A geospatial approach to identifying optimal adolescent mental health service locations in Toronto

10.17975/sfj-2025-014

Danny Zhang[1], Jed Lin[1*], Hudson Haas[1*], Darius Aul[1]

[1]Crescent School, Ontario, Canada

*Authors contributed equally

ABSTRACT: In Toronto, disparities in access to mental health services across neighbourhoods highlight inequalities in geospatial placement, a critical factor influencing service utilization. As adolescents in Toronto continue to face high rates of mental health challenges, evidence-based resource allocation offers a solution for more equitable access to appropriate services. By analyzing geospatial infrastructure and census data for the city's 140 neighbourhoods, an algorithm was developed to identify optimal placements for mental health services, prioritizing underserved areas. Min-max normalization was applied to public transportation route density, median after-tax income, adolescent population, and existing service density, assigning neighbourhoods a score from 0 to 1 to indicate service need. DBSCAN clustering was then used to identify clusters of high-need neighbourhoods in close proximity. The top cluster, Thorncliffe Park and Flemingdon Park, was further analyzed using a fixed-radius search to identify an optimal 500m placement radius that maximizes accessibility and ensures well-distributed services for adolescents. The service needs heatmap aligned with other studies on mental health service accessibility in Toronto, with the highest-need neighbourhoods in this model corresponding to those with the lowest accessibility. The clustering algorithm achieved a silhouette score of 0.602, indicating moderately strong clusters with room for improvement. It is recommended that policymakers use this algorithm with real-time data and adjusted weightings to identify service "cold spots" for placement. Future research should incorporate additional filtered data and account for off-limits zones in the placement optimization process. This model is adaptable to similar urban environments, provided consistent factor datasets are available.

KEYWORDS: resource optimization, mental health disparities, DBSCAN clustering, geospatial analysis

INTRODUCTION

Adolescence is a key period in human development that is characterized by increased neuroplasticity, during which the rapid development of strong cognitive, emotional, and social skills occurs [1]. Insufficient mental health support during this stage can have far-reaching and long-lasting consequences on one's future educational, career, and social opportunities [2]. This can contribute to the emergence and perpetuation of disparities in well-being across communities and demographic groups [3,4]. In Ontario alone, one in five children and youth are dealing with some form of mental illness at any given time [5]. However, less than one in six of those suffering actually receive the specialized care they need [6]. This underscores the need for early identification of mental illness, and more importantly, interventions that can lead to a better healthier life. Unfortunately, existing services currently fail to meet the needs of Ontario youth [7].

In Toronto, certain areas exhibit more mental health service usage than others, indicating wide disparities in access across neighbourhoods [8,9]. Research has shown that geographic proximity and spatial access to mental health services are key enabling factors in service utilization and preventing adverse outcomes [10,11]. Additionally, factors such as socioeconomic disparities, population density, service distribution, and transportation access can also influence the accessibility and effectiveness of mental health services [12-15].

Thus, our research aims to answer: How can geographic, socioeconomic, and transportation data be analyzed to optimize the spatial distribution of mental health services for adolescents in Toronto? Answering this question could be crucial to ensuring equitable mental health outcomes across various demographics. By focusing on variables such as Neighbourhood Median After-Tax Income, Neighbourhood Adolescent Count, Facilities per Population, and TTC Density (routes per km2), we aimed to provide empirical insights into where services are most needed.

We first created a heatmap of service needs in each of

Toronto's 140 neighbourhoods, then used DBSCAN clustering to identify groups of close-proximity neighbourhoods that are most underserved. Finally, we applied a location-optimization algorithm was used to identify the most accessible site within a given neighbourhood cluster. This multi-step approach helps rationally distribute resources that ensure that underserved communities receive the attention that they require.

METHODS

Programming Environment and Libraries

Multiple libraries and algorithms were used to conduct this research and to extract and process data. We used Python 3.1.0 in a Google Collaboratory environment. Some libraries used to process the data were: Pandas, NumPy, Matplotlib, GeoPandas, and Shapely.

Data Collection

This study focuses on Toronto due to the authors' familiarity with the local context, which enables more informed data interpretation and collection. Additionally, Toronto being an urban city demonstrates many advantages as more populated areas tend to yield greater variability in data, urbanicity is linked to elevated rates of mental health disorders and has more statistical power through a larger population pool [16].

The shapefiles that contained the geographical data for city and neighbourhood boundaries, transportation routes, and existing mental health facility locations were all found on the Toronto Open Data Portal, an open-source site containing official data published by the City of Toronto [17-21]. The transportation data contained both bus and subway routes which are part of the Toronto Transit Commission (TTC), which is the city's primary public transport agency. The Youth Asset Mapping Project selects the existing mental health facilities in collaboration with FindHelp/211 funded by the Ministry of Children and Youth Services, and consists of non-profit health centres that provide mental health support for youth. The census data for the average income and adolescent

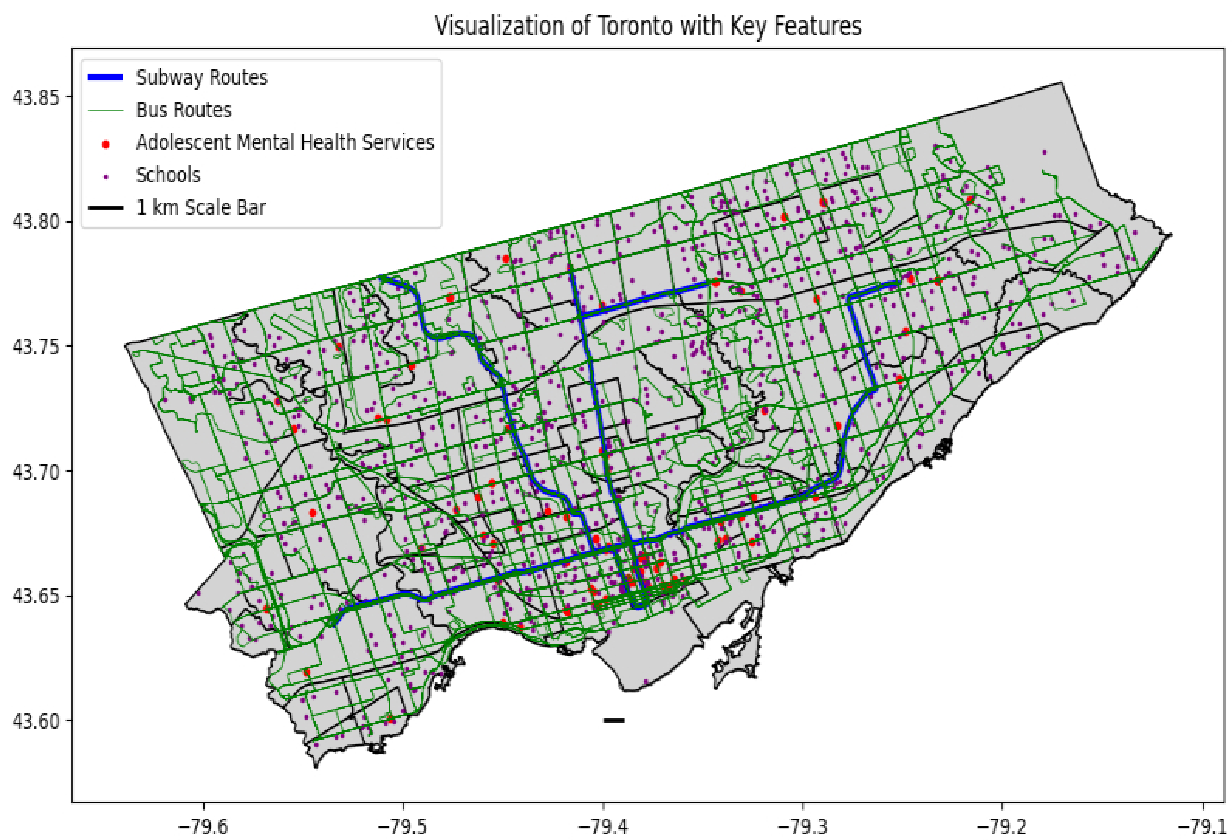

Figure 1. A visualization of key features of Toronto as used in our algorithm, including TTC routes, existing adolescent mental health facilities, neighbourhood boundaries and schools.

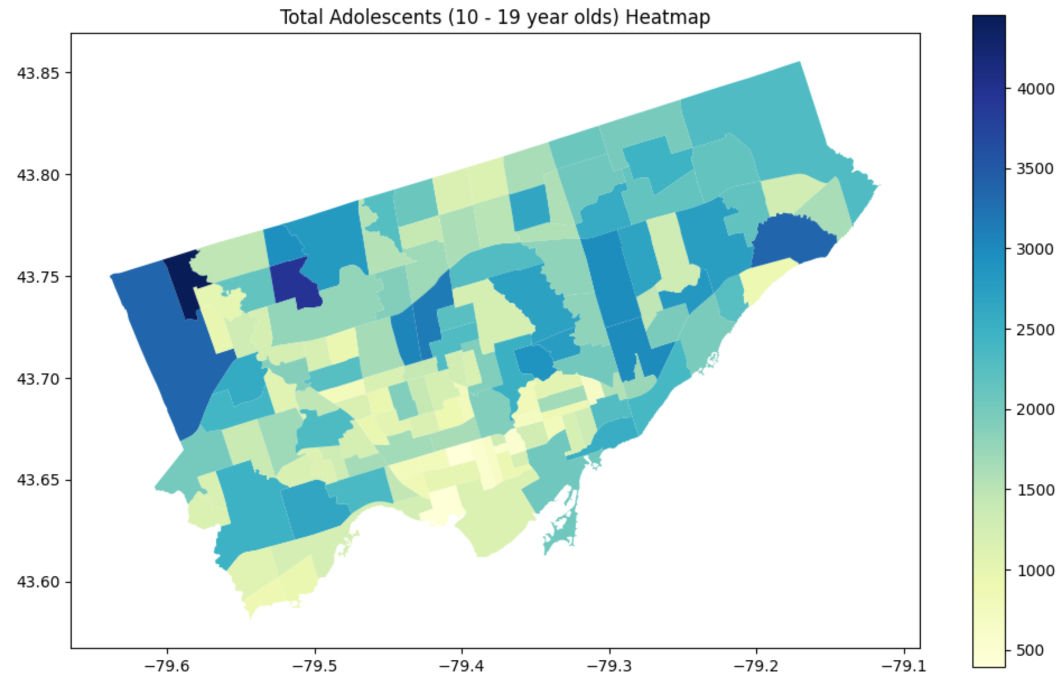

Figure 2. A visualization of total adolescent population distribution (as used in our study) represented on a neighbourhood heatmap.

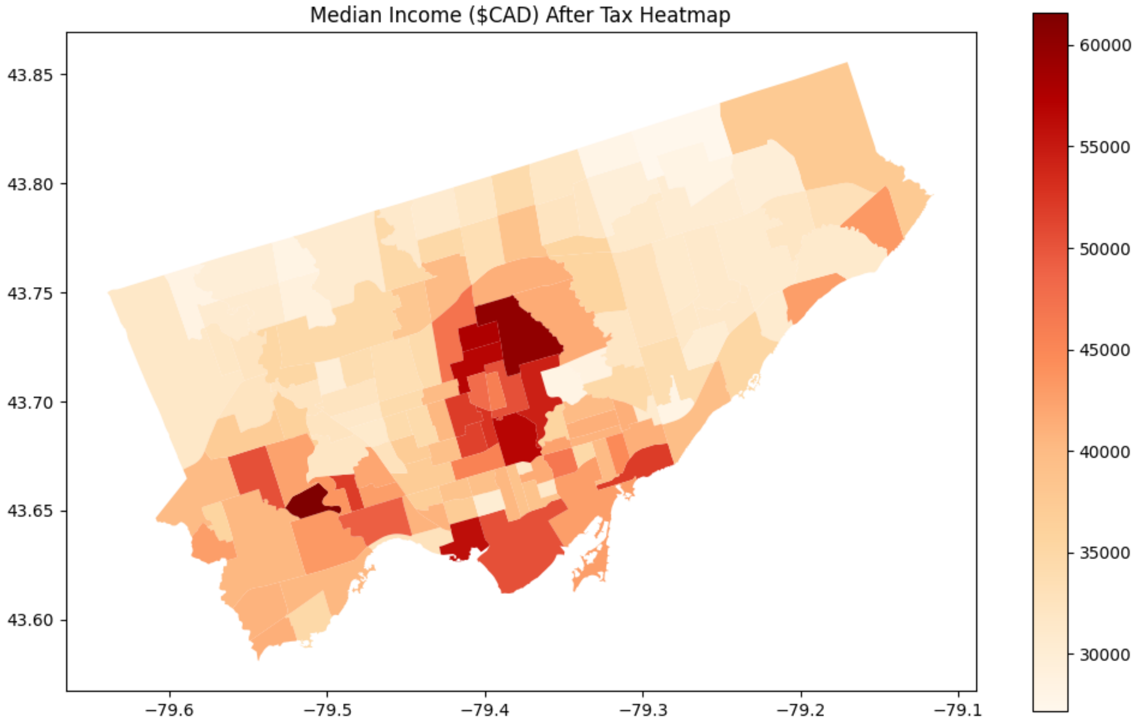

Figure 3. A visualization of household yearly income distribution (as used in our study) represented on a neighbourhood heatmap.

population for each neighbourhood were also available in the portal [22].

Preprocessing

The neighbourhood data did not align with the neighbourhood boundaries due to an update in April 2022 where many of the city's previous split off into smaller neighbourhoods. As a result, the census data contained 158 neighbourhoods, while the corresponding shapefile for neighbourhood boundaries contained 140 neighbourhoods. To resolve this discrepancy, we determined which neighbourhoods had split and recombined them by summing the adolescent populations and averaging the Median After-Tax Incomes of the smaller neighbourhoods. The adolescent population was calculated by adding the population of 10-14-year-olds and the population of 15-19-year-olds from the census data. The number of facilities per adolescent was measured by dividing the number of facilities in each neighborhood by the number of adolescents. To determine the transport route density of each neighbourhood, we performed a spatial join on the neighbourhood boundaries and transportation routes GeoData frames. We then calculated the number of times each neighbourhood boundary was intersected by a route, and divided that number by the number of square kilometers that make up the area of that neighbourhood. Although an alternative approach to finding the route density would have been to find the length of each route that intersected the neighbourhood boundary and divide by area, this method is preferred as it accounts for how accessible a neighbourhood is from both within and from outer areas.

Normalization

The data for each neighbourhood were computed using the preprocessed variables (adolescent population, transportation density, existing centres per adolescent population, and median after-tax income). Each variable was then normalized using min-max normalization to give a

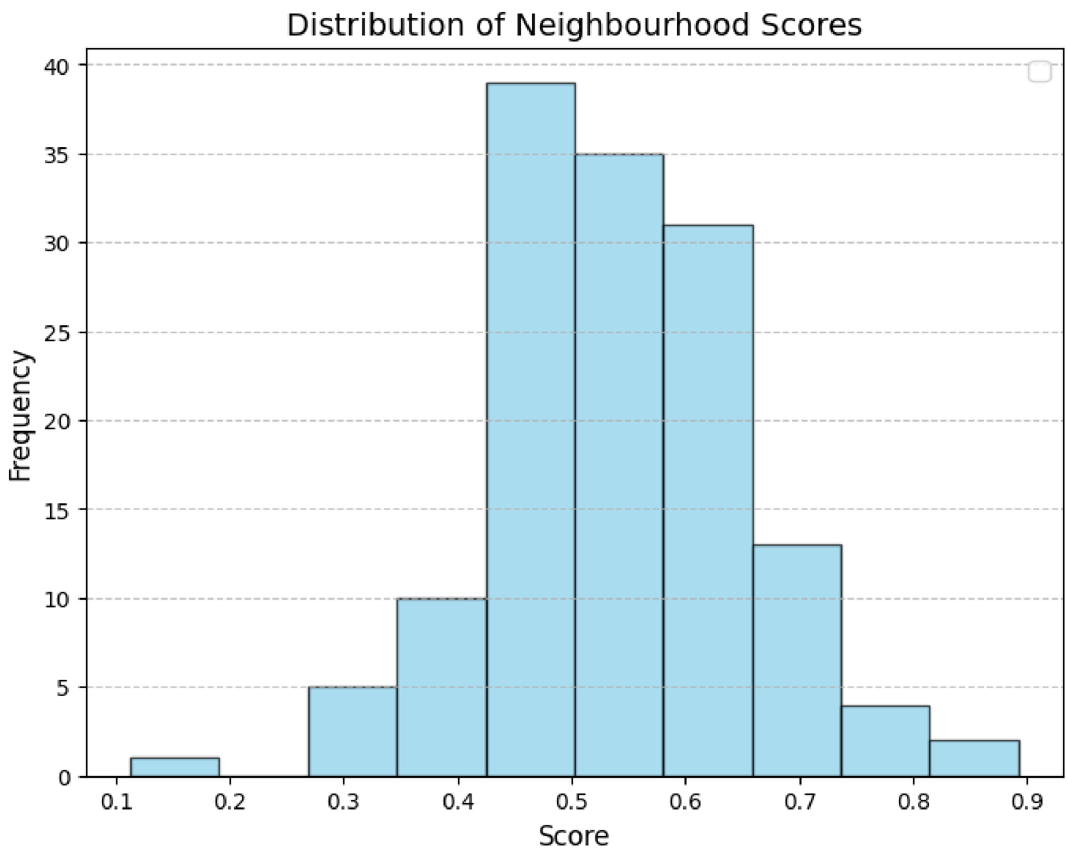

Figure 4. Distribution of scores after data was normalized using a series of differing weights to produce a set of more normalized values to improve data integrity by reducing the impact of outliers and extreme cases.

score between 0 and 1. Each factor score was then multiplied by their respective weighting and added together to get a final score for each neighbourhood. A higher score indicates that the neighbourhood should receive more priority. The model prioritized neighbourhoods with higher adolescent populations and transportation density and gave lower scores to neighbourhoods with a high existing centre per adolescent population and median after-tax income. To inverse the min-max normalized scores for centres per population and income, it was simply subtracted from 1 before being multiplied by the weighting. The factors were assigned different weightings because some factors would likely be considered more than others; adolescent population was weighted 0.4, income was weighted 0.3, facilities per population was weighted 0.2, and transportation density was weighted 0.1.

The adolescent population was given the highest weighting because it directly reflects demand [23-25]. Income was weighted as second highest, given its strong influence on mental health outcomes and service accessibility [26-28]. Facilities per population were weighted above transportation density to emphasize unmet service needs [29].

These weights are adjustable and can be tailored to align with the priorities of mental health organizations using the model.

Clustering

A clustering algorithm was created that could identify groups of close-proximity, high-scoring neighbourhoods. First, we determined the centre point of each neighbourhood and calculated the longitude and latitude of each of those points. The neighbourhoods with a score in the top 70th percentile were considered for the clustering model. Second, the scores were multiplied by weight so that the model prioritizes high-scoring neighbourhoods over geographical proximity. This allows the model to identify large neighbourhoods whose centres may be further apart even if their borders are physically touching. Third, a 3D dataset that contains the longitude, latitude, and weighted score for each neighbourhood was created. Finally, DBSCAN clustering was applied to the dataset, which creates a list of clusters and identifies certain neighbourhoods as noise. DBSCAN was chosen over other algorithms, such as K-Means, because the number of clusters was unknown. Clusters were ordered based on the average neighbourhood score of each cluster, and the top N clusters were selected. A silhouette score was used to analyze whether clusters were well organized and compact, indicating cluster quality.

Location Selection

The cluster of neighbourhoods with the highest priority was further analyzed to identify an optimal location for a new facility. The algorithm used proximity to bus stops to measure transportation efficiency, and schools as a proxy for adolescent population density. Locations that were close to schools and transportation routes were prioritized.

Using an ordered grid of 100 evenly spaced points around the top neighbourhood cluster, each point was analyzed for the number of schools and simulated bus stops within a 500-meter radius. To avoid redundancy, the algorithm avoided locations that were within 300 meters of an existing mental health facility. Bus stops were stimulated by placing points at approximately 350-meter intervals along bus routes, based on a study done on Canadian and American bus stop spacing [30]. The radius located on the point with the greatest number of bus stops and schools was selected to be the recommended location within that neighbourhood cluster.

RESULTS

The neighbourhoods in the top cluster were Thorncliffe Park and Flemingdon Park. Both of these neighbourhoods did not have any existing adolescent mental health services. Thorncliffe Park had approximately median values for after-tax income, adolescent count, and TTC density (51st, 51st, and 53rd percentile respectively). Flemington Park had low income (33rd percentile) and a high neighbourhood adolescent count (87th percentile), but low TTC density (36th percentile). The average normalized score of these two neighbourhoods was in the 95th percentile, and the distance between their centroids was 1.6 km. This confirms that they stand out significantly compared to the other neighbourhood clusters. The next two clusters had normalized scores in the 90th and 87th percentile, with all of these neighbourhood clusters sharing a border. These neighbourhood clusters show comparable features and also stand out; the silhouette score of the DBSCAN clustering algorithm was 0.602, showing that it was decently effective but could still be improved [31]. This score suggests that most neighbourhoods were classified correctly, but there may have been overlapping between clusters or neighbourhoods that were misclassified. The 500-meter radius that the program recommended within

STEM Fellowship Journal

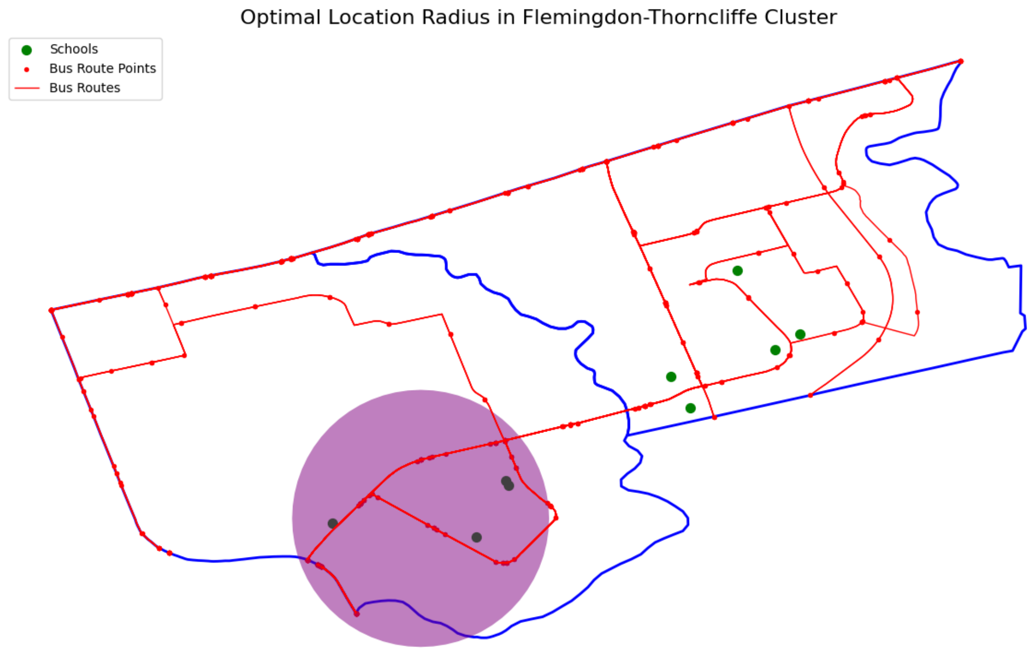

Figure 5. Heatmap of all normalized rating values based on 1) adolescent population, 2) yearly household income, 3) facilities per population, and 4) public transportation density. The prioritized areas were outlined in purple, blue, and green respectively for the 1st, 2nd, and 3rd priorities.

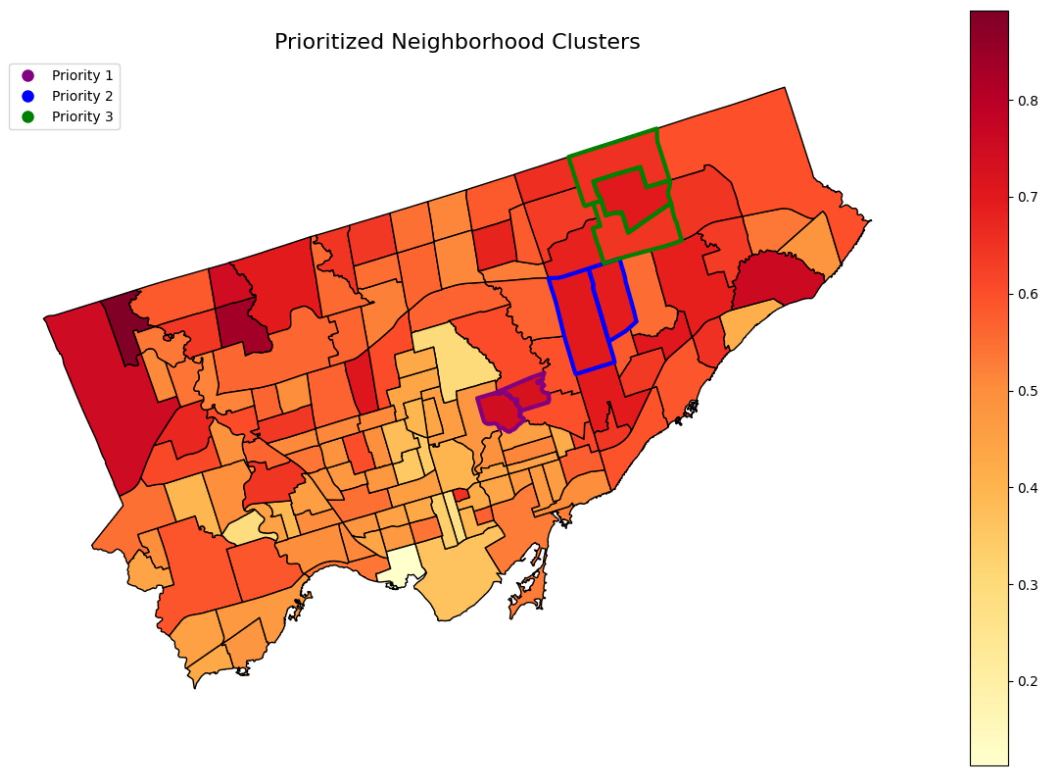

Figure 6. Representation of the optimized area to place a mental health service in the Flemingdon-Thorncliffe cluster. The algorithm prioritizes accessibility and adolescent population density.

the top cluster was located in lower Thorncliffe Park. This was the location with the greatest number of simulated bus stops and schools within a 500-meter radius.

DISCUSSION

Interpretation and Policy Recommendation

Our study further confirms that access to mental health services in Toronto faces strong disparities, highlighted by groups of high-scoring neighbourhoods that are condensed in certain parts of the city: mostly Scarborough, Northern Etobicoke, and East York. Our research also found three key neighbourhood clusters that were optimal for mental health centre placement, the top one being Flemingdon Park and Thorncliffe Park. Although some areas of Etobicoke had very high scores, they were not considered optimal, likely because their centroids were further apart than other neighbourhoods with similar or even lower scores. This would explain why the neighbourhood clusters that our model chose consisted of smaller neighbourhoods with closer centroids. Still, prioritizing smaller neighbourhoods could be beneficial, because neighbourhoods that have closer centroids are easier to travel between, service placement in said neighbourhoods would be more accessible compared to a new service location in larger neighbourhoods. The 500-meter radius selected in lower Thorncliffe Park appears to cover less bus stops and schools than central Flemingdon Park, but in reality, there are numerous overlapping bus routes in that area, meaning more bus stops that are very close in proximity to each other.

We recommend that policy makers prioritize Thorncliffe Park and Flemingdon Park for resource allocation–in particular, the placement of a new mental health service that serves adolescents. Not only are the areas highlighted as clusters by our algorithm, but they are also validated as significant due to their high adolescent populations and lack of existing services. Additionally, their median income percentiles are generally low, and their transportation accessibility is decent. The radius in Thorncliffe Park in particular is centered around many overlapping bus routes and schools, meaning it is highly accessible for adolescents to travel to. Policymakers may also assign their own specific weightings for each factor depending on what they deem most important, and integrate real-time data for improved accuracy.

Comparison

Wang et al. analyzed spatial accessibility to mental health services in Toronto and found that the peripheral neighbourhoods of Scarborough, East York, and Etobicoke had lower accessibility scores [10]. This closely aligns with our heatmap of service needs, which found that similar areas are hotspots for youth mental health service needs, thereby validating the weightings for each neighbourhood need factor. Despite this, the top three neighbourhood clusters from our algorithm were all located in Scarborough and East York, despite the statistical significance of Etobicoke being higher in Wang et al.'s study [10]. The reason for this difference could be that Wang and colleagues analyzed all mental health services, including those for adults, while our study focused on adolescents specifically [10].

Limitations

There were several limitations to our study related to data and calculation methods. First, the TTC data was collected in 2019, and significant extensions to the Vaughan line have been made since. However, more recent data was not available, so this limitation applies broadly to any research using TTC data. Secondly, the majority of the data was sourced from Toronto, therefore, it would be difficult to extrapolate that the model's weightings would remain valid in a different environment. Lastly, the method we used to calculate transportation density is not all-encompassing, as it is only a good measure of inter-neighbourhood accessibility. Using the alternative method to measure transportation density, using the ratio between km of road and area in km2, would also not be all-encompassing due to the possibility that the closest venue is located in another neighbourhood. There are also confounding factors such as the frequency of buses on a given route, timeliness of buses on that route that are omitted due to lack of data. Lastly, location-specific recommendations for new mental health support services are suggested only as a 500m radius, meaning that within the suggested area, variability remains as to where said centre should be established. This may reduce the model's effectiveness, depending on the land availability and budget constraints faced by implementing agencies.

STEM Fellowship Journal

Future Work

Some mental health service providers in our dataset are not open to all demographics or have strict limitations on specific demographics. A dataset of non-profit organisations that serve all adolescents in Toronto would be better suited for our study, but to our knowledge, does not exist. This dataset could be compiled through web scraping. The location placement algorithm should take into account which areas can have construction and buildings, in order to prevent placement in an area that is not suitable such as green space or housing.

CONCLUSION

As the global mental health crisis reaches an all-time high, our study demonstrates the potential of geospatial analyses algorithms in addressing disparities within access to mental health support for adolescents in Toronto. By integrating socioeconomic indicators, population demographics, and transportation density, our model identifies high-priority neighbourhoods where mental health facilities could have the greatest impact.

Our model provides a replicable framework with adjustable weightings that allow urban planners and policymakers to optimize resource distribution equitably and efficiently. The inclusion of dynamic elements, such as public transit accessibility, ensures that our recommendations are not only equitable but also practical and implementable.

While this study attempts to highlight the efficacy of big data in public health planning, it also identifies areas for further refinement such as more recent data, real-time demographic changes, and community feedback. Moreover, the methodology is highly scalable and applicable to other metropolitan areas facing similar challenges. Finally, our data is consistent with other studies focusing on spatial equality to mental health in Toronto. By bridging gaps in infrastructure and fostering greater equity in mental health service provisioning, we hope to contribute to a healthier and more inclusive future for youth in Toronto.

ACKNOWLEDGEMENTS

We would like to thank the Big Data Challenge for giving us the opportunity to engage in this research, as well as for providing the resources and educational tools we needed to implement our idea. We would like to thank Mr. Stephens for organizing the Statistics Club at Crescent and for giving us advice and resources. We would also like to thank Mr. Daniel Nucci and our peers David Shan and James Tian for providing us with valuable suggestions for our research. Finally, thanks to all our friends and family for making this experience possible through their support.

REFERENCES

1. Beers CW. A mind that found itself: An autobiography. Longmans, Green; 1917.
2. Schlack R, Peerenboom N, Neuperdt L, Junker S, Beyer A-K. The effects of mental health problems in childhood and adolescence in young adults: Results of the KIGGS cohort. J Health Monitor. 2021;6(4):3.
3. Parry M. From a patient's perspective: Clifford Whittingham Beers' work to reform mental health services. Am J Public Health. 2010;100(12):2356-2357.
4. Cairney J. The mental health of children and youth in Ontario. desLibris; 2015.
5. Georgiades K, Duncan L, Wang L, Comeau J, Boyle MH, 2014 Ontario Child Health Study Team. Six-month prevalence of mental disorders and service contacts among children and youth in Ontario: evidence from the 2014 Ontario Child Health Study. Can J Psychiatry. 2019;64(4):246-255.
6. Blakemore S-J. Adolescence and mental health. Lancet. 2019;393(10185):2030-2031.
7. Kourgiantakis T, Markoulakis R, Lee E, Hussain A, Lau C, Ashcroft R, et al. Access to mental health and addiction services for youth and their families in Ontario: perspectives of parents, youth, and service providers. Int J Ment Health Syst. 2023;17(1):4.
8. Law J, Perlman C. Exploring geographic variation of mental health risk and service utilization of doctors and hospitals in Toronto: a shared component spatial modeling approach. Int J Environ Res Public Health. 2018;15(4):593.
9. Steele LS, Glazier RH, Agha M, Moineddin R. The gatekeeper system and disparities in use of psychiatric care by neighbourhood education level: results of a nine-year cohort study in Toronto. Healthcare Policy. 2009;4(4):e133.
10. Wang L, Ariwi J. Mental health crisis and spatial accessibility to mental health services in the city of Toronto: A geographic study. Int Health Trends Perspect. 2021;1(2):191-213.
11. Tadmon D, Bearman PS. Differential spatial-social accessibility to mental health care and suicide. Proc Natl Acad Sci. 2023;120(19):e2301304120.
12. Kirkbride JB, Anglin DM, Colman I, Dykxhoorn J,

Jones PB, Patalay P, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry. 2024;23(1):58.

13. Maconick L, Sheridan Rains L, Jones R, Lloyd-Evans B, Johnson S. Investigating geographical variation in the use of mental health services by area of England: a cross-sectional ecological study. BMC Health Serv Res. 2021;21:1-10.

14. Noorain S, Scaparra MP, Kotiadis K. Mind the gap: a review of optimisation in mental healthcare service delivery. Health Syst. 2023;12(2):133-166.

15. Lyeo JS, Tiznado-Aitken I, Farber S, Brown HK, Spence N. Predictors of transportation-related barriers to healthcare access in a North American suburb. J Public Health. 2024;32(8):1359-1370.

16. Heinz A, Deserno L, Reininghaus U. Urbanicity, social adversity and psychosis. World psychiatry. 2013 Oct;12(3):187-97.

17. City of Toronto. Regional municipal boundary, 2019. Accessed 2025 Jan 10.

18. City of Toronto. Neighbourhoods, 2024. Accessed 2025 Jan 10.

19. City of Toronto. TTC routes and schedules, 2024. Accessed 2025 Jan 10.

20. City of Toronto. TTC subway shapefiles, 2019. Accessed 2025 Jan 10.

21. City of Toronto. Wellbeing youth mental health, 2024. Accessed 2025 Jan 10.

22. City of Toronto. Neighbourhood profiles, 2022. Accessed 2025 Jan 10.

23. National Academies of Sciences E, Division H and M, Education D of B and SS and, Board on Children Y, Applications C on the N and SS of AD and I, Backes EP, et al. Health System [Internet]. www.ncbi.nlm.nih.gov. National Academies Press (US); 2019 [cited 2022 Oct 10]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545489/

24. Cadigan JM, Lee CM, Larimer ME. Young adult mental health: A prospective examination of service utilization, perceived unmet service needs, attitudes, and barriers to service use. Prevention Science. 2019 Apr 15;20(3):366-76.

25. Mori Y, Sourander A, Mishina K, Ståhlberg T, Klomek AB, Kolaitis G, Kaneko H, Li L, Huong MN, Praharaj SK, Kyrrestad H. Unmet need for mental health care among adolescents in Asia and Europe. European Child & Adolescent Psychiatry. 2024 Dec;33(12):4349-59.

26. Ravensbergen SJ, Bouter DC, de Neve-Enthoven NG, Hagestein-de Bruijn C, Hoogendijk WJ, Grootendorst-van Mil NH. Low household income and adolescent mental health. Development and Psychopathology. 2025:1-2.

27. Zimmerman FJ. Social and economic determinants of disparities in professional help-seeking for child mental health problems: Evidence from a national sample. Health services research. 2005 Oct;40(5p1):1514-33.

28. Steele L, Dewa C, Lee K. Socioeconomic status and self-reported barriers to mental health service use. The Canadian Journal of Psychiatry. 2007 Mar;52(3):201-6.

29. Lankila T, Laatikainen T, Wikström K, Linna M, Antikainen H. Association of travel time with mental health service use in primary health care according to contact type—a register-based study in Kainuu, Finland. BMC Health Services Research. 2022 Nov 30;22(1):1458.

30. Devunuri S, Lehe LJ, Qiam S, Pandey A, Monzer D. Bus stop spacing statistics: Theory and evidence. J Public Transp. 2024;26:100083.

31. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53-65.