Published online on the journal's webpage: **http://journal.ittelkom-pwt.ac.id/index.php/dinda**

Journal of Dinda

Data Science, Information Technology, and Data Analytics

Vol. 4 No. 2 (2024) 35 - 42 E-ISSN: 2809-8064

Hangout Places Recommendation System Using Content-Based Filtering and Cosine Similarity Methods

Abdul Raihan[1], Ahmad Ibrahim A.M[2], Alfian Akbar Ghozali[3*]

[1,2,3] Telkom University, Bandung, Indonesia

[1] abdulraihan@student.telkomuniversity.ac.id, [2] ahmadmubaroq@student.telkomuniversity.ac.id,

[3*]alfian@telkomuniversity.ac.id

Abstract

Coffee shops are becoming the new normal to hang out. Selecting the ideal location to hang out can be exceedingly difficult. There are too many choices, and it can be difficult to know where to begin. Based on this problem, a web application that responds to the growing need for an easy method of finding local hangouts is named Nongkies. This platform uses a recommender system to find cafes and restaurants easily. This system recommends places based on user preferences. This system was developed using the cosine similarity method, which is a systematic approach that uses a similar method based on cosine angles. Extensive testing has confirmed the reliability of Nongkies, offering user-friendly and accurate search results. Moreover, this app helps users find local hangouts and directions to those locations, especially university students, and the selection of places to socialize has a significant effect on students' learning experiences. By providing a tool that simplifies the decision-making process, Nongkies not only saves time but also enhances the social lives of its users, making it easier to discover new and exciting places to visit. The app's intuitive interface and efficient recommendation algorithm make it a valuable resource for anyone looking to explore their city's coffee shops and restaurants. With Nongkies, finding the perfect spot to relax, study, or meet friends becomes a hassle-free experience, ultimately contributing to a more enjoyable and productive lifestyle for its users.

Keywords: *Hangout, Recommender System, Content-Based Filtering, Cosine Similarity*

1. Introduction

The campus is like a learning hub where students can level up their skills and knowledge [1]-[4]. To improve learning, it is hoped that combining students' participation in campus activities with out-of-classroom experiences and engaging approaches that are relevant to everyday life will support effective learning [5]-[6] Besides being a place for education, campuses are like little versions of the entire world, with people from all kinds of backgrounds coming together. Consequently, the campus functions as a vibrant hub for intercultural discussions [7].

Commencing studies at a new university, particularly one situated away from one's hometown, puts everyone in a similar situation. Various universities are positioned in diverse regions of Indonesia, each with its advantages that may influence an individual's choice of where to pursue their education [8]. Every place has its unique vibe, so adjusting the way of speaking might be necessary to align with local customs. However, it's not just about changing speech patterns. Sometimes it is necessary to get accustomed to new methods or even new traditions. This adjustment can be overwhelming and may sometimes lead to feelings of homesickness or being out of place. [9].

Hanging out with friends or coworkers at coffee shops is becoming the new normal, not just for work meetings but for any kind of get-together [10]-[11]. It has become a cool thing to do, a whole new trend spreading like wildfire. And it makes sense - Indonesia's coffee scene is booming, and these cafes are like comfortable living rooms away from home [12]-[13]. The chill vibes and relaxed atmosphere boost people's creativity and make them more productive when they need to brainstorm or just unwind [14]. So, next time to catch up with someone or just need a change of scenery, grab a latte and head to the local café - it is the perfect place to relax, connect, and even get inspired.

Choosing the perfect hangout spot for a college student can be a real headache. There are just too many options,

Received: 27-05-2024 | Accepted: 21-06-2024 | Published: 01-08-2024

and it can be hard to know where to even start [15]. We know that the places students hang out outside of class are important for their growth and development, so we studied two rural high schools to see how the buildings, the rules, and how students interact with each other all come together to create the best learning environment [16]. What we found is that schools should think about everything when it comes to how they design and manage their spaces. They should give students more control over their environment and create chill pockets where everyone feels good and can learn and be themselves at the same time.

The goal is to provide these students with a tailored solution that considers their preferences, interests, and location, thus simplifying the decision-making process and ensuring they find the perfect hangout spot [17]. By using technology and data analysis, this personalized recommender system aims to alleviate confusion and enhance the overall hangout experience for university students.

In a neutral context, a cheap coffee shop provides the perfect environment for focus, teamwork, and friend socialization [18]. Although campus life offers exceptional chances for academic advancement and intellectual exploration, discovering suitable venues for socializing and relaxation can pose a challenging endeavor for university students. Given the abundance of cafes, co-working spaces, and alternative hangout choices spread both on and off campus premises, the question of "where to go" may become a common dilemma [19].

The choice of a socializing venue significantly influences the student experience [20]. Apart from providing a break from academic challenges, these places encourage social bonds, support creativity, and contribute to mental well-being. Deep conversations, idea-sharing, and forming lifelong friendships can happen at a favorite hangout spot. For those attending a university far from home, finding the perfect place can foster a sense of belonging. Spending time at local cafes can ease introductions to new cultures, create a welcoming atmosphere, and enhance the overall college experience. This paper consists of four sections: Section 1 discusses student behavior socialization, hangout research, university students' problem finding hangout places, and the recommender system for university students. Section 2 discusses hangout places in universities, the recommender system for hangout places, and categorizes hangout places with a filtering system and content-based filtering recommender system.

2. Research Method

This chapter describes the methodologies employed in the proposed system. It goes through the steps of collecting information, making a decomposition matrix, finding similarities between objects, making recommendations, and assessing the outcomes. Every stage is essential to understanding the operation of the system and the efficacy of its recommendation-making capabilities. Let us examine each of these techniques in more detail to figure out how they affect the system's overall operation. The research approach workflow is shown in Figure 1.

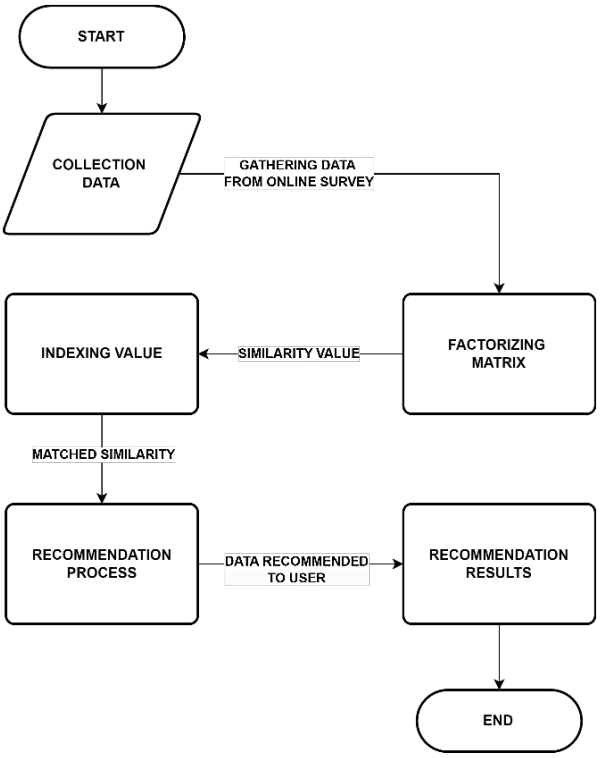

Figure 1. Workflow of Proposed Recommender Hangout Places System.

2.1. *Data Collection*

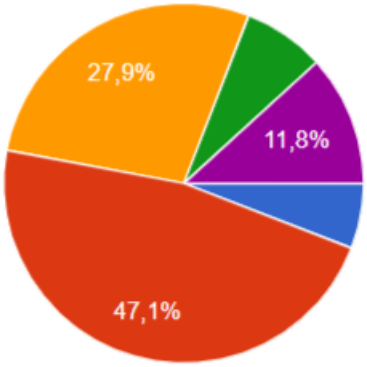

Figure 2. Survey Data Result

An online survey was performed to collect information about how often university students hang out, as shown in the following Figure 2. 47,1% of students preferred locations for socializing and the frequency of their hangouts is among the topics covered in the survey. The survey questions are straightforward to understand. Students will be asked to select their preferred hangout locations from a list of options in addition to being questioned about the frequency of hangouts.

These locations might be cafes, libraries, and more. Social media sites, student forums, and online community organizations were used to give the survey broadly. Students are welcome to take part in the survey and are guaranteed the confidentiality of their answers. After gathering survey answers, here is a simple analysis to see how students hang out and where they prefer to meet up. This study will be using the cosine similarity method.

2.2. *Factorizing Matrix Using Cosine Similarity*

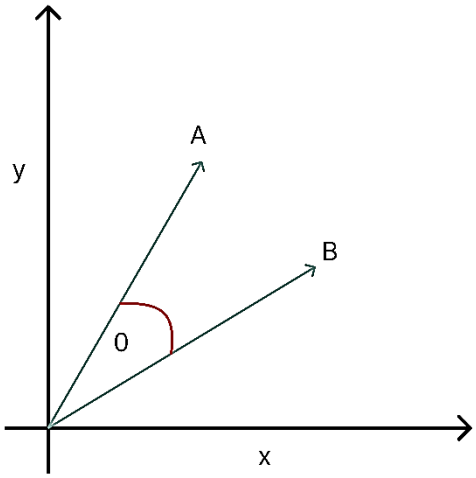

Figure 3. Cosine Similarity Graph

$$Sim(A,B) = \cos(\theta) = \frac{A \cdot B}{\|A\| B\|}$$

$$Sim(A,B) = \cos(0) = \frac{A \cdot B}{\| A \| B \|}$$

In this research, we use the cosine similarity method to construct a recommender hangout system as content-based filtering, to figure it out a graph is used as shown in Figure 3.

As seen in Figure 3, the angle between A and B is displayed in the diagram above. The smaller the angle between them, the more similar A and B are. Therefore, if the angle is small, A and B are almost the same, and if the angle is large, they are very different. Assuming A and B are two non-zero vectors, the dot product of the vectors is shown as ".", and the magnitude of a vector is shown as "|| ||". The dot product of vectors A and B is the product of their respective components. The magnitude of a vector is the square root of the sum of the squares of its parts.

Vectors are considered the same or very similar when their cosine similarity score is 1. If the angle between them is 90°, the cosine of 90° is 0, and a score of -1 means they are opposite. If the angle is 0°, the cosine of 0° is 1, indicating they are at right angles to each other. The cosine similarity scale ranges from -1 to 1. Cosine similarity measures the angle between two vectors. Therefore, two vectors are very similar if their cosine similarity is close to 1, and very dissimilar if it is close to -1.

2.3. *Indexing Values to Find Similarity Between Items*

Matrix factorization in this research is used in content-based filtering to detect commonalities. As part of this procedure, the matrix is dissected to find patterns or relationships in the data, and the similarity between these elements is then evaluated. The matrix is first factorized to expose underlying structures.

$$Tf(t) = \frac{frequency\ occurence\ of\ term\ t\ in\ document}{total\ number\ of\ terms\ in\ document}$$

$$Idf(t) = log10\frac{total\ number\ of\ document}{number\ of\ documents\ containing\ termt}$$

Figure 4. TF-IDF Formula

In this research, the TF-IDF (Term Frequency – Inverse Document Frequency) method is used to transform textual data into numerical vectors as can be observed in Figure 4.

Term frequency (TF) is the number of times a word appears in a document divided by the total number of words in that document.

Inverse document frequency (IDF) shows how many documents in the entire collection contain that word. The first step involves preprocessing the text data. Ultimately, a matrix with a fixed number of rows and columns is produced.

2.4. *Recommendation Process*

```
similarity_matrix
```

```
array([[1.        , 0.16903085, 0.1754116 , ..., 0.3       , 0.21081851,
        0.        ],
       [0.16903085, 1.        , 0.22237479, ..., 0.08451543, 0.17817416,
        0.08451543],
       [0.1754116 , 0.22237479, 1.        , ..., 0.1754116 , 0.2773501 ,
        0.        ],
       ...,
       [0.3       , 0.08451543, 0.1754116 , ..., 1.        , 0.10540926,
        0.2       ],
       [0.21081851, 0.17817416, 0.2773501 , ..., 0.10540926, 1.        ,
        0.        ],
       [0.        , 0.08451543, 0.        , ..., 0.2       , 0.        ,
        1.        ]])
```

Figure 5. Similarity Data Matrix

After that, straightforward methods like value comparisons and distance measurements are used to assess how similar these structures are to one another. To find similarity, the cosine similarity method is applied to the TF-IDF computation matrix as can be observed in Figure 5. Figure 5 shows matrix equations arrayed created from numbers that were converted using the TF-IDF methods. This process involves turning text into numbers, such as a list of cafe names with different category tags. Therefore, the figure's 2D array shows that each row in the café corresponds to a different category term. Some of these tags are the same, so some cafes end up with the same numbers.

In summary, cosine similarity and TF-IDF methods classification is a useful method for finding links between similarities in data. It makes it possible to create customized analyses and recommendations.

2.5. *Recommendation Result*

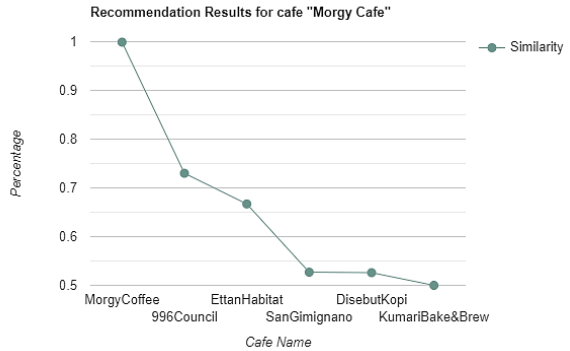

Figure 6. Graph Data Similarity Results

Based on the studies conducted, a content-based filtering method using cosine similarity in the recommendation system, the implementation of the recommendation system has been executed effectively, and the outcomes of the system's suggestions have been seen in Figure 6. The recommendation system works well in suggesting hangout places by considering what users like and the characteristics of the items. The feedback received from the system's suggestions proves a significant level of precision and appropriateness in recommending items to users.

3. Results And Discussion

This study's results and discussion section describe how the researchers collected a dataset and used cosine similarity to apply content-based filtering. The research findings are displayed in a table based on the research. Furthermore, the study offers a thorough analysis and interpretation of this research, emphasizing its importance and ramifications.

3.1 *Dataset*

The datasets that were obtained were gathered from Leonel Hertawan Trakter. id website, where the coffee shop list was obtained. From the data below it has 55 data which includes all the details about cafes in Bandung. After obtaining the dataset, the author will use content-based filtering and cosine similarity methods to develop it as a recommendation application. The example dataset can be seen in Table 1.

Table 1. Example Datasets

i/n	name	category	start price	street
0	Jati Kopi DU	nyaman, bangunan unik, cita rasa enak, populer, estetik, semi outdoor, viral	20	Jl. H. Wasid No.31
1	Tjimanoek 43	nyaman, viral, cita rasa enak, pelayan ramah, semi outdoor	17	Jl. Cimanuk No.43
2	996 Council	kopi enak, wfc able, pelayan ramah, instagramable, bangunan unik, populer, nyaman, cita rasa enak, barista	22	Jl. Trunojoyo No.4
...	Morgy Coffee	instagramable, cita rasa enak, bangunan unik, suasana alam, wfc able, asri	27	Jl. Anggrek No.46

55	Ettan Habitat	ambience, cita rasa enak, indoor, outdoor, wfc able, nyaman, instagramable, unik	20	Jl. Haya m Wuru k No.24

3.2 *Estimated Recommendation Values*

The author uses the TF-IDF algorithm and cosine similarity techniques in the computation outputs that are produced through an implementation in the Python programming language to produce recommendations for hangout places depending on on the user's chosen café. The author uses two different sets of input data to measure the effectiveness. Two recommended hangout locations are retrieved from the dataset for each test, arranged in decreasing order of similarity.

In this paper, the use of the cosine similarity method and TF-IDF in the recommendation system is an effective way to find connections between data similarities. To see the results percentage obtained can be seen in Table 2.

Table 2. First Data Input Similarity Result

i/n	Cafe Name	Similarity	Similarity Category
4	Morgy Coffee	1.0000%	11 words
2	996 Council	0.7302%	8 words
53	Ettan Habitat	0.6674%	7 words
24	San Gimignano	0.5270%	5 words
49	Disebut Kopi	0.5270%	5 words
13	Kumari BakeBrew	0.4000%	4 words

For the example usage, a cafe named “Morgy Coffee” has the top 5 recommendation results, there are “996 Council, Ettan Habitat, San Gimignano, Disebut Kopi, Kumari Bake & Brew”, which has similarities related to “Morgy Coffee” in terms of category or shared attributes and the same characteristic features of the place. “Morgy Coffee” has categories which consist of, *(“Instagram able, cita rasa enak, bangunan unik, suasana alam, wfc able, asri”*) Based on user recommendations, the following cafes are recommended: “996 Council, Ettan Habitat, San Gimignano, Disebut Kopi, Kumari Bake & Brew". These cafes are recommended to users based on their similarity to the coffee they chose, which in this case is "Morgy Coffee".

This system makes recommendations for cafés based on the choices and tastes of the user. The algorithm in the previously stated café example works by suggesting cafes based on the user's choices and interests. As displayed in the preceding table “Morgy Coffee” has 100% similarity which is the user's choice and then the next information, which includes suggested cafés that are comparable to the chosen café which in this case is "Morgy Coffee".

Essentially, the TF-IDF algorithm can translate language into numbers, meaning that related results will show up in the top suggestions regardless of how much text or number is displayed or similar in some café categories. As can be seen “996 Council” has 8 words similar categories with “Morgy Coffee” which is why it ranks first among the top suggestions and carried on with “Ettan Habitat” which has 8 words similar categories to “Morgy Coffee”, “San Gimignano and Disebut Kopi” has 5 words categories similar with “Morgy Coffee”, and “Kumari Bake & Brew” has 4 words categories similar with “Morgy Coffee” which places it last among the top suggestions.

The next testing filtering is based on start price and wifi conditions. the results percentage obtained can be seen in Table 3.

Table 3. Second Input Similarity Results Based on Price & Network

i/n	name	start_price	wifi_status	wifi_speed
1	Tjimanoek 43	17	Available	Very Fast
38	Saraga Coffe	18	Available	Very Fast
42	Kopi Karya Karsa	18	Available	Very Fast
3	ManA Social Cafe	21	Available	Very Fast
5	Arah Coffee	21	Available	Very Fast
31	Kozi Budara	21	Available	Very Fast

As displayed in the preceding table, Filtering the matrix based on pricing and network was how the testing was done. The top five recommendations that are most comparable to the chosen café—in this example, "Tjimanoek 43"—were shown.

Table 4. Second Data Input Similarity Result Details

i/n	Cafe Name	Similarity
1	Tjimanoek 43	1.0000 %
38	Saraga Coffe	0.9989 %
42	Kopi Karya Karsa	0.9989 %
3	ManA Social Cafe	0.9875 %
5	Arah Coffee	0.9875 %
31	Kozi Budara	0.9875 %

Additionally as shown in Table 4, "Saraga Coffee, Kopi Karya Karsa, ManA Social Cafe, Arah Coffee – Pasundan, Kozi Budara" and the chosen cafe are the most similar. In terms of price and network condition, "Saraga Coffee and Kopi Karya Karsa" is the closest match, coming in at 0.9989% similar. And so on with

"ManA Social Cafe, Arah Coffee – Pasundan, Kozi Budara," which, regardless of cost or network condition, has a 0.9875% comparable rating.

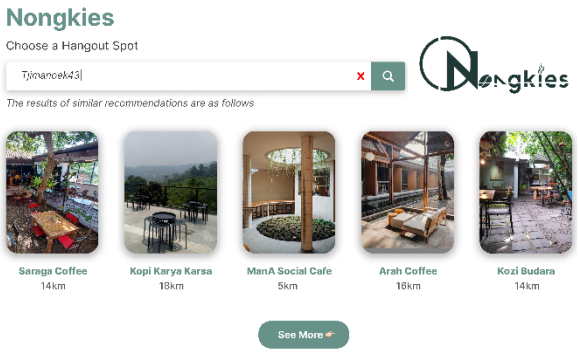

Figure 7. Mockup Similarity Results Design

Overall, the system showed excellent precision and efficacy. The mockup design can be seen in Figure 7.

Data obtained from testing the recommendation system on users shows that, among students, 52.3% hang out once a week, 18.5% hang out twice a week, 10.8% hang out three times a week, and 6.2% hang out more than three times a week. This research suggests that students often hang out during their free time or when working on assignments with friends.

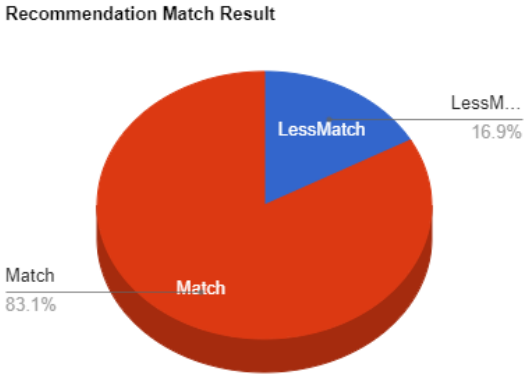

Figure 8. Recommendation Data System

Based on the results of surveys administered to students, the authors received an average rating of 4.153 out of 5,000 for how proper and suitable the hangout search recommendations are with users.

4. Conclusion

From the result of research and discussion Using content-based filtering along with the cosine similarity method made the hangout location suggestion system highly accurate and effective. Our innovative approach used item details and user preferences to offer relevant suggestions, boosting user satisfaction and engagement. The recommendations closely matched users' interests because the system effectively paired user preferences with the features of hangout spots. Cosine similarity in conjunction with the content-based filtering method proven scalability and flexibility. The recommendation algorithm relies on the cosine similarity method, which makes precise comparisons between user profiles and hangout spot attributes to find similarity scores. The great accuracy and dependable operation of the system were supported by these strategies. This accuracy was shown by the positive reviews and high acceptance rate of the suggested locations. The success of recommender systems using cosine similarity shows how effective they are. As the system develops, users will enjoy even more engaging and relevant hangout experiences, thanks to improved algorithms and more data sources. This innovative approach sets a new standard for recommendation systems, offering users a valuable tool for finding personalized, high-quality hangout spots.

References

[1] W. B. Michael and E. L. Boyer, “Chapter II: Campus Environment,” *Rev Educ Res*, 1965, doi: 10.3102/00346543035004264.

[2] O. Dakhi, J. JAMA, and D. IRFAN, “Blended learning: a 21st century learning model at college,” *International Journal Of Multi Science*, 2020, [Online]. Available: https://multisciencejournal.com/index.php/ijm/article/view/92

[3] Hernandez-de-Menendez, Marcela, Carlos Escobar Díaz, and Ruben Morales-Menendez. "Technologies for the future of learning: state of the art." *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 2020, doi: 10.1007/s12008-019-00640-0.

[4] K. Syauqi, S. Munadi, and M. B. Triyono, “Students’ Perceptions toward Vocational Education on Online Learning during the COVID-19 Pandemic,” *International Journal of Evaluation and Research in Education*, 2020, [Online]. Available: https://eric.ed.gov/?id=EJ1274581

[5] D. M. De Sawal, A. Peck, and D. M. Rosch, “Defining and Addressing Professional Burnout in Campus Activities.,” *Journal of Campus Activities Practice and Scholarship*, 2022, [Online]. Available: https://eric.ed.gov/?id=EJ1359330

[6] Y. P. Astuti and M. M. AR, “Implementation of the Campus Teaching Program Batch 3 in

Building Scientific Literacy in Elementary Schools," *Jurnal Penelitian Pendidikan IPA*, 2023, [Online]. Available: https://garuda.kemdikbud.go.id/documents/detail/3671877

[7] M. Groulx, N. Nowak, K. Levy, and A. Booth, "Community needs and interests in university–community partnerships for sustainable development," *International Journal of Sustainability in Higher Education* , 2021, doi: 10.1108/IJSHE-03-2020-0086.

[8] Y. Mclane and N. Kozinets, "Spatiality, experiences, and the formation of place attachment at campus student life centers," *Coll Stud J*, 2019, [Online]. Available: https://www.ingentaconnect.com/content/prin/csj/2019/00000053/00000001/art00010

[9] W. Mardhiyah and A. S. Kusuma, "Capital in Adaptation to Culture Shock in Overseas Students of Muhammadiyah University Surakarta," *Proceedings Series on Social Sciences & Humanities*, 2023, [Online]. Available: http://conferenceproceedings.ump.ac.id/index.php/pssh/article/view/822

[10] E. Kleszewski and K. Otto, "The perfect colleague? Multidimensional perfectionism and indicators of social disconnection in the workplace," *Pers Individ Dif*, 2020, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191886920302051

[11] Y. Zhou, X. He, and B. Zikirya, "Boba Shop, Coffee Shop, and Urban Vitality and Development—A Spatial Association and Temporal Analysis of Major Cities in China from the Standpoint of," *Remote Sens (Basel)*, 2023, [Online]. Available: https://www.mdpi.com/2072-4292/15/4/903

[12] A. Agusetyaningrum and N. H. Kistanto, "Hanging Out Phenomenon at Coffee Shop as An Environmental-based Economic tool in Covid-19 Pandemic Era," *E3S Web of Conferences*, 2021, [Online]. Available: https://www.e3s-conferences.org/articles/e3sconf/abs/2021/93/e3sconf_icenis2021_01044/e3sconf_icenis2021_01044.html

[13] K. Wang, B. Zhao, X. Peng, Z. Zhu, and ..., "Cafe: Learning to condense dataset by aligning features," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, [Online]. Available: http://openaccess.thecvf.com/content/CVPR2022/html/Wang_CAFE_Learning_To_Condense_Dataset_by_Aligning_Features_CVPR_2022_paper.html

[14] T. Tantra, A. M. Suyanto, and R. A. Jackie, "A Study on Consumer Preference Towards Braga Art Café in Bandung City," *practitioners. In order to reach the*, 2021, [Online]. Available: https://www.academia.edu/download/92625733/Proceeding_Book_ISCLO_2021.pdf#page=395

[15] M. Wrede-Jäntti, S. Aapola-Kari, and J. Lahtinen, "Correction to: This Is Where We Love to Hang Out! Interactive Maps as a Method of Gathering Information on the Everyday Lives of Finnish Young People in a," *Involving Methods in Youth*, 2021, doi: 10.1007/978-3-030-75941-4_13.

[16] E. Altenburger and B. R. Wellenreiter, "Where to hang out: Interplay between school building characteristics, authority structures, and school micro-climates," *Child Youth Environ*, 2021, [Online]. Available: https://muse.jhu.edu/pub/330/article/848927/summary

[17] V. Riel, "' They Just Hang Out with Their Own:' Organized Racial Inequality in a Rural Southern High School," *Journal of Negro Education*, 2020, [Online]. Available: https://muse.jhu.edu/pub/417/article/802543/summary

[18] E. Franke, *Creation of a Coworking Space in Lancaster, Ohio: A Business Plan on Pivot LLC.* rave.ohiolink.edu, 2023. [Online]. Available: https://rave.ohiolink.edu/etdc/view?acc_num=ouhonors1682695780720697

[19] M. Wrede-Jäntti, S. Aapola-Kari, and J. Lahtinen, "This Is Where We Love to Hang Out! Interactive Maps as a Method of Gathering Information on the Everyday Lives of Finnish Young People in a Longitudinal," *Involving Methods in Youth*, 2021, doi: 10.1007/978-3-030-75941-4_10.

[20] S. H. J. Hwang, E. Cappella, and M. J. Kieffer, "'Let's hang out!': Understanding social ties among linguistically diverse youth in urban afterschool programs," *Social*, 2022, doi: 10.1111/sode.12531.