Suicide Phenomena in South Korea from 2011 to 2020: Text Mining and Network Analysis of News Using Big Data

Jungeun Lee [1] and Jiyoung Lyu [2,*]

[1] Institute of Life and Death Studies, Hallym University, Chuncheon, South Korea; Researcher; ichocoi11@naver.com
[2] Institute of Aging, Hallym University, Chuncheon, South Korea; Associate Professor; jyyoo200@hallym.ac.kr
* Correspondence

Abstract: *This study aimed to identify latent clusters underlying suicide phenomena in South Korea from 2011 to 2020, a period marked by the country's highest suicide rates. To achieve this, 12,570 news articles were collected from BIG KINDS, a news article database of the Korea Press Foundation, and analyzed using big data techniques. Text mining was applied to article titles using Textom, followed by CONCOR analysis in UCINET6. Results were visualized using NetDraw. Through frequency analysis, 7,542 keywords were extracted. Of them, 86 high-frequency keywords were selected for network analysis. The CONCOR analysis revealed seven key thematic clusters: school, public officials, military, family, anomie, suicide attempts, and suicide locations. This study contributes to a deeper understanding of interconnected socio-cultural factors influencing suicide dynamics in South Korea. By examining a large, diverse dataset over a ten-year period, this research offers new insights into the evolution of suicide-related discourse and the role of media in shaping public attitudes. Findings of this study provide valuable implications for suicide prevention strategies, policy-making, and future research on the role of media in shaping societal perceptions of suicide.*

Keywords: Suicide; News Big Data; Text Mining; Network Analysis; South Korea

1. Introduction

South Korea (hereafter, Korea) consistently reports one of the highest suicide rates among OECD countries, highlighting suicide as a serious and persistent social concern. Following the 1997 Asian financial crisis, Korea's suicide rate increased significantly, peaking at 31.9 deaths per 100,000 individuals in 2011—the highest among OECD member nations at the time [1].

Suicide, defined as the intentional act of ending one's own life [2], is a multifaceted phenomenon influenced by a range of biological, psychological, and sociological factors [3]. However, beyond the individual dimension, suicide must also be understood through a socio-structural lens, as its impacts extend beyond the individual to affect families, communities, and society at large.

In recent years, researchers in Korea have increasingly turned to unstructured text data, such as social media and news articles, to explore suicide-related issues within broader societal contexts [4-6]. For instance, Song (2013) developed predictive models of suicide causation using social big data [4], while Kim (2021) analyzed two decades of media reporting on suicide to trace shifts in public discourse and propose possible solutions [5]. Similarly, Lee and Lyu (2021) assessed the feasibility of suicide prevention policies using news big data analysis [6].

News big data, in particular, has drawn attention for its ability to capture dynamic social change, reflect evolving public concerns, and mitigate sampling bias. With its archival structure and wide societal reach, news content provides a valuable resource for in-depth analysis using advanced techniques such as natural language processing (NLP) and semantic network analysis [7, 8].

Although some prior studies have applied semantic network analysis to suicide-related news—for example, Kim and Choi (2016) examined ideological frames in suicide reporting from two major Korean newspapers

between 2005 and 2014 [9]—these efforts have typically focused on limited sources or short time spans. As a result, they offer only partial insights into the broader evolution of suicide discourse during periods of sustained national concern.

To address this gap, the present study applies semantic network analysis to a large-scale, multi-source news big data corpus spanning a full decade (2011–2020)—a period during which Korea's suicide rate remained consistently high. Specifically, this study aims to (1) identify key thematic clusters and patterns in suicide-related discourse, (2) examine how public narratives surrounding suicide have shifted over time, and (3) offer a more comprehensive understanding of the socio-cultural dimensions of suicide in Korea.

The unique contribution of this research lies in its comprehensive scope, longitudinal perspective, and the application of semantic network analysis to an extensive news dataset. These features enable a deeper and more nuanced exploration of suicide-related discourse, advancing existing literature and providing new insights beyond previous studies that were limited by narrower media coverage or shorter analytical timeframes.

2. Materials and Methods

2.1 Data Collection

In this study, news big data was collected through the news collection system of BIG KINDS (https://www.bigkinds.or.kr), provided by the Korea Press Foundation. A total of 64,364 Korean news articles published between 2011 and 2020 were gathered using "suicide" as a search term, from 11 nationwide daily newspapers and 28 regional newspapers. After excluding articles that were irrelevant to the study, 12,570 articles were selected for analysis. The unit of analysis was the title of the news article, as it offers a concise and implicit representation of the article's content [10], fulfilling the role of journalism by conveying the core message and values of the information [11].

2.2 Data Extraction and Preprocessing

Using Textom, a big data analysis program, only nouns were extracted as keywords through morphological analysis and data preprocessing. The extracted keywords were then refined by integrating those with similar meanings and contexts. This step is essential because when keywords with similar meanings and contexts are not integrated, their frequency is diminished, and the meaning of the keywords becomes fragmented or underestimated, which can affect the results of the network analysis. After this refinement process, high-frequency keywords were selected through a frequency analysis of the refined keywords.

2.3 Network Analysis

In this study, CONCOR (CONvergence of iteration CORrelation) analysis was used to visualize the relationships between high-frequency keywords. CONCOR analysis is a method that measures the similarity in the status and role of nodes by grouping them into clusters based on the relationships between nodes in the network. This approach emphasizes the correlation of connection patterns rather than focusing solely on direct or indirect connections across the entire network. Additionally, block-modeling identifies blocks of nodes to analyze the relationship structure of subpopulations hidden within complex network populations [12]. As a result, the clusters identified by CONCOR analysis reflect keywords with a certain level of similarity. CONCOR analysis was performed using UCINET6 and visualized using NetDraw.

## 3. Results

3.1 Keyword Frequency

A total of 7,542 keywords were extracted through frequency analysis, and 86 high-frequency keywords were selected for network analysis, as shown in Table 1. Among these, the top 10 high-frequency keywords, including "suicide," were: "suicide attempt," "discovery," "throwing oneself," "police," "female," "rescue," "death," "male," and "selecting extremes." These results suggest a high prevalence of discussions around suicide attempts and frequent police rescues in the media. The keyword "throwing oneself" indicates that suicide by jumping or throwing oneself is a prominent topic in news articles. The appearance of the keyword "selecting extremes" may reflect the influence of the 'Media Guidelines for Reporting on Suicides' in Korea.

Table 1. High Frequency Keywords related to Suicide Phenomena in Korea

No.	Keyword	Freq.	No.	Keyword	Freq.	No.	Keyword	Freq.	No.	Keyword	Freq.
1	suicide	5684	23	sexual crime	337	45	couple	196	67	missing	129
2	suicide attempt	1571	24	apartment	335	46	public servant	193	68	military	129
3	discovery	1549	25	incident	327	47	disability	189	69	executive	128
4	throwing oneself	1158	26	probe	312	48	prosecutor	185	70	friend	123
5	police	993	27	mother	307	49	vitality	184	71	by oneself	122
6	female	864	28	life	304	50	elderly	180	72	debt	120
7	rescue	697	29	hanging	275	51	father	176	73	prison	120
8	death	673	30	vehicle	274	52	employee	172	74	corruption	119
9	male	671	31	firearms	271	53	taking poison	172	75	tragedy	116
10	selecting extremes	565	32	Daegu[1]	266	54	school violence	170	76	Son Ho Young[2]	114
11	murder	545	33	assailant	265	55	charge	170	77	NIS[3]	112
12	suicide note	513	34	men and women	264	56	hardship of life	165	78	CEO	112
13	high schooler	442	35	student	248	57	Incheon[4]	162	79	report	110
14	middle schooler	435	36	violence	241	58	suspicion	149	80	motel	107
15	soldier	433	37	uproar	237	59	professor	147	81	bridge	107
16	assumption	428	38	policeman	233	60	briquette	142	82	school	107
17	family	427	39	home	231	61	attempt	141	83	accident	106
18	accompanying	417	40	KAIST[5]	229	62	weapon	136	84	university	105
19	wife	380	41	quickness	214	63	dead body	136	85	hospital	104
20	son	365	42	suspect	210	64	Han River[6]	136	86	Busan[7]	104
21	pessimism	355	43	husband	209	65	shock	132			
22	investigation	348	44	depression	204	66	burning oneself	130			

3.2 CONCOR Analysis

Through the CONCOR analysis, seven clusters were identified: school, public officials, military, family, anomie, suicide attempt, and suicide location. The results of the CONCOR analysis are visualized in Figure 1.
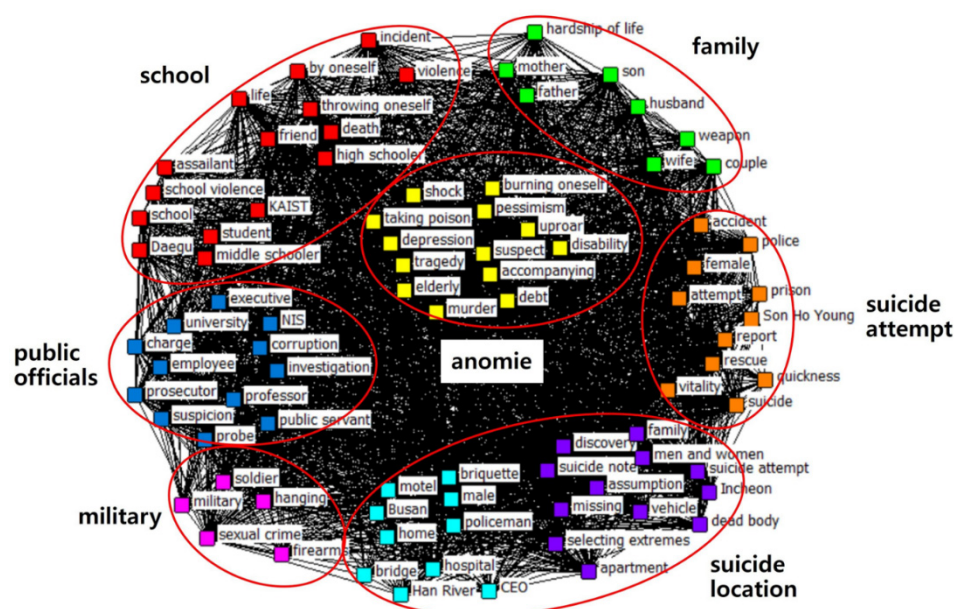
Figure 1. Clusters of Suicide Phenomena in Korea

---

[1] The region in Korea
[2] A celebrity of Korea
[3] National Intelligence Service
[4] The region in Korea
[5] The university in Korea
[6] The river in Korea
[7] The region in Korea

Table 2 summarizes each cluster, highlights the main keywords associated with it, and provides a brief interpretation.

Table 2. Summary of Suicide Clusters and ain Keywords

Cluster	Main Keywords	Interpretation
School	Daegu, school, middle schooler, student, school violence, assailant, KAIST, life, by oneself, friend, throwing oneself, incident, violence, death, high schooler	School violence is a significant cause of suicides in middle and high schools.
Public Officials	Executive, NIS, corruption, investigation, professor, public servant, probe, suspicion, prosecutor, employee, university, executive	Suicides among public officials are linked to corruption, such as bribery and abuse of power.
Military	Military, soldier, hanging, sexual crime, firearms	High suicide rates in the military are associated with the availability of firearms.
Family	Hardship of life, mother, father, son, husband, weapon, wife, couple	Family suicides often follow the murder of a family member, either ascendant or descendant.
Anomie	Shock, taking poison, depression, tragedy, elderly, murder, accompanying, suspect, burning oneself, pessimism, disability, debt, uproar	Reflects societal chaos, with a lack of common values or norms contributing to suicides.
Suicide Attempt	Accident, female, police, attempt, prison, Son Ho Young, report, rescue, vitality, quickness, suicide	Suicides are often prevented or responded to quickly by the police.
Suicide Location	Motel, briquette, Busan, home, policeman, bridge, hospital, Han River, CEO, discovery, family, suicide note, men and women, assumption, Incheon, vehicle, selecting extremes, apartment, dead body	Suicides occur in various locations, including common sites like bridges, motels, and homes, with notable trends in cities like Busan and Incheon.

4. Discussion

This study used CONCOR (CONvergence of iteration CORrelation) analysis to identify latent clusters in suicide-related news article titles in South Korea from 2011 to 2020. By applying network analysis to high-frequency keywords, this study visualizes the socio-cultural contexts and relational patterns surrounding suicide phenomena in Korea. These findings emphasize that suicide is not merely an individual problem but is deeply embedded in societal structures and institutional dynamics.

The keyword "suicide attempt" was ranked second only to "suicide," reaffirming prior research that positions suicide attempts as a key precursor to suicide death [13]. Its strong relational links to keywords such as "police," "rescue," and "throwing oneself" suggest a pattern of crisis interventions, especially by law enforcement, in immediate suicide prevention efforts [14]. These associations highlight the importance of timely detection and response systems in preventing suicide deaths.

CONCOR analysis yielded seven clusters, and this result suggests that suicide phenomena in Korea can be characterized by a high prevalence in specific groups, such as schools, public officials, the military, and families, with major causes of suicide being anomie. Discussions regarding seven clusters derived through CONCOR analysis are as follows.

First, the "school" cluster highlights that student suicides are primarily attributed to school violence and intense academic competition. Stressors related to university entrance and interpersonal conflict within the school setting are closely associated with these incidents [15]. This reflects the broader success-oriented culture in Korean society, where relentless academic pressure and high expectations contribute significantly to adolescents' mental health struggles [16]. Beyond anti-bullying programs, comprehensive educational reforms are required—such as restructuring university admissions criteria, expanding psychological support services within schools, and fostering inclusive school environments that prioritize emotional well-being.

Second, the cluster of "public officials" suggests that public officials often take their own lives under the weight of corruption scandals. Corruption, as a systemic issue arising from structural modernization processes [17], highlights the need for societal-level interventions. Therefore, suicide prevention strategies must include mechanisms to prevent corruption and reduce stigma by implementing fair investigative procedures, psychological support systems for officials under investigation, and structural transparency reforms.

Third, the "military" cluster represents suicides resulting from internal violence and sexual crimes. Within hierarchical military subcultures, behaviors such as beatings, verbal abuse, and harassment are prevalent. The increase in female soldiers has also been paralleled by rising instances of sexual violence, often perpetrated by superiors exploiting their authority [18]. To address these issues, it is imperative to establish mandatory independent military ombudsman offices, enforce zero-tolerance policies for abuse, and ensure safe reporting mechanisms. Institutionalization of gender-sensitive reforms is also crucial to promote safer military environments.

Fourth, the "family" cluster reveals that suicides frequently stem from familial conflicts and caregiving burdens. Economic difficulties are a major catalyst for family feuds, which can escalate into domestic violence, divorce, or abandonment [19, 20]. Additionally, caregiving responsibilities—especially among family members caring for elderly or chronically ill relatives—have led to a rise in murder-suicide cases. This trend reflects a growing societal issue in Korea, where the aging population places an increasing burden on informal caregivers [21]. Preventive measures should include caregiver support programs, subsidies for medical expenses, and community-based respite services.

Fifth, the cluster of "anomie" reflects the impact of social anomie, wherein the erosion of dominant norms amid rapid societal change leads to weakened social integration and increased suicide risk. Following the neoliberal reforms triggered by the 1997 financial crisis, Korea has experienced growing employment insecurity, poverty, inequality, and polarization [22-25]. These structural disruptions exacerbate feelings of alienation and hopelessness. National suicide prevention strategies should thus incorporate broader social policies—such as employment stabilization, universal healthcare, and equal access to education—to restore cohesion and reduce psychological distress.

Sixth, the "suicide attempt" cluster indicates that such incidents are particularly prevalent within prison environments. Contributing factors include the isolating nature of incarceration, separation from family, physical illness, and loneliness [26]. These findings suggest the need for integrated mental health services in correctional facilities. This includes routine mental health screenings, the establishment of crisis intervention teams comprising trained correctional staff and clinical professionals, and ongoing psychological support programs for at-risk inmates.

Seventh, the cluster of "suicide location" highlights common suicide locations—such as motels, homes, apartments, bridges, and vehicles—with a notable prevalence of suicides from high-rise apartment buildings. The correlation between increased urban apartment living and jumping suicides has been well-documented [27]. Architectural interventions are therefore essential, including installing protective barriers, restricting rooftop access, enhancing surveillance at high-risk sites, and deploying public signage to raise awareness of mental health resources.

5. Limitations

While network analysis provides valuable insights, the findings of this study are constrained by the quality and representativeness of the news articles analyzed. Several limitations should be acknowledged. First, this study was limited to analyzing news articles available through the BIG KINDS news collection system. Future research should aim to include a broader range of sources, encompassing all national and regional newspapers to enhance coverage and representativeness. Second, network analysis in this study focused on high-frequency keywords, which may have led to the exclusion of less frequent but potentially important terms. Particularly, the underrepresentation of suicides among socially disadvantaged groups—often overshadowed by media attention to celebrity suicides—may hinder deeper analysis of structural and societal causes of suicide. Third, the use of "suicide" as the sole search term may have excluded relevant articles that did not explicitly use the term. Since the release of the 'Media Guidelines for Reporting on Suicides ver. 3.0' in 2018, there has been a growing tendency in the media to avoid direct mention of the word "suicide." Future research should consider employing a more diversified set of search terms to capture a broader range of relevant content.

6. Conclusions

This study provides a novel perspective on suicide phenomena in South Korea by applying semantic network analysis to news big data from 2011 to 2020. By identifying and interpreting latent clusters within news discourse, the research highlights the complex, interconnected socio-cultural factors underlying suicide in Korean society. Although the analysis is confined to the Korean context, the study's significance lies in its affirmation that suicide is not merely an individual issue but a deeply social phenomenon—shaped by broader societal dynamics and structures. The findings indicate that key social groups contributing to elevated suicide

rates include students in school settings, public officials, military personnel, and family units. Notably, the prevalence of themes related to anomie suggests that societal disintegration, normlessness, and instability are major contributors to suicidal behavior. These insights underscore the need for systemic, context-sensitive policy responses. Suicide prevention efforts must extend beyond individual-level interventions to address structural inequalities, improve institutional environments, and foster societal cohesion. This study thus serves as a foundation for both future research and the development of more effective, socially informed suicide prevention strategies.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] OECD, OECD Health Statistics 2022, 2022. [Online] Available: http://www.oecd.org/health/health-data.htm

[2] American Psychiatric Association, Diagnostic and statistical manual of mental disorders: DSM-5 (Vol. 5, No. 5), Washington, DC, 2013.

[3] A. A. Leenaars, "Suicide notes at symbolic ages," Psychological Reports, vol. 78, no. 3, pp. 1034-1034, Jun. 1996, doi: https://doi.org/10.2466/pr0.1996.78.3.1034.

[4] T. M. Song, "Development of suicide search prediction model through analysis of social big data," Health and Welfare Policy Forum, vol. 8, no. 202, pp. 74-86, Aug. 2013, doi: http://dx.doi.org/10.23062/2013.08.8.

[5] D. Kim, "Exploring inter-media agenda-setting effects: Network agenda-setting model by using big-data analysis," Journal of the Korea Convergence Society, vol. 12, no. 5, pp. 121-126, May. 2021, doi: https://doi.org/10.15207/JKCS.2021.12.5.121.

[6] J. Lee and J. Lyu, "Exploring the suicide phenomena in Korea using news big data analysis," The Journal of the Korea Contents Association, vol. 21, no. 3, pp. 33-46, Mar. 2021, doi: https://doi.org/10.5392/JKCA.2021.21.03.033.

[7] S. Bloch-Elkouby, B. Gorman, A. Schuck, S. Barzilay, R. Calati, L. J. Cohen, F. Begum, and I. Galynker, "The suicide crisis syndrome: A network analysis," Journal of Counseling Psychology, vol. 67, no. 5, pp. 595-607, Oct. 2020, doi: https://doi.org/10.1037/cou0000423.

[8] G. Coppersmith, R. Leary, P. Crutchley, and A. Fine, "Natural language processing of social media as screening for suicide risk," Biomedical Informatics Insights, vol. 10, pp. 1-11, Aug. 2018, doi: https://doi.org/10.1177/1178222618792860.

[9] D. W. Kim and M. I. Choi, "Semantic network analysis on suicide reporting from 2005 to 2014: The Chosun Il-bo and the Hankyoreh," Korean Journal of Journalism & Communication, vol. 60, no. 2, pp. 178-208, Apr. 2016, doi: http://dx.doi.org/10.20879/KJJCS.2016.60.2.007.

[10] B. Brooks, G. Kennedy, D. Moen, and D. Ranly, News Reporting and Writing (7th ed.), Bedford/St. Martin's Press, New York, 2002.

[11] D. Dor, "On newspaper headlines as relevance optimizers," Journal of Pragmatics, vol. 35, no. 5, pp. 695-721, May. 2003, doi: https://doi.org/10.1016/S0378-2166(02)00134-0.

[12] S. Wasserman and K. Faust, Social network analysis: Methods and applications, Cambridge University Press, Cambridge, 1994.

[13] E. D. Klonsky, A. M. May, and B. Y. Saffer, "Suicide, suicide attempts, and suicidal ideation," Annual Review of Clinical Psychology, vol. 12, pp. 307-330, Jan. 2016, doi: https://doi.org/10.1146/annurev-clinpsy-021815-093204.

[14] Y. L. Kim, G. B. Kim, S. E. Lee, and J. S. Lee, "A study of the status and improvement measures of suicide response by police," Journal of Korean Public Police and Security Studies, vol. 19, no. 4, pp. 61-82, Nov. 2022, doi: http://dx.doi.org/10.25023/kapsa.19.4.202211.61.

[15] J. W. Chung and J. H. Kang, "Play subculture as a generating milieu of school delinquency," Korean Journal of Criminology, vol. 25, no. 2, pp. 35-73, Aug. 2013, doi: http://dx.doi.org/10.36999/kjc.2013.25.2.35.

[16] M. R. Lee, "Relationships of high school seniors' coping strategies used to handle examination stress with levels of depression: In relationship of additional life event stress," Journal of the Korean Home Economics Association, vol. 33, no. 5, pp. 89-98, Apr. 1995.

[17] J. T. Kang, "Modernization theory and rural environmental problem: From 'Progressive social change theory' toward 'Circular social change theory'," Journal of Agricultural Extension & Community Development, vol. 3, no. 2, pp. 285-297, Dec. 1996.

[18] J. Min, "An identification and analysis on the military organization's culture," Korean Review of Organizational Studies, vol. 8, no. 3, pp. 91-121, Dec. 2011, doi: http://dx.doi.org/10.21484/kros.2011.8.3.91.

[19] C. S. Chung, "Economic crisis and family life," Korea Journal of Population Studies, vol. 24, no. 1, pp. 91-121, Jun. 2001.

[20] A. J. Kposowa, "Divorce and suicide risk," Journal of Epidemiology & Community Health, vol. 57, no. 12, pp. 993-993, Dec. 2003, doi: https://doi.org/10.1136/jech.57.12.993.

[21] S. W. Park, "A study on the causes and countermeasures of elderly care crime," The Journal of Legal Studies, vol. 29, no. 2, pp. 129-160, Apr. 2021, doi: http://dx.doi.org/10.35223/GNULAW.29.2.6.

[22] M. H. Shin, "How and why is Korean society being polarized?: Its relation to globalization and povertization," The Journal of Asiatic Studies, vol. 49, no. 1, pp. 7-34, Mar. 2006.

[23] H. G. Lee, "The polarization of labour market and social integration -social integrative law & institutions and labour market policy," Korea Social Policy Review, vol. 20, no. 2, pp. 261-304, Jun. 2013, doi: http://dx.doi.org/10.17000/kspr.20.2.201306.261.

[24] D. Y. Gim, "A study on the effects of socioeconomic status on health level and medical service use," Institute for Humanities and Social Sciences, vol. 16, no. 4, pp. 329-369, Nov. 2015, doi: http://dx.doi.org/10.15818/ihss.2015.16.4.329.

[25] I. S. Nahm, "A study on socio-economic bipolarization and educational inequality in Korea," The Korean Journal of Humanities and the Social Sciences, vol. 35, no. 3, pp. 15-38, Sep. 2011.

[26] R. L. Bonner, Isolation, seclusion, and psychosocial vulnerability as risk factors for suicide behind bars. In R. W. Maris, A. L. Berman, J. T. Maltsberger, and R. I. Yufit (Ed.), Assessment and prediction of suicide, The Guilford Press, New York, pp. 398-419, 1992.

[27] D. Lester, "Suicide by jumping in Singapore as a function of high-rise apartment availability," Perceptual and Motor Skills, vol. 79, no. 1, pp. 74-74. doi: https://doi.org/10.2466/pms.1994.79.1.74.