Data in Brief 57 (2024) 110877

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Machine learning-ready mental health datasets for evaluating psychological effects and system needs in Mexico city during the first year of the COVID-19 pandemic

Carlos Rodrigo Garibay Rubio[a,*], Katsuya Yamori[b], Genta Nakano[b], Astrid Renneé Peralta Gutiérrez[c], Silvia Morales Chainé[d], Rebeca Robles García[e], Edgar Landa-Ramírez[f], Alexis Bojorge Estrada[g], Alejandro Bosch Maldonado[h], Diana Iris Tejadilla Orozco[i]

[a] *Graduate School of Informatics, Kyoto University, Yoshidahonmachi, Sakyo Ward, Kyoto 606-8317, Japan*
[b] *Disaster Prevention Research Institute, Gokasho, Uji, Kyoto 611-0011, Japan*
[c] *Independent researcher, Gokasho Sanbanwari 37, Obaku shiei jutaku, Uji, Kyoto 611-0011, Japan*
[d] *Faculty of Psychology, National University of Mexico, Circuito Ciudad Universitaria Avenida, C.U., 04510 Ciudad de México, Mexico*
[e] *Research Center for Global Mental Health, National Institute of Psychiatry "Ramón de la Fuente Muñiz", Calz México-Xochimilco 101, Colonia, Huipulco, Tlalpan, 14370 Ciudad de México, CDMX, Mexico*
[f] *Ministry of Health, "Hospital General Dr. Manuel Gea González", Calz. de Tlalpan 4800, Belisario Domínguez Secc 16, Tlalpan, 14050 Ciudad de México, CDMX, Mexico*
[g] *Ministry of Health, Psychiatric Services, Av. Marina Nacional 60, Tacuba, Miguel Hidalgo, 11410 Ciudad de México, CDMX, Mexico*
[h] *General Directorate of Community Attention, National Autonomous University of México, 04510 Mexico City, CDMX, Mexico*
[i] *Ministry of Health, Child Psychiatric Hospital "Juan N Navarro" Av. San Fernando 86, Belisario Domínguez Secc 16, Tlalpan, 14080 Ciudad de México, CDMX, Mexico*

ARTICLE INFO

Article history:
Received 9 April 2024
Revised 7 June 2024
Accepted 19 August 2024
Available online 28 August 2024

ABSTRACT

The prevalence of mental health problems constitutes an open challenge for modern societies, particularly for low and middle-income countries with wide gaps in mental health support. With this in mind, five datasets were analyzed to track mental health trends in Mexico City during the pandemic's first year. This included 33,234 responses to an

* Corresponding author.
E-mail address: garibay.rodrigo.42s@st.kyoto-u.ac.jp (C.R. Garibay Rubio).

https://doi.org/10.1016/j.dib.2024.110877

Dataset link: Microtrends of mental health reactions for better decision making Generals (Original data)
Dataset link: Mental_health_cdmx_2020 (Original data)

Keywords:
Disaster recovery curve
Psychological response in disasters
Mental health support systems
Stress response in emergencies
Acute stress response
Earthquake early warning
Pandemic mental health effects

online mental health risk questionnaire, 349,202 emergency calls, and city epidemiological, mobility, and online trend data.

The COVID-19 mental health risk questionnaire collects information on socioeconomic status, health conditions, bereavement, lockdown status, and symptoms of acute stress, sadness, avoidance, distancing, anger, and anxiety, along with binge drinking and abuse experiences. The lifeline service dataset includes daily call statistics, such as total, connected, and abandoned calls, average quit time, wait time, and call duration. Epidemiological, mobility, and trend data provide a daily overview of the city's situation.

The integration of the datasets, as well as the preprocessing, optimization, and machine learning algorithms applied to them, evidence the usefulness of a combined analytic approach and the high reuse potential of the data set, particularly as a machine learning training set for evaluating and predicting anxiety, depression, and post-traumatic stress disorder, as well as general psychological support needs and possible system loads.

Specifications Table

Subject	Health and Medical Sciences / Data Science
Specific subject area	Psychiatry and Mental Health / Applied Machine Learning
Type of data	Raw data (Excel and CSV files), Translation and File structure (PDF) Analysed (CSV files), Parsing, Preprocessing, Optimization, Dimensionality Analysis, Machine learning applications and Validation (Python 3 files), Processed and visualization (Jupyter Notebook file). Processed and visualization (Tableau public viz).
Data collection	The online questionnaire is publicly available on the internet through the web page https://www.misalud.unam.mx/, and was disseminated through mass media as part of the COVID-19 countermeasures. The period covered consecutively from 13/04/2020 to 07/12/2020, and both the original language and its English translation can be found in a file within the reservoir. The Lifeline emergency phone call service set reports the calls received and attended by mental health emergency phone operators, and it is part of the archives of the Ministry of Health. The dataset retrieved the calls from May 24, 2020, to December 31, 2020, in a consecutive manner, with additional +−72 hours' time periods surrounding earthquake events with public early warnings between 13/04/2020 and 10/04/2023 resulting in 280 days between consecutive and non-consecutive days). The epidemiological data set contains gender, death from COVID-19, suspected and confirmed cases, and a date. It was retrieved from the Ministry of Health (Mexico). Mobility (https://www.google.com/covid19/mobility/) and popularity trends (https://trends.google.com/trends/) were retrieved from Google services.
Data source location	The dataset of the questionnaire for the detection of risks to mental health COVID-19 corresponds to Mexico City in Mexico, and it was originally stored in the servers of the General Direction for Community Attendance (DGACO) of the National University of Mexico (UNAM) (General Directorate of Community Attention, National Autonomous University of México, Mexico City 04510,). The

(*continued on next page*)

	data set of the lifeline emergency phone call service, as well as the epidemiological records, are stored in servers of the Ministry of Health of Mexico (*Av* Costera Miguel Alemán 276, Hornos, 39355 Acapulco de Juárez, Gro., Mexico), while the city mobility and popularity trends datasets were retrieved from Google services. For convenience, all raw and processed datasets, as well as the file structure, parsing, preprocessing, machine learning coding, translation in English, and other analysis files mentioned here, were made available publicly at a Zenodo repository (doi:DOI:10.5281/zenodo.11501598) linked to a Github release named "DiB_v3.0", and an interactive file can be accessed through a tableau public viz named "Microtrends of mental health reactions for better decision making Generals". Analyzes were performed at Kyoto University address: Yoshidahonmachi, Sakyo Ward, Kyoto, 606-8501
Data accessibility	**Repository name**: Mental_health_cdmx_2020 **GitHub release:** https://github.com/ro-riskreduction/Mental_health_cdmx_2020/releases/tag/DiB **Zenodo reservoir:** https://zenodo.org/records/11501598 **DOI:** 10.5281/zenodo.11501598 **Repository name Tableau Public**: Microtrends of mental health reactions for better decision making Generals Data identification number: NA Direct URL to data: https://public.tableau.com/app/profile/c.rodrigo.garibay.rubio/viz/MicrotrendsofmentalhealthreactionsforbetterdecisionmakingGenerals/GeneralResults?publish=yes The dataset instructions are referred to in the README file as well as in the File structure file.

1. Value of the Data

- *Global Trends and Resource Optimization:* Global trends in disaster frequency highlight the critical need for high-quality data to optimize resources and provide timely and adequate mental health support, especially in low—and middle-income countries. The datasets retrieved contribute to this objective.
- *Valuable Information from Data Sets:* The information from the datasets can support mental health strategies, aid in decision-making, and predict future stress-related disorder prevalence during other disaster events.
- *Machine Learning Suitability:* The datasets are highly suitable for machine learning operations. Scaling and normalization strategies have been explored, but they have yielded no further improvement.
- *Integrated Perspective on Mental Health Effects:* By integrating different databases, decision-makers gain a comprehensive perspective on the mental health effects within the city. The interactive integrative tools presented illustrate this.
- *Coding, Replication, and Flexibility:* The provided code alongside the dataset enables easy replication, validation, and visualization. It also offers flexibility for conducting further analysis on these or other data sets.

2. Background

The COVID-19 pandemic has brought about significant changes in general livelihoods and mental health worldwide, necessitating the identification and support of those in need. This urgency is particularly pronounced in middle- and low-income countries, where the mental health gap between perceived support needs and actual access to it remains high, as reported by the World Health Organization [1]. Stress and disaster frameworks posit that, under these conditions, a substantial portion of the population will experience normal psychological effects over a short period, while a smaller segment is at risk of developing long-lasting psychological disorders if left unattended. Therefore, efficient and rapid assessment is crucial.

The datasets themselves offered valuable information for understanding prevalent mental health conditions and were used to support individuals through cognitive-behavioral self-help strategies and phone-based attention. Simultaneously, integrating epidemiological, mobility, and information-seeking trends datasets assisted decision-makers in refining ongoing mental health strategies by providing a comprehensive view. Furthermore, these machine learning-ready datasets are useful for predicting the prevalence of stress-related disorders during future disaster events.

3. Data Description

3.1. Datasets description

The questionnaire for the detection of risks to mental health COVID-19 was a tool designed by the National University of Mexico and the National Institute of Psychiatry of Mexico, to help with the pandemic and was promoted through mass media (TV conferences), social media, and institutional dissemination in the form of an electronic survey composed of 100 questions to be completed in 15–20 min. It was available electronically for anyone who wanted to monitor their mental health condition and receive appropriate support depending on its results. The dataset retrieves all cases reported for Mexico City for the period between April 13, 2020, and December 7, 2020. The loss of answers due to partial responses was minimal (239 out of 36,721 raw).

The raw dataset comprised 36,721 applications, while the tidy and processed dataset available consisted of 33,234 applications once underage users were removed. The questionnaire covers several sections, including socioeconomic self-report, present chronic health conditions, COVID-19 symptoms and diagnosis, loss of loved ones, lockdown status, and a mental health screen with 27 questions. These questions can be categorized into acute stress (seven items), avoidance and sadness (five items), distancing and anger (seven items), generalized anxiety (four items), health-related anxiety (four items), and additional factors: binge alcohol consumption (three items) and experience of abuse (two items). Respondents answered using a Likert format, selecting a response from a range between 0 and 10 integers [2].

The “Questionnaire for the Detection of Risks to Mental Health during COVID-19” is based on previous instruments [3–8] and implemented during the pandemic (corresponding to the dataset in “analysis_clean.csv”), used to evaluate the psychological condition of the users (for an in-depth review of the development of the questionnaire, see [9]). Sections 4.2.2 and 4.2.3 of the present article address the mechanisms used for scaling and creating the compound variables referencing psychological problematics.

The questionnaire has a Cronbach alpha of 0.96, commonalities $>$ 0.30 for all items, and an explained variance of 64 %. All scales related to mental health problematics behaved adequately (RMSEA $\langle$0.080, SRMR $<$ 0.06, CFI y TLIs$\rangle$ 0.950, and alphas $>$ 0.60), as reported in a previous publication [2], mentioning adequate psychometric properties for acute stress, avoidance and sadness, distancing and anger, as well as for generalized anxiety.

The questionnaire items, together with their translation, reference to its columns within the main file (“analysis_clean.csv”), and anchors to the response options, can be found in the reservoir file (Mental_health_cdmx_2020/Databases and coding/Databases/Translation of questions answers, and encoding.pdf).

The dataset retrieved from the Lifeline Emergency Phone Call Service comes from their record and report system, consisting of 349,202 calls, grouped daily between April 24, 2020, and December 31, 2020, consecutively, and then $+-72$ h periods surrounding earthquakes with early warnings, in Mexico City during 2021, and 2022 up to 6/04/2023. The vectors within this data set represent the number of calls (enterqueue), connected calls, number of abandoned attempts, average time (in seconds) before quitting an attempt, average waiting time for connection, and average time of completed calls.

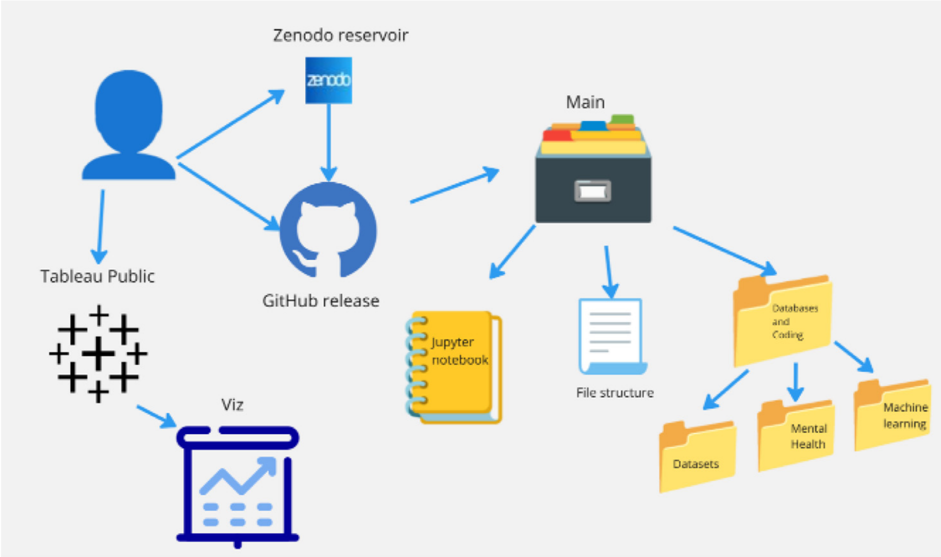

Fig. 1. Structure of the information, files, and datasets.

This dataset can provide insights into the relationships between mental health demands over time and correlations with other vectors, which can be useful for determining thresholds of system overflow. In particular, this dataset is well-suited for unsupervised learning.

The epidemiological effects dataset reports daily numbers of confirmed COVID-19 cases, suspected cases, and death tolls between April 13, 2020, and December 7, 2020.

The popularity trends, derived from news appearances and web searches retrieved from Google Trends (available at this link https://trends.google.com/trends/), cover the period between April 13, 2020, and December 7, 2020. These trends focus on five key concepts: COVID-19, quarantine, epidemiological traffic light (which establishes appropriate health security measures for work, educational activities, and public spaces), pandemic, and mental health. Popularity is measured daily on a 0–100 scale relative to its maximum value during the specified period.

The mobility results, obtained from this link https://www.google.com/covid19/mobility/, evaluate the percent change from baseline for various locations between April 1, 2020, and December 31, 2020. These locations include residences, workplaces, transit stations, parks, grocery and pharmacies, and retail and recreation places.

3.2. Structure of information

The data description and analysis of the data sets are guided through two interactive pathways: a Jupyter notebook and a Tableau public viz, which guide the reader through the datasets. Complementary a "File structure" file maps the file locations and provides a short description of the use of each one of the files (Fig. 1).

All datasets are located in the subfolder "Databases and coding/Databases," whose analysis got integrated through a Jupiter notebook and a Tableau Public viz.

3.3. Jupyter notebook with machine learning applications

The Jupyter notebook titled "Database_machine_learning_applications.ipynb" available through the GitHub reservoir (https://github.com/ro-riskreduction/Mental_health_cdmx_2020/tree/main) [10], guides the reader through the packages, modules, coding, analysis, and visualizations (files located in Databases and coding/ Coding_Database_machine_learning_applications)

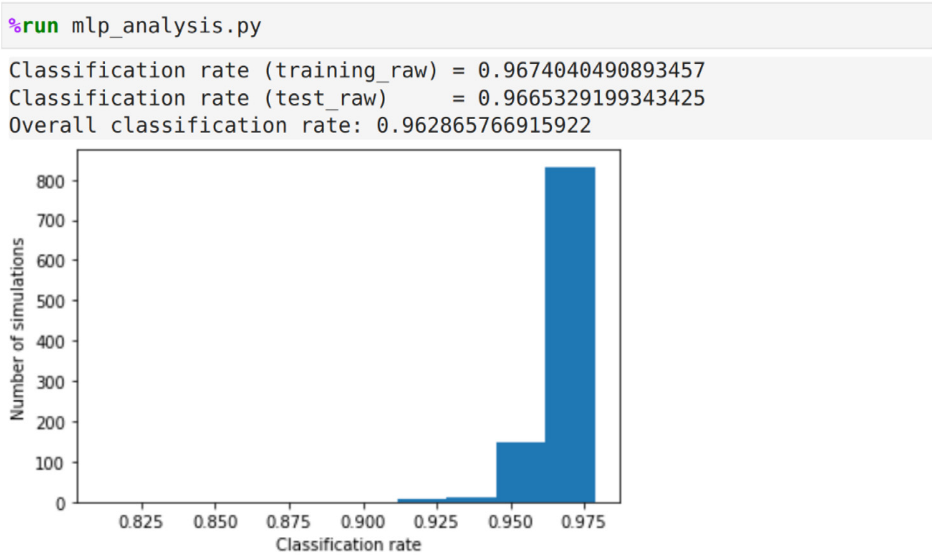

Fig. 2. Example showing the organizing structure for each section inside the main Jupyter Notebook file.

useful for understanding the datasets of the questionnaire for the detection of risks to mental health COVID-19 and the lifeline phone service database in an easy and integrated manner.

The Jupyter Notebook structure consistently begins with an objective section, followed by references to the relevant scripts, and closes with results and/or visualizations (Fig. 2). Within the notebook, sections cover parsing, descriptive statistics, database distribution, item dispersion (with a focus on analyzing psychological effects related to PTSD, depression, and anxiety disorders), time distributions, and various analyses (including principal component analysis, one-way ANOVA, and percentage change analysis). Additionally, the notebook includes machine learning steps such as preprocessing strategies, scaling, mapping approaches, and supervised and unsupervised algorithms, along with validation.

3.4. Tableau public with interactive trend analysis

Concurrently to the Jupyter Notebook, the Tableau Public visualization available at this link https://t.ly/VJbcq provides an in-depth exploration of the descriptive data from the "Questionnaire for the Detection of Risks to Mental Health during COVID-19". This visualization offers an integrated and interactive view of the interrelationships among the questionnaire responses, the lifeline phone service, epidemiological data, mobility patterns, quarantine adherence, and trending results related to critical concepts during the pandemic (Fig. 3). Furthermore, it highlights differences between information-seeking and avoidance behaviors.

The Tableau Public visualization enables real-time analysis of psychological trends—focusing on anxiety, avoidance, and acute stress—for the fourth quartile of the sample, representing the most heavily affected segment.

3.5. Significance of the datasets

The data sets play a crucial role in understanding psychological changes during the pandemic in a highly urbanized area of a middle-income country, specifically Mexico City. These datasets

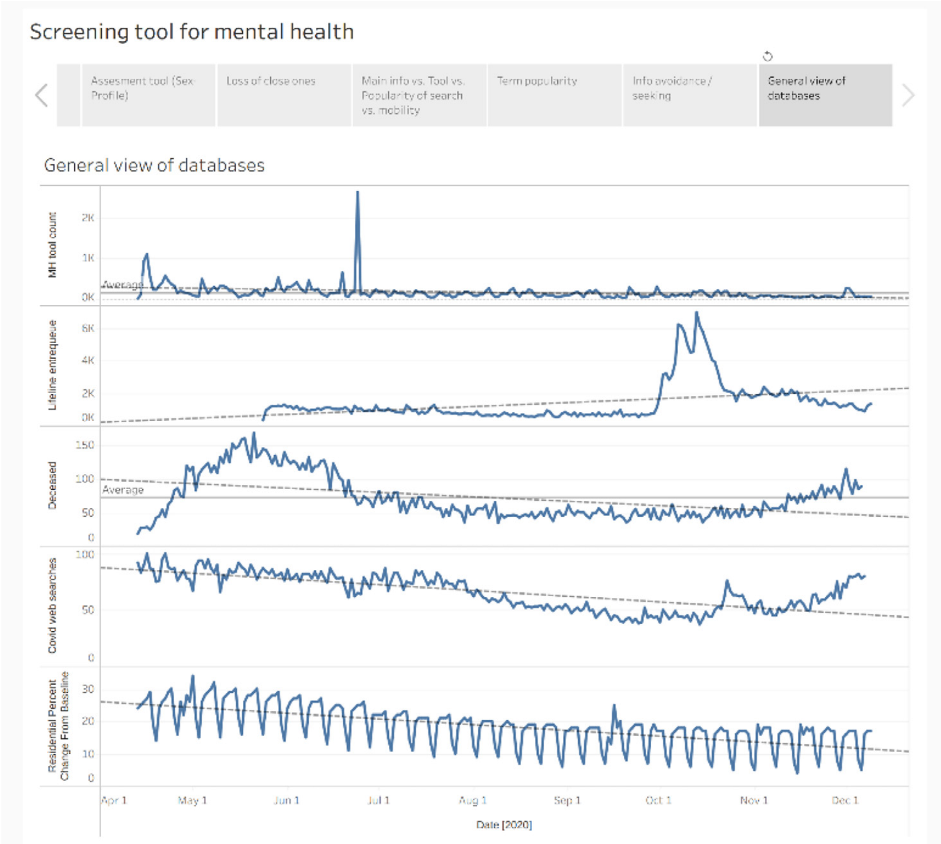

Fig. 3. Interaction of the different data sets using days as the X-axis on Tableau public viz.

shed light not only on the mental health effects but also on the demands made to the first-level mental health response systems. Notably, this is the first time that the “Questionnaire for the Detection of Risks to Mental Health during COVID-19” dataset and the “Lifeline Emergency Phone Call Service” have been integrated and made publicly available for open research. Remarkably, there are no existing records of the integration of similar databases in Mexico City.

4. Experimental Design, Materials and Methods

4.1. Data gathering

Mexico City, one of the world’s largest cities, experienced a significant impact from the COVID-19 pandemic, resulting in diverse psychological effects on its population, which changed rapidly over time. To identify individuals at risk of developing anxiety, depression, or post-traumatic stress disorder and to provide timely cognitive-behavioral self-help and direct mental health support, various first-level mental health strategies were developed. Among these, the “Questionnaire for the Detection of Risks to Mental Health COVID-19” and the “Lifeline Emergency Phone Call Service” emerged as central components.

The “Questionnaire for the Detection of Risks to Mental Health COVID-19” is an online application system, accessible via the link here https://www.misalud.unam.mx/. Based on individual scores, it triggered the delivery of cognitive-behavioral self-help strategies (including videos, infographics, training courses, and relaxation activities), and in cases with high psychological risk, a phone contact number for mental health support was provided, along with an option for immediate contact with a professional for brief psychological intervention. The responses constituted the elements of the dataset.

The "Lifeline Emergency Phone Call Service" is a call center supervised by the Ministry of Health, accessible via the number "01 800 911 2000," which offers guidance and psychological phone interventions when needed, as well as referrals to specialized mental health support systems. Its daily reports were retrieved and incorporated into the dataset.

The epidemiological information, sourced from the entire health sector in Mexico City, was publicly available and updated daily during the COVID-19 pandemic. This information provides a current and comprehensive view of the situation. An up-to-date visualization of this data can be found at https://www.imip.org.mx/imip/node/129[1].

The mobility and trends datasets were retrieved from Google services for the selected period and integrated into the overall analysis in a novel manner. This integration served to crosscheck general quarantine adherence and evaluate information-seeking behaviors related to concepts directly associated with the emergency. Specifically, this relates to the information-seeking and avoidance components of the "Questionnaire for the Detection of Risks to Mental Health COVID-19".

4.2. Analysis and machine learning possibilities

4.2.1. Supervised learning

The "Questionnaire for the Detection of Risks to Mental Health COVID-19" dataset provides meaningful insights into mental health and is well-suited for evaluating short-term psychological responses within the community. It allows to evaluate the effects of events such as changes in public policies, pandemic measures, earthquake alerts, and festivities. Patterns emerge through statistical analysis (including ANOVA, variations in percentual changes, and principal component analysis), making this dataset useful for training machine learning classifiers in supervised and unsupervised settings.

Furthermore, the dataset was enriched with labels (composite variables columns: DX (PTSD), DQ (Anxiety), and DR (Depression)) based on response thresholds for different psychological disorders identified in the questionnaire, including post-traumatic stress disorder, anxiety, and depression.

4.2.2. Scaling

The scaling code is presented in Tables 1, 2, and 3:

Post-traumatic stress disorder

Anxiety

Depression

4.2.3. Evaluation criteria

The following diagnostic criteria were implemented for developing symptom labels:

1) The PTSD evaluation was done according to DSM-5 diagnostic criteria as follows [11].
   A. Exposure to stressful event
   B. One or more items of questions 37 to 40 with answers of 2 or above.
   C. One or more items of questions 41 to 42 with answers of 2 or above.
   D. Two or more items of questions 43 to 48 with answers of 2 or above.
   E. Two or more items of questions 49 to 53 with answers of 2 or above.

[1] Requires vpn for access outside of Mexico.

Table 1
Scaling code for PTSD according to the DSM-V.

	Not at all Code=0	A bit Code=1	Moderately Code=2	A lot Code=3	Extremely Code=4
I repeatedly imagine or think I am going to get sick	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have repeated nightmares about this illness	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I feel worried when people mention this illness	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have unpleasant physical reactions when I think about this illness (such as strong heartbeats, breathing problems, sweating)	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I avoid thinking, feeling or talking about this illness	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I avoid seeing or reading official information about this illness	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have difficulty remembering what the authorities recommend for coping with risk or illness	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I think that if I or someone in my family gets sick it is my fault	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have lost interest in activities I used to enjoy.	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have felt distant from the people I spend time with, since the risk of this illness began	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have difficulty feeling affection for those I love	0	1, 2, 3	4, 5, 6	7, 8,9	10
I feel my future is uncertain since the risk of suffering from this illness began	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I feel I want to do things to harm myself	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I have difficulty falling or staying asleep (Code insomnia if the answer is four or more)	0	1, 2, 3	4, 5, 6	7, 8,9	10
I feel angry	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I find it hard to pay attention	0	1, 2, 3	4, 5, 6	7, 8, 9	10
I feel frightened about the risk of suffering from this illness	0	1, 2, 3	4, 5, 6	7, 8, 9	10

Table 2
Scaling for anxiety according to the 5-item Anxiety Scale for ICD-11 PHC field studies.

	Never	Several days	Over half the days	Nearly every day
	Code=0 (no)			Code=1 (yes)
I feel nervous, anxious or about to burst	0	1,2,3,4	5,6,7,8	9,10
I have felt unable to control my worries	0	1,2,3,4	5,6,7,8	9,10
I have felt so worried I have been unable to keep still	0	1,2,3,4	5,6,7,8	9,10
I have found it hard to relax	0	1,2,3,4	5,6,7,8	9,10
I have felt scared something terrible was going to happen	0	1,2,3,4	5,6,7,8	9,10

Table 3
Scaling for depression according to the Patient Health Questionnaire-2 (PHQ-2).

	Never	Several days	Over half the days	Nearly every day
	Code=0	Code=1	Code=2	Code=3
I have felt little interest or pleasure in doing things	0	1,2,3,4	5,6,7,8	9,10
I have felt down, depressed or hopeless	0	1,2,3,4	5,6,7,8	9,10

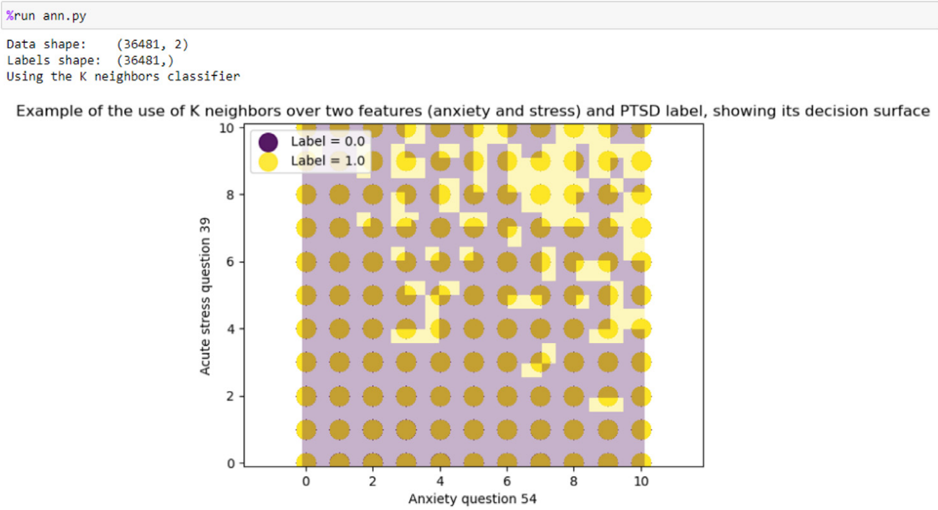

Fig. 4. Simple classification algorithm with decision surface for PTSD.

Table 4
Classification rates of a multi-layer perceptron classifier applied for post-traumatic stress disorder labels.

Strategy	Classification rate mean running 1000 iterations with random test samples of 0.3 size using MLP algorithm with 10,000 iterations of the neural network
Maintaining the data without transformations	0.897
Normalizing the data	0.668
Scaling using the standard scaling algorithm	0.910
Scaling using the robust scaling algorithm	0.909

2) A generalized anxiety label was ascribed if a total score of three or more results from the sum of the items [4].
3) For depression, the recommended cut-off point was a score of three or over based on the sum of the items [12].

4.2.4. Classification example

Based on the provided labels, a classification task using supervised machine learning algorithms can be performed, effectively mapping the relationship between features. Fig. 4 shows a straightforward application of the K Nearest Neighbours (KNN) algorithm and its decision surface.

4.2.5. Preprocessing optimization

Regarding the preprocessing steps for the “Questionnaire for the Detection of Risks to Mental Health COVID-19”, Table 1 reveals that the dataset does not benefit from applying further optimization strategies. In fact, such strategies only reduce the classification rates during machine learning training. A multi-layer perceptron classifier was used as a validation mechanism (Table 4).

4.2.6. Unsupervised learning

The dataset related to the Lifeline Emergency Phone Call Service is well-suited for unsupervised learning, such as using the Support Vector Machine (SVM) to compare real versus predicted increases in calls. Significant events, like strong earthquakes, tended to trigger higher predicted values than regular events, which is interesting for further study (Fig. 5).

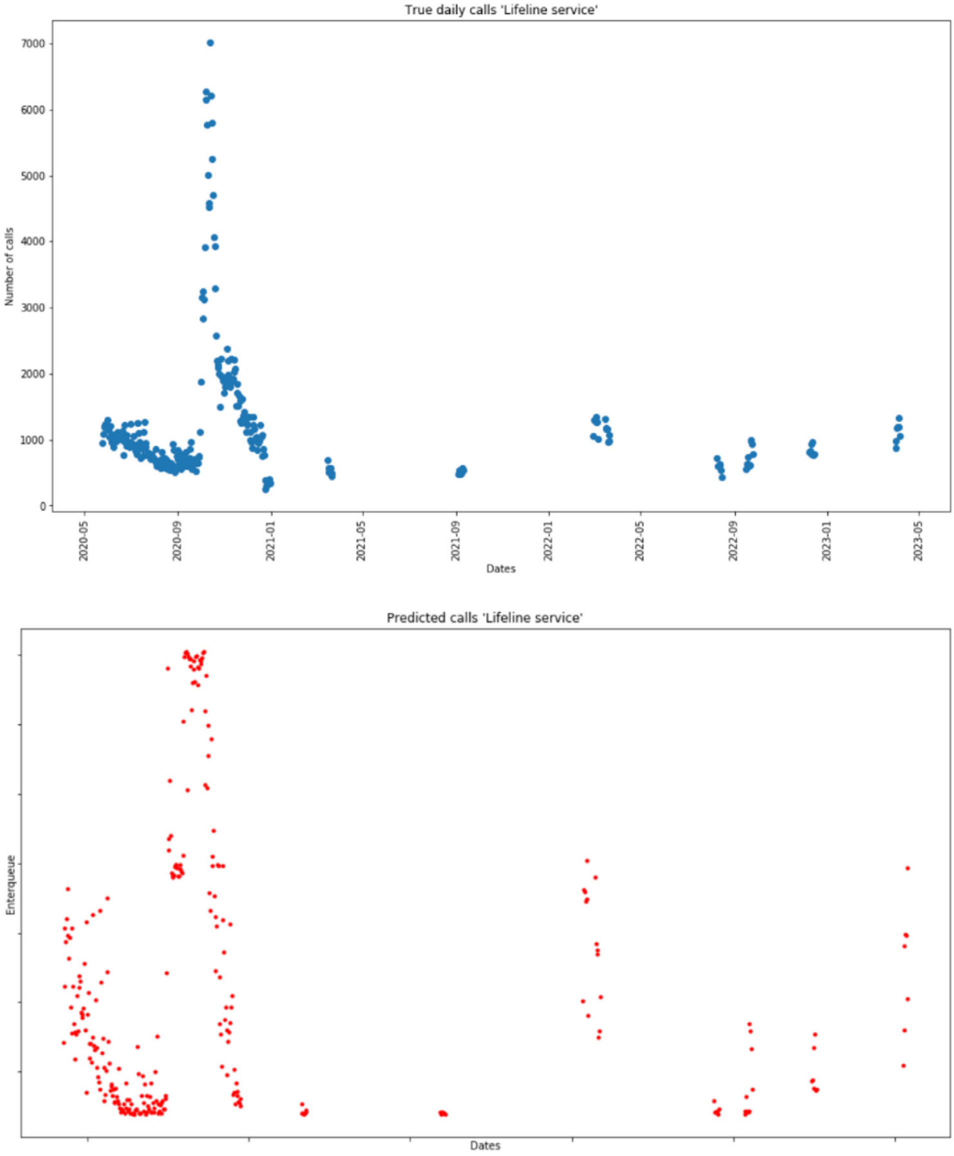

Fig. 5. The figure shows the real data points of the Lifeline phone service along with all retrieved dates in the data set compared to the predicted ones using a Support Vector Machine algorithm.

Limitations

The data sets have some limitations that are needed to be considered:

Not all databases have the same time ranges. The dataset for the lifeline emergency phone call service started a bit later than the others (as seen in Fig. 4), and after 2020, only dates related to earthquakes were retrieved from the lifeline service. Similarly, the questionnaire for the Detection of Risks to Mental Health COVID-19 was established as a countermeasure once the pandemic effects started to be felt on April 13, 2020. There are no records available for the questionnaire for the detection of risks to mental health COVID-19 further than December 2020.

The tool was publicly available for anyone who desired and had access to the internet, so there are unequal records in regard to gender (women 67.14% vs. men 32.86 %).

The tool's target population was divided into health personnel, university students, and the general population, so no deeper classification can be performed in this regard.

Finally, in regard to age, ranges below 13 years old were grouped as one; the same happened with records over 85 years old.

Despite the previous limitations, the datasets reflect the characteristics and demands imposed on Mexico City's first level of mental health support systems during the first year of the pandemic. Statistical manipulations can be performed on the datasets to mitigate the effects of the sample gender distribution.

Ethics Statement

The datasets were retrieved from several agencies under authorization, and the study was conducted in accordance with the Declaration of Helsinki and also held following the Code of Conduct for Scientists established by the Science Council of Japan.

The database from the "questionnaire for the detection of risks to mental health COVID-19" was constructed following the protocol approved by the Ethics Committee from the Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz (Approval Code: CEI/C/010/2020), and the Faculty of Psychology (UNAM), where Informed consent was obtained from all subjects, as mentioned in a previous article https://www.frontiersin.org/articles/10.3389/fpubh.2021.656036/full. The questionnaire and the statement for the protection of personal information can be accessed at https://misalud.unam.mx/.

The database from the Lifeline support system does not retrieve any personal information (maintains anonymity) from the start and according to Mexican regulations established by the National Council from the National System for Transparency (Published in the Official Diary of the Federation (May 4, 2016), normative available at: https://home.inai.org.mx/?page_id=1870&mat=a) they are publicly available for statistical analysis under demand. All other databases are public and retrievable from the corresponding sources mentioned in the article.

Data Availability

Microtrends of mental health reactions for better decision making Generals (Original data) (Tableau Public).

Mental_health_cdmx_2020 (Original data) (Zenodo).

CRediT Author Statement

Carlos Rodrigo Garibay Rubio: Conceptualization, Writing – original draft, Visualization, Data curation, Formal analysis, Software; **Katsuya Yamori:** Supervision, Writing – review & editing, Resources; **Genta Nakano:** Supervision, Writing – review & editing, Resources; **Astrid Renneé Peralta Gutiérrez:** Writing – original draft, Visualization, Data curation, Formal analysis, Software; **Silvia Morales Chainé:** Conceptualization, Methodology, Investigation, Supervision, Writing – review & editing, Resources; **Rebeca Robles García:** Writing – original draft, Conceptualization, Methodology, Investigation, Supervision, Writing – review & editing, Resources; **Edgar Landa-Ramírez:** Writing – review & editing; **Alexis Bojorge Estrada:** Writing – original draft, Conceptualization, Methodology, Investigation, Supervision, Writing – review & editing, Resources; **Alejandro Bosch Maldonado:** Conceptualization, Methodology, Software, Data curation; **Diana Iris Tejadilla Orozco:** Writing – review & editing.

Acknowledgments

We want to thank Dr. Juan Manuel Quijada Gaytan, head of the National Commission for Mental Health and Addictions (Mexico), for helping obtain the necessary data for this research.

We also want to thank Dr. Todd Pataky, Associate professor at the Graduate School of Medicine, Kyoto University, for helping and sharing coding and data management strategies during his lessons.

We also thank the Japan International Cooperation Agency (JICA) for their support.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

he authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] WHO, "World Mental Health report," 2022.
[2] S.M. Chainé, et al., Mental health symptoms, binge drinking, and the experience of abuse during the COVID-19 lockdown in Mexico, Front Public Health 9 (2021) [Online]. Available:https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2021.656036.
[3] R.C. Cuevas, R.E. Cano, and S. Morales-Chainé, "Análisis de datos del instrumento de tamizaje (Facultad de Psicología, UNAM)," Acciones en el marco de la respuesta frente al sismo del 19 de Septiembre de 2017. Lecciones aprendidas y buenas prácticas (énfasis en salud mental), pp. 99–103, 2018.
[4] D.P. Goldberg, et al., Screening for anxiety, depression, and anxious depression in primary care: a field study for ICD-11 PHC, J Affect Disord. 213 (2017) 199–206.
[5] S. Velasco, M. Ruiz, C. Álvarez-Dardet, Modelos de atención a los síntomas somáticos sin causa orgánica: de los trastornos fisiopatológicos al malestar de las mujeres, Rev. Esp Salud Publ. 80 (2006) 317–333.
[6] J. Arrieta, et al., Validity and utility of the patient health questionnaire (PHQ)-2 and PHQ-9 for screening and diagnosis of depression in rural Chiapas, Mexico: a cross-sectional study, J. Clin. Psychol. 73 (9) (2017) 1076–1090.
[7] B. Arroll, et al., Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population, Ann. Fam. Med. 8 (4) (2010) 348–353.
[8] A.J. Mitchell, M. Yadegarfar, J. Gill, B. Stubbs, Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies, BJPsych Open 2 (2) (2016) 127–138.
[9] S.M. Chainé, et al., Mental health conditions during the COVID-19 pandemic, Rev. Int. Investig. Adicc. 6 (2) (2020) 11–24, doi:10.28931/riiad.2020.2.03.
[10] C.R. Garibay Rubio, Mental_health_cdmx_2020, GitHub Reservoir, Zenodo, 2024 vol. V1, doi:10.5281/zenodo.11501598.
[11] C.A. Blevins, F.W. Weathers, M.T. Davis, T.K. Witte, J.L. Domino, The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation, J. Trauma Stress 28 (6) (2015) 489–498.
[12] K. Kroenke, R.L. Spitzer, J.B.W. Williams, The patient health questionnaire-2: validity of a two-item depression screener, Med. Care 41 (11) (2003) 1284–1292.