

코로나 19와 서울 소상공인 상권의 상관관계 분석

김재호¹ · 김장영^{2*}

The Analysis of Correlation Between COVID-19 and Seoul Small Business Commercial Districts

Jae-Ho Kim¹ · Jang-Young Kim^{2*}

¹Graduate Student, Department of Computer Science, The University of Suwon, Hwaseong, 18323 Korea

^{2*}Associate Professor, Department of Computer Science, The University of Suwon, Hwaseong, 18323 Korea

요약

현재 국내든, 해외든 코로나19로 인해 많은 소상공인들이 피해를 입고 있고, 많은 점포들이 문을 닫고 있는 것이 현실이다. 국가재난지원금을 통해 소비자들의 소비를 격려하면서 어느 정도 피해를 막으려고 하지만, 소상공인들의 폐업을 막는 것은 힘들게 되었다. 2020년 9월 서울기준, 코로나19사태로 인해 점포2만곳 이상이 폐점되었고, 코로나 19블루로 인해 우울증에 호소하는 사람들도 많아졌다. 이 문제는 서울, 대한민국뿐만이 아닌, 전 세계적으로 코로나 19 사태에 피해입은 전 지역에 대한 문제다. 코로나19의 환자 수가 증가할수록 점포 수는 꾸준히 줄어들고 있다. 이를 피어슨, 스피어만, 켄달의 상관계수를 분석해 코로나19 환자 수와 점포 수의 음의 상관관계를 나타낸다는 것을 제시한다.

ABSTRACT

Currently, whether in a domestic or international sphere, many small businesses are suffering due to COVID-19. The grim reality is that several businesses are shutting down. While the national disaster relief grant was used to contain the damages by encouraging consumer spending, it has become difficult to prevent closures of small businesses. As of September 2020, more than 20,000 stores have closed in Seoul due to the COVID-19 pandemic. There has also been an increase in the number of people with depression caused by the COVID-19 blues. This issue is not only confined to Seoul in the Republic of Korea, but is influencing all other areas affected by the pandemic. As the number of COVID-19 patients increase, the number of open stores is decreasing steadily. The analysis of the correlation coefficient of Pearson, Spearman, and Kendall suggests a negative correlation between the number of COVID-19 patients and the number of stores in business.

키워드 : 코로나19, 상관관계, 피어슨 상관계수, 스피어만 상관계수, 켄달의 상관계수

Keywords : COVID-19, Correlation, Pearson correlation coefficient, Spearman correlation coefficient, Kendall's correlation coefficient

Received 22 February 2021, Revised 24 February 2021, Accepted 8 March 2021

* Corresponding Author Jang-Young Kim(E-mail:jykim77@suwon.ac.kr, Tel:+82-31-229-8345)

Associate Professor, Department of Computer Science, The University of Suwon, Hwaseong, 18323 Korea

Open Access <http://doi.org/10.6109/jkiice.2021.25.3.384>

print ISSN: 2234-4772 online ISSN: 2288-4165

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

1.1. 구현동기

지금 우리는 뉴스 기사나, 그 외의 많은 매체에서 코로나19사태와 함께 그로 인한 피해를 보도하고 있다. 길 거리에 나가봐도 우리는 쉽게 매장들이 사라진 것을 볼 수 있다. 서울기준 2020년 2분기 점포 수는 2만여개가 폐업을 하였다[1]. 자영업, 소상공인들의 수입이 70% 가까이 줄어들었다[그림1]. 헬스장 같은 경우는 영업정지와 형평성 논란으로 어려움을 호소하고 있다. 실제로 코로나19블루로 인해 우울감이 증가하고 있고, 자살을 하는 기사를 접하면서 심각성을 강조하고자 한다.

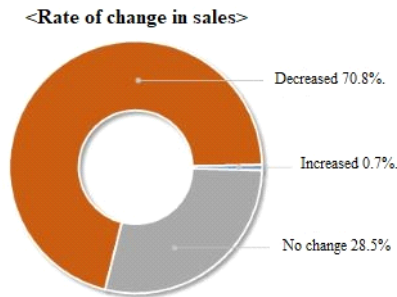


Fig. 1 Rate of change in sales

1.2. 코로나19 전염병 발생

코로나19는 SARS-CoV-2라는 신종 코로나 바이러스에 의해 발생하는 질병이다. WHO는 2019년 12월 31일 중국 우한에서 발생한 '바이러스성 폐렴' 사례 집단보고에 이어 신종 바이러스에 대해 처음 알게 되었다. 코로나19는 이전에 인간에서 확인되지 않은 신종 코로나 바이러스로 인한 질병이다. 대부분의 경우 코로나19는 마른기침, 피로감 및 열을 포함한 경미한 증상을 유발하지만, 일부 노인에게는 열이 증상이 아닐 수 있다. 다른 경미한 증상으로는 통증과 코 막힘, 콧물, 인후통 또는 설사가 있다. 어떤 사람들은 감염되었지만, 증상이 나타나지 않았다. 대부분의 사람들은 특별한 치료 없이 질병에서 회복된다. 코로나19에 걸린 6명 중 1명이 중병에 걸리고 호흡 곤란을 겪는다. 코로나19는 주로 감염된 사람이 다른 사람과 밀접한 접촉을 할 때 사람 간에 퍼지는 SARS-CoV-2바이러스에 의해 발생한다. 바이러스는 감염된 사람이 기침, 재채기, 말, 노래 또는 심호흡할 때 작은 액체 입자로 입이나 코에서 퍼질 수 있다. 이 액체 입자는 더 큰 '호흡기 방울'에서 더 작은 '에어로졸'에 이르

기까지 크기가 다르다. 다른 사람들은 바이러스가 입, 코 또는 눈에 들어갔을 때 코로나19에 감염될 수 있으며, 이는 사람들이 감염된 사람과 직접 또는 가까운 접촉을 할 때 발생할 가능성이 더 높다[2].

1.3. 국내, 해외 코로나 실태

2021년 국내 코로나에 감염된 인구는 61,759명, 사망 인구수는 917명으로 위험한 수준이며, 국가기관에서도 방역에 힘쓰고 있다. 2021년 1월 1일 전 세계 기준 코로나에 발병된 인구는 81,947,506명, 사망 인구수 1,808,041명으로 WHO 보건기구에 의해 보고되며, 전 세계적으로 COVID19는 계속해서 늘어나는 추세다[2].

1.4. 최근동향

아스트라제네카, 모더나, 화이자등 여러 제약회사에서 백신을 개발하였고, 접종을 시작하였고, 대한민국도 2021년 11월까지 코로나19의 백신 접종을 완료할 것으로 발표했다. 반면, 외국에서는 화이자 백신을 접종하고 6일 만에 코로나19에 감염된 사례, 각종 부작용과 변이 바이러스가 나오는 사례들이 나타나고 있다.

1.5. 코로나19로 인한 상권붕괴

코로나19로 인해 많은 소상공인들이 피해를 입었고, 폐업까지 이어지면서 엄청난 경제타격을 입었다. 전염성이 강하기 때문에 경제활동이 적어졌기 때문이다. 배달업체 경우 어느 정도의 피해를 상쇄시킬 수 있겠지만, 모든 업종들이 배달로 해결되지 않는다. 코로나19 사태와 같이 팬데믹이 발생했을 경우 어떠한 업종들이 큰 타격을 입는지 데이터를 통해 알아보려고 한다.

II. 기존연구

기존 소상공인, 자영업자들의 업종별 폐업, 개업에 요인을 미치는 것에 관한 많은 연구들이 존재한다[3-5]. 본 논문은 코로나19 사태 전후로 소상공인, 자영업자들의 폐업 수를 제시한다.

2.1. 서울시 최저임금에 따른 업종별 개업 및폐업의 영향 요인 분석

서울시라는 대표성이 있는 지역을 채택하여 서울시

표준 45개 소상공인 업종의 변화에 대해 모두 밝히고자 하였으며, 상업지역비율과 도로비율이라는 지리적 특성을 기준으로 지역을 구분하여 업종별 개 폐업률을 관찰한다. 기술통계를 통해 지역에 따라 나타나는 지역적 특색과 소상공인 생태계 규모의 차이를 관찰하였으며, 점포밀도, 프랜차이즈비율, 개업률, 폐업률의 차이를 확인하였다. 또한, 분산분석을 통해 2017년 2Q와 2019년 1Q 기간 사이 소상공인 생태계에 변화가 일어난 것을 발견하였다. 한편, 최저임금의 상승은 인건비 부담을 초래해 소상공인의 폐업률이 증가하고, 개업률이 감소한다는 것이 사회적인 통념이다. 그러나 실험결과에서 업종과 지역에 따라 그 현상이 다르게 나타나는 것이 입증되었다. 최저임금 상승으로 유의미하게 폐업률이 상승한 업종은 없었으며, 개업률의 감소에서 뚜렷하게 나타났다.

2.2. 지역상권 특성이 자영업자 폐업률에 미치는 영향에 관한 연구

연구 목적은 개인과 기업의 특성을 함께 보유하고 있는 자영업자의 신용위험 평가에서 이들 점포가 위치한 상권 또는 지역 정보가 이들의 신용위험(폐업률)을 설명하는데 유용한 정보를 포함하고 있는지 등을 살펴보고자 하는 것이다. 분석결과, 외식업의 경우 프랜차이즈 점포 수 증가율(상권의 성장성), 1년 생존률(상권의 변화도), 5년 생존률(상권의 성숙도), 개업률 등이 이들의 폐업률을 유의미하게 설명하고 있었다. 서비스업은 이에 더해 일반점포 수(상권의 규모)와 유동인구 증가율(상권의 다양성)도 폐업률에 유의한 영향을 미치고 있음을 살펴볼 수 있었다.

2.3. 업종별 창업 및 폐업의 지리적 특성 분석

본 연구에서는 개인과 법인의 창업 및 폐업에서 나타나는 지리적 분포 특성과 실제 창업 비중이 높은 생활서비스 및 생활소비와 관련 있는 생활밀접업종을 대상으로 세부 업종별 창업과 폐업의 증감에 따른 지리적 분포 양상을 살펴보고 이들 분포에 작용하는 사회경제적 요인을 파악하고자 하였다. 이를 위하여 전국 227개 시·군·구를 대상으로 개인과 법인 형태의 신규창업과 폐업의 공간 분포에서 나타나는 특징을 비교하고, 이 중 100대 생활밀접업종을 중심으로 창업과 폐업이 두드러지는 세부 업종을 추출해 이들의 지리적 증감의 분포 패턴을 살펴보았다.

III. 분석 알고리즘

기존에 있던 Pearson 상관계수와 Spearman 상관계수의 비모수 버전이라 볼 수 있는 Spearman 상관계수, Kendall 상관계수, Chi-squared test, 단순회귀분석을 비교해서 코로나감염과 점포 수에 관해서 분석한다.

3.1. Pearson 상관계수

Pearson 상관계수는 두 변수 간의 선형성이 얼마나 강한지 측정하기 위해 사용된다. 한 변수의 변화가 다른 변수의 변화에 비례적으로 연관되어 있을 경우 선형 관계가 있다. 두 변수 간에 선형성이 강하다는 것은 변수들의 관계가 직선에 의해 잘 모델링 된다는 뜻이다[6].

3.2. Spearman 상관계수

Spearman 상관계수는 단조성을 평가하기 위해 사용된다. 단조성이 좋다는 것은 한 변수의 값의 크기가 커지면(또는 작아지면) 다른 변수의 크기도 커진다(또는 작아진다)는 뜻이다. 절댓값이 1에 가까울수록 단조 상관성을 가지고, 0에 가까울수록 단조 상관성이 없는 것이다[6].

3.3. Kendall 상관계수

Kendall 상관계수도 단조성을 평가하기 위해 사용된다. 변수는 각각 일치하는 쌍들의 수와 일치하지 않는 쌍들의 수를 가리킨다. 일치하는 쌍들의 수가 많을수록 Kendall 상관계수는 1에 가까워질 것이다. Kendall 상관계수도 절댓값이 1에 가까울수록 단조 상관성을 가지고 있고, 0에 가까울수록 단조 상관성이 없는 것이다[6].

3.4. Chi-squared test

Chi-squared test는 교차분석이라고 불리며 두 범주형 변수에서 관찰된 빈도가 기대 빈도와 의미 있게 다른지를 검증하기 위해 사용된다. 즉, 두 범주형 변수 A, B가 종속 사건인지 독립 사건인지 판별하는 것이다. 기대 빈도가 5 미만인 셀이 전체의 20%를 초과할 때에는 가설 검정 결과가 정확하지 않다. 따라서 많은 표본을 확보하는 것이 결과의 신뢰도를 크게 높일 수 있다[7].

3.5. 단순 회귀분석 (Regression analysis)

단순선형 회귀는 독립변수와 종속변수 사이의 선형

관계를 파악, 예측하는 데 사용한다. 수집 데이터를 관찰하여 관계식을 만들고, 현재 수집된 데이터의 추세를 반영하는 것이 목표다[8].

3.6. Pearson 및 Spearman, Kendall 계수의 비교

상관계수란 2개의 변수 간의 연관성을 보여주는 지표이다. 여기서 r은 상관계수를 의미한다. 모든 상관계수는 다음을 따른다. 상관계수 r의 범위는 $-1 < r < 1$ 이다. 절댓값이 1에 가까울수록 상관성이 크다는 것을 의미하며, 값이 1에 가까울수록 두 변수 간 움직임이 비슷함을 뜻하고, -1에 가까울수록 움직임이 역방향으로 비슷하지 않다는 것을 의미한다.

- r>0: 양의 상관관계(x가 증가하면 y도 증가)
- r<0: 음의 상관관계(x가 증가하면 y는 감소)
- r이 0에 가까울수록 상관관계가 적다.
- r이 1이나 -1에 가까울수록 상관관계가 높다.

Pearson = -1, Spearman, Kendall = -1에 가까울 때, 한 변수가 감소하면 다른 변수가 증가하지만, 양이 일정하지 않은 관계인 경우, Pearson 상관 계수는 음수이지만 -1보다 크다. 이 경우 Spearman 계수는 여전히 -1이다.

Pearson = -0.799, Spearman, Kendall = -1에 가까울 때, -1 또는 1의 상관 값은 원의 반지름과 원둘레의 관계처럼 정확한 선형 관계를 의미한다. 그러나 상관 값의 실제 값은 완전한 관계를 나타내는 값보다 작다.

IV. 실험결과 및 분석

4.1. 코로나19 발생 전후 점포 수 변화

코로나19라는 질병이 발생한 후, 엄청난 전염성 때문에 세계적으로 모든 가게들이 피해를 보고 폐점하는 경우가 많이 발생했다. 따라서 우리나라 서울기준 코로나 전후로 점포 수 변화를 기술하고자 한다.

4.2. 코로나19 이전 점포 수

2019년 3월과 2019년 9월의 점포 수를 관광/여가/오락, 부동산, 생활서비스, 소매, 숙박, 스포츠, 음식, 학문/교육으로 대분류해서 살펴볼 수 있다. 6개월 사이, 2019년 3월 전체 서울 전체 점포 수는 382,245개, 2019년 9월 점포 수는 407,127개로 전체적인 점포 수는 증가하는 추세였던 것을 알 수 있다.

4.3. 코로나 이후 점포 수

코로나19 발생일은 2019년 12월 12일, 우리나라 첫 코로나 발생한 2020년 1월 20일이다. 따라서 가장 눈여겨보아야 할 데이터는 2019년 9월과 2020년 3월 데이터다. 모든 대분류 점포 수들이 줄었다. 6개월 사이에 15,627개의 점포 수가 줄었다. 2020년 9월의 점포 수는 367,534개로 총 39,593개의 점포가 줄었다. 코로나19로 인해 큰 폭으로 감소하고 있는 추세이다[표1].

Table. 1 Number of store

Category/Date	2019.09.30	2020.03.31	2020.09.30
Sightseeing/Leisure/Entertainment	11,613	11,714	9,594
Real estate	15,600	15,778	15,226
Life service	84,809	68,002	64,221
Retail	125,767	126,954	117,790
Accommodation	5,305	3,131	2,937
Sports	167	240	538
Food	132,425	134,041	128,789
Academics/Education	31,441	31,640	29,439
Totally	407,127	391,500	367,534

표1은 대분류별 점포 수이다. 경우에 따라 스포츠분야가 생활서비스 분야에 들어가는 경우도 있어 스포츠 분야의 수가 적게 나왔다.

4.4. 코로나 감염자 수와 점포 수의 상관관계

Pearson 상관계수를 이용하여 코로나 감염자 수와 점포 수의 상관관계를 다음의 그림으로 제시한다[그림2, 표2].

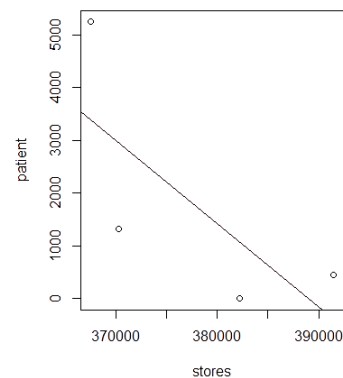


Fig. 2 Corona Infection and Number of Stores Correlation

Table. 2 Analysis algorithm value

	P-value	Coefficient of correlation/Explanation
Pearson	0.1963	-0.6910099
Spearman	0.08859	-0.7378648
Kendall	0.07697	-0.8207827
Chi-squared test	0.2133093	12
Regression analysis	0.2405	0.3652

Pearson 계수, Spearman 계수, Kendall 계수, Chi-squared test 분석, 회귀분석을 위 표와 같이 값을 도출하였다. p-value 값이 낮을수록 신뢰성이 높다는 것을 의미한다. 회귀분석은 p-value 값이 상대적으로 제일 높았다. Chi-squared test의 p-value 값은 데이터가 적기 때문에 결과 값이 정확하지 않다. 따라서 Pearson, Spearman, Kendall 계수를 서로 분석하는 것이 좋다고 판단했다. Pearson 계수 알고리즘 값(r)은 -0.6910099, Spearman 계수는 -0.7378648, Kendall 계수는 -0.8207827을 가진다. 본 실험은 Pearson과 Spearman, Kendall 계수를 비교, 분석하는데 있고, 분석은 음의 상관관계를 따른다. Pearson과 Spearman, Kendall 계수비교에서 언급했듯 정확한 선형관계를 의미한다고 볼 수 있다. 코로나 감염자 수와 점포 수의 상관성은 강력한 음의 상관관계, 선형관계로 감염자 수가 증가할수록 점포 수는 줄어든다는 것을 보여준다.

V. 결론 및 향후연구

본 논문은 기존 연구들과는 달리 코로나로 인한 점포 폐점을 수치화하고 있다. 코로나19로 인해 2019년 9월부터 2020년 9월의 전체 점포 수를 나타내고, 총 39,593 개의 점포 수가 줄어든 것을 알 수 있다. 이로써 코로나19 때문에 발생한 소상공인들의 경제적 타격을 수치화하여 알 수 있다. Pearson 상관계수와 Spearman, Kendall 상관계수의 관계를 분석하여 코로나감염자 수와 점포 수는 음의 상관관계를 가진다는 것을 도출하였다.

향후연구로 일반적으로 날씨, 환경, 건강 부분에서도 상관관계 분석이 입증할 수 있으므로 이 논문에 쓰인 Pearson, Spearman, Kendall, Chi-squared test, 회귀분석을 p-value 값에 따라 다른 분야에 적용하는 것을 제시한다. 예를 들면, 맥락과 수명에 대한 상관관계, 풍향에 따른 미세먼지 농도 상관관계 등 많은 분야에 사용될 수 있다.

REFERENCES

- [1] Federation of Small Business Association, Concerning the survey of the impact of small business owners on Corona19, [Internet]. Available: <http://www.kfme.or.kr/home/homeIndex.do>.
- [2] WHO Coronavirus Disease (COVID-19), [Internet]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>.
- [3] S. L. Bae, "Factors Affecting Small Business Opening and Closure : Comparison before and after the Rise of Minimum Wage Using Big Data of Seoul Metropolitan," Hanyang University Graduate School, Department of Business Informatics, Feb. 2020.
- [4] S. Y. Yoon, "A study on the impact of local trademark characteristics on self-employed business rates: mainly in 25 autonomous regions of Seoul," *Business and Innovation Research*, vol. 42, no. 3, pp. 21-39, Sep. 2019.
- [5] K. S. Lee and S. H. Park, "Geographical Characteristics of Business Start-up and Closing Business according to the Type of Industry," *Korean Economic Geography Society Publication*, vol. 22, no. 2, pp. 178-195, Jun. 2019.
- [6] Comparison of Pearson and Spearman correlation methods, Minitab, LLC. [Internet]. Available : <https://support.minitab.com/ko-kr/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/correlation-and-covariance/a-comparison-of-the-pearson-and-spearman-correlation-methods/>.
- [7] Y. S. Suh, "Historical Review of Independence Test: From Cai-Square Test to Boost Wrapping Method," *Korea Educational Evaluation Society of Education Evaluation Research*, vol. 33, no. 2, pp. 323-351, Jun. 2020.
- [8] S. J. Oh, *Data Analysis with R, 1th ed*, HanBit Academy, ch. 11, pp. 381-419, 2019.



김재호(Jae-Ho Kim)

수원대학교 컴퓨터학부 학사
수원대학교 컴퓨터학부 석사과정
※관심분야 : 인공지능



김장영(Jang-Young Kim)

연세대학교 컴퓨터과학 공학사
Pennsylvania State Univ. 공학석사
State University of New York 공학박사
University of South Carolina 교수
수원대학교 컴퓨터학부 교수
※관심분야 : Big data, AI, Cloud computing, Networks