼普尔邦的人口密度是每平方公里 128 人。恰尔坎德邦的人口密度也是曼尼普尔邦的三倍多。所以我们可以假设, 在其他因素相似的情况下, 由于贾坎德邦人口更密集, 人们更有可能接触到 Covid-19, 因此更有可能感染 Covid-19。

表 1 曼尼普尔邦与贾坎德邦各情况表

州/中央直辖区	人口密度	人均 GDP	机场数量	港口数量	疫苗接种率	总确诊人数
曼尼普尔邦	128	1022	1	0	2.76%	99,331
贾坎德邦	414	1111	0	0	3.69%	347,200

表 2 古吉拉特邦与喀拉拉邦各情况表

州/中央直辖区	人口密度	人均 GDP	机场数量	港口数量	疫苗接种率	总确诊人数
古吉拉特邦	308	2885	3	1	10.18%	824,900
喀拉拉邦	860	2983	4	1	11.19%	3,411,489

我们可以通过比较人均 GDP 约 2900 美元的古吉拉特邦和喀拉拉邦进一步证明这一点, 表 2 显示它们都有相似的国际机场和国际港口数量, 而且它们的第二次疫苗接种率都在 10% 左右。然而, 古吉拉特邦和喀拉拉邦的病例数量相差四倍多。对比这两个邦的人口密度可以看出, 喀拉拉邦的人口密度是古吉拉特邦的近 3 倍, 进一步证明了人口密度与确诊病例总数呈正相关。

五、结束语

本文从影响 COVID-19 传播的宏观因素进行了分析, 得出政策、疫苗、经济、人口密度对 COVID-19 的传播有一定的影响, 使政府能够更好地控制 COVID-19 的传播和未来的损害。接下来, 我们根据个人个体情况的差异研究被影响感染的概率微观因素, 据根个人情况和症状运用决策树模型来显示个人感染 COVID-19 的概率。

参考文献:

- [1] Alenezi, M., Al-Anzi, F., & Alabdulrazzaq, H. Building a sensible SIR estimation model for COVID-19 outbreak in Kuwait. *Alexandria Engineering Journal*, 2021, 60(3), 3161-3175.
- [2] Bartsch, S. M., O' Shea K. J., & Wedlock, P. T., et al. The Benefits of Vaccinating With the First Available COVID-19 Coronavirus Vaccine. *American Journal of Preventive Medicine*, 2021, 60(5), 605-613.
- [3] Bassi, G. L., Suen, J., & Barnett, A. G., et al. Design and rationale of the COVID-19 Critical Care Consortium international, multicentre, observational study. *BMJ Open*, 2020,

10(12), 1-6.

[4] Becker, B., Greenberg, M. R., & Madsen, T., et al. Sex-or Gender-specific Differences in the Clinical Presentation, Outcome, and Treatment of SARS-CoV-2. *Clinical Therapeutics*, 2021,43(3), 557-571.

[5] Chokshi, A., DallaPiazza, M., & Zhang, W. W., et al. Proximity to international airports and early transmission of COVID-19 in the United States—An epidemiological assessment of the geographic distribution of 490,000 cases. *Travel Medicine and Infectious Disease*, 2021, 40.

[6] Faura, C., J. COVID-19 Mortality Rate and Its Incidence in Latin America:

Dependence on Demographic and Economic Variables. *International Journal of Environmental Research and Public Health*,2021,18(13), 6900.

[7] Kassem, A. Z. E. Does Temperature Affect COVID-19 Transmission? *Front Public Health*, 2020,8.

[8] Mozumder M., Amin M., & Uddin M., et al. (2021). Coronavirus COVID-19 outbreak and control: Effect of temperature, relative humidity, and lockdown implementation. *Archives de Pédiatrie*,2021, 28(2), 111-116.

[9] Rahman, M. S., Azad, M. A. K., & Hasanuzzaman, M., et al. How air quality and COVID-19 transmission change under different lockdown scenarios? A case from Dhaka city, Bangladesh. *Science of The Total Environment*, 2021, 762. .

[10] Shringi, S., Sharma, H., & Rathie, P. N., et al. (2021). Modified SIRD Model for COVID-19 Spread Prediction for Northern and Southern States of India. *Chaos, Solitons & Fractals*.2021.

[11]Sun, Z., Zhang, H., & Yang, Y., et al. December 1). Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. *Science of The Total Environment*, 2020,746.

[12] WHO. WHO Coronavirus (COVID-19) Dashboard. *World Health Organization*.2021,