# SafeConnect

## Connecting Help-Seekers and Volunteers with Database Management

Annan Fu, Danni Hu, Jiani Guo, Xing Gao

Northeastern University, Vancouver, BC, Canada

## Abstract

The SafeConnect database project aims to create an efficient system for an application that connects help seekers with appropriate volunteers and resources, facilitating timely assistance when help seekers cannot find resources elsewhere, and enhancing the support network for victims of various violence. The database is designed with 19 tables to map our clear relationships between volunteers and help seekers as well as needs, resources, incidents, and interventions. With assumptions based on related research, our team set up hundreds of mock data inputs. We created 2 stored procedures and 7 queries to use data visualization to learn more about the insights that we can potentially gain when we are able to receive actual user input.

## Introduction

The app SafeConnect aims to connect help-seekers who are experiencing all forms of violence (domestic violence, intimate partner violence, workplace violence, bullying, etc.) with volunteers. Based on our volunteer work in emergency cases connecting with women and LGBTQ+ victims suffering from domestic violence, this app hopes to allow individuals to connect with their communities immediately and effectively to support needs that cannot be fulfilled solely through governmental or institutional aid. The resources provided by the volunteers include but are not limited to emergency shelter, counseling, accompanying, employment, legal, and translation assistance.

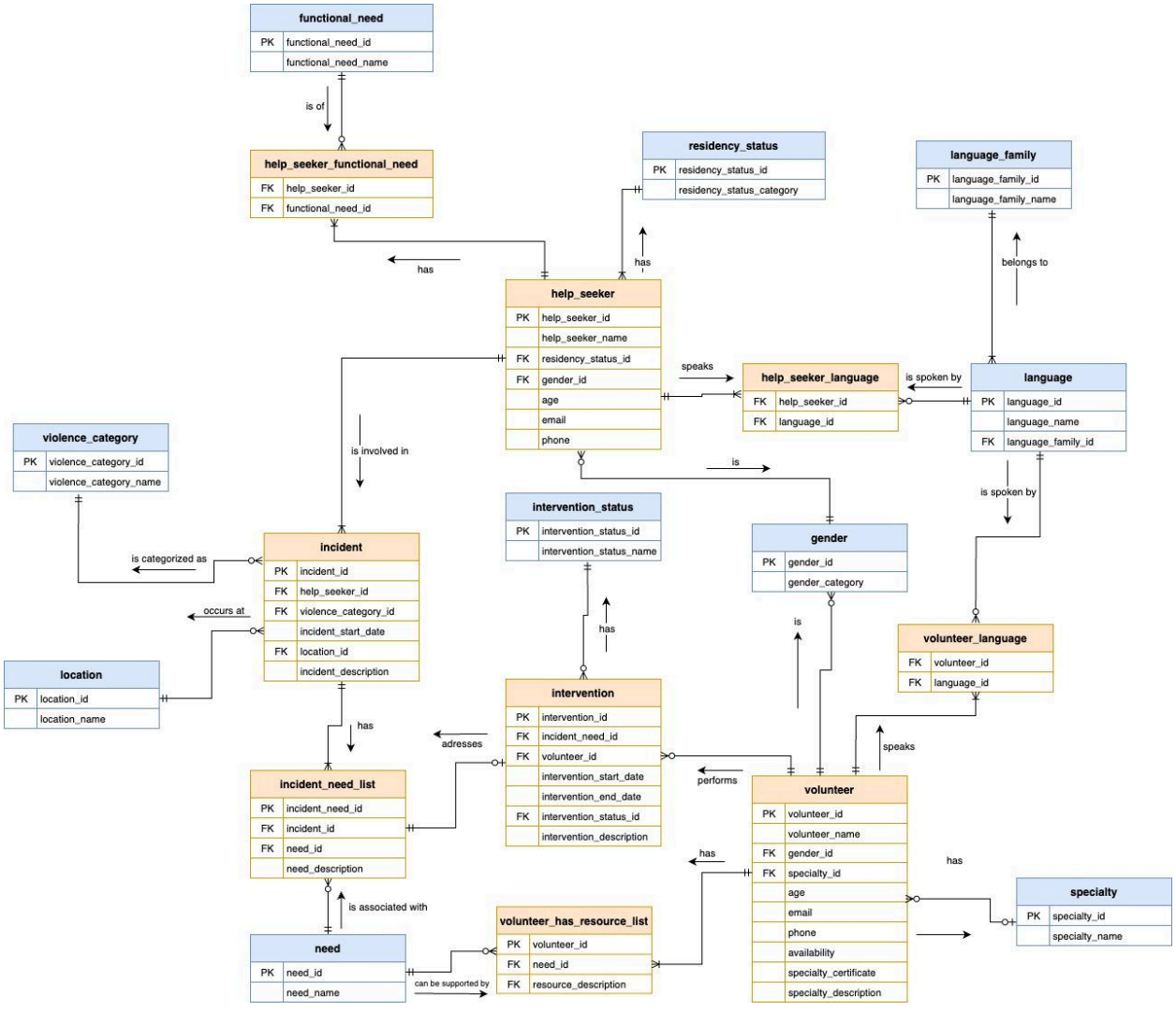
Our current goal is to design a database that can be used to manage user information for help seekers and volunteers, store help seeker's incidents and needs, store intervention conducted by volunteers that involve important information, for example, the intervention status, start and end date. The join tables (incident_need_list, volunteer_has_resource_list, help_seeker_functional_need) are critical in realizing our potential app feature of automatically connecting volunteers with specific resources to those with corresponding needs.

While we have not made the mobile application or have actual user data input, we researched studies on different categories of violence and abuse, the demographics who experience these

violent incidents, intervention effectiveness, and help seekers with disabilities to create a database with assumptions from the current research. We also designed 2 stored procedures (to filter out invalid interventions and match volunteers with incident need) and ran queries on these data to present some insightful visualization with the current assumptions which could be updated once we have real user input.

## Database Design

Below is the conceptual model of the 19 tables, their columns, and the relationships between them. The cardinality and participation constraints are specified in the ER Diagram.




- **Basic entities (in blue color):**

There are 10 tables that store the basic information about the model. Each table consists of two columns: a primary key id storing the unique auto-incrementing sequence, and the key information column that stores the category name of that information, such as violence category, gender, location, need category, etc. (Exception: Language table is associated with Language Family table)

- **Key user case entities (in orange color):**
  There are 9 join tables that store user case information. Each of these join tables is associated with two or more basic entities to reflect the one-to-many or many-to-many relationships. These entities capture essential user case aspects of the app functionality, such as information about help seekers, incidents of violence, needs associated with the incidents, volunteers and resources, and their interactions reflected in the interventions.

- **A user case walkthrough:**

  Help seeker registers personal information(basic information and language) -> reports an incident -> selects functional needs and general needs that are relevant to the incident

  Volunteer registers personal information(basic information and language) -> selects the resources that could be provided to match the need categories

  The application checks the resources and availability of the volunteers -> activates an intervention corresponding to an incident-need pair -> keeps track of the progress of each intervention and records the effectiveness of the result with an end date.

## Data Sources and Methods

**Data Source and Assumption**

We have 19 entities in total in our database design: 10 of them store basic settings and information that would be used in the user cases, and the other 9 entities including the join tables are mock user cases generated by AI based on our assumptions. In this application prototype, we assume the user scope of users is focused on Vancouver, Canada.

- **Basic entities (10 tables)**

  We define the basic category of information data based on objective facts, social norms, related publications, and public statistics. Considering the scalability limitation of this application prototype, we may select representative values for application presentation.

1. **violence**[1]: Defines the types of violence.

   Based on the "Types of Violence and Abuse" introduced on the Canadian Government website, we categorize violence into four types as the data value: Domestic Violence, Intimate Partner Violence, Bullying, and Workplace Violence.

2. **location**[2]: The location where the incident that the help seeker is involved in occurs.

   Based on the official data from the Destination Vancouver website, we list all the cities and regions within Greater Vancouver, including 16 locations as the data value: Bowen Island, Burnaby, Coquitlam, Delta, Langley, Maple Ridge, New Westminster, North Vancouver, Pitt Meadows, Port Coquitlam, Port Moody, Richmond, Surrey, Vancouver, West Vancouver, White Rock.

3. **need**[3]: Lists categories of needs that help seekers might have.

   According to Amanda L. Vasquez and Jaclyn Houston-Kolnik's report, we categorize the common resources sought by help seekers into 9 types as the data value: shelter, food, medical care, legal assistance, psychological support, financial assistance, counseling, housing, and employment.

4. **specialty**[4]: Categorizes the specialties of the volunteers to intervene in special cases.

   By referencing the resources provided by the New York Anti-Violence Project, we define the specialties of volunteers in 5 types as the data value: crisis intervention, legal, social_work, counseling, translation/interpretation.

5. **gender**[5]: Stores gender categories for inclusivity.

   Based on the inclusive language guidelines provided by California State University San Marcos (CSUSM), we include a diverse range of 10 gender identities as the data value: man, woman, non-binary, genderqueer, agender, bigender, two-spirit, transgender, other, prefer not to respond.

6. **language**[6]: Lists languages spoken by help seekers and volunteers.

   Based on the statistics provided by the Canada BC provincial government, we define 18 languages as the data value: Mandarin Chinese, Cantonese Chinese, English, French, Punjabi, Hindi, Urdu, Spanish, Russian, Tagalog, Tamil, Arabic, Cree, Ojibwe, Salish, Dene, Korean, Farsi.

7. **language_family**[7]: Categorizes languages into families.

   We put the above 18 languages into 10 categories: Sino-Tibetan, Indo-European, Afro-Asiatic, Austronesian, Dravidian, Algonquian, Salishan, Athabaskan, Korean, and Iranian.

8. **residency_status**[8]: Defines residency statuses of the help seekers.

   Based on the definitions provided by the Canadian Government website, we categorize residency status(immigrant status) into 3 types as the data value: Non-immigrant citizen, Immigrant citizen or permanent resident, and Other temporary resident permit.

9. **functional_need**[9]: Describes different functional needs of the functional abilities of help seekers.

   Based on the Disability Language Guide by Stanford Disability Initiative Board, common functional requirements, we categorize functional needs from help seekers mainly into 6 types: physical aid, hearing aid, vision aid, medical aid, communication aid, and mobility aid.

10. **intervention_status**: Reflects the current status of the intervention performed.

    Based on the application design on intervention progress control, we define intervention status in 5 situations: Pending, In Progress, Effective, Escalated, and Closed.

- **User case entities (9 tables)**

  The remaining 9 entities are mock user case data generated using generative AI(ChatGPT) with our specific prompts for each entity based on the basic entity attribute value. The data is not derived from real-world incidents or actual individuals but was created artificially to simulate realistic scenarios for analysis and visualization.

  **Mock Data-generation Steps:**

  (1) Set up the database based on the above information with “create table” of the 19 entities and “insert value” of the above 10 basic entities, please refer to “group07_backup.sql”. With the table cardinality, participation constraints, primary key, and foreign keys predefined and basic information provided, provide the backup(setup) file for AI to analyze the basic information and relationship.

(2) For each entity out of the 9, define the assumptions as prompts provided for AI in terms of sample size, value output example, and constraints. For Value format, ask AI to generate the row value except for the primary key id which is defined as auto-increment. For constraints, quantify the approximate distribution and percentage of the value in certain attributes based on our research on the related source and statistics, which are significant to reflect the current situation to render further data analysis.
(3) Due to the limitation of AI in processing complex database models and relationships, make manual adjustments to clean the generated mock data in order to better facilitate the presentation of further analysis.
(4) Insert the value of the final version into the database.
(5) Use a stored procedure to further filter out the invalid data based on the function logic of the application.

**Assumptions:**

11. **help_seeker:** Stores information about individuals seeking help.

**Sample size:**

We include a total of 100 sample records for help seekers to ensure an adequate sample size for statistically significant data analysis.

**Value format:**

SQL statement: INSERT INTO help_seeker (help_seeker_name, residency_status_id, gender_id, age, email, phone) VALUES;

Example value: ('Alice', 1, 2, 30, 'alice@example.com', '123-456-7890').

**Constraints:**[10][11][12][13][14][15]
Based on the table constraint design(foreign keys) itself, the data value should be random but must adhere to the following requirements:
According to related statistics reports by the Canadian government on family violence, intimate partner violence, and workplace violence:
1. Gender: women and other genders are reported to experience much higher rates compared to men. We designed our sample data with a gender ratio of approximately 4 (women): 2 (men): 2 (other genders) to reflect these findings.
2. Residency status: Non-immigrant residents report higher rates of violence in bullying and workplace violence, while immigrants are more prone to domestic violence and intimate partner violence. Due to the lack of reported cases or analysis specific to temporary residents, the proportion of non-immigrants and immigrants should be much higher compared to temporary residents.

3. Age: The age of victims is predominantly between 12 and 44 years. Nearly half of women aged 60 and older have experienced some form of abuse since turning 55. We include certain older female help seekers in our samples.

**12. incident:** Records each incident of violence.

**Sample size:**

In this application prototype, we assume one incident corresponds to one help seeker, the data sample size for the incident table is also 100.

**Value format:**

SQL statement: INSERT INTO incident (help_seeker_id, violence_category_id, location_id, incident_date, incident_description) VALUES;

Example value: (1, 1, 14, '2023-01-01', 'Incident of domestic violence reported by Alice in Vancouver.').

**Constraints:**

Based on the table constraint design(foreign keys) itself, the data value should be random but must adhere to the following requirements:
1. The incidents are associative with the correlation between help seekers and the violence, which should take into account the distribution of gender and residency status of the help seekers: [10][11][12][13][14][15]
(1)Domestic violence and intimate partner violence (more gender-based): immigrant > nonimmigrant;
(2)Bullying(more youth): immigrant PR/citizen : nonimmigrant = 1 : 2;
(3)Workplace: immigrant: nonimmigrant = 2 : 3;

2. Locations: [16]
We assume the probability of incident occurrence is the same across all the locations, then the incident distribution should be proportional to the population of the locations:
tier 1: Vancouver - 25%; Surrey - 20%; Richmond - 10%; Burnaby - 10%;
tier 2: North Vancouver, Coquitlam, Langley, Delta, Maple Ridge (around 7-8%);
tier 3: Other cities, immaterial, could be 1 - 2 samples or just leave.

3. Incident date: Starts from 2023-01-01.

**13. incident_need_list:** Lists the needs associated with each incident.

**Sample size:**

By referencing a Victim Need Report by an Illinois Authority, victim service providers identified the needs of violent victims, which included basic needs(shelter, food, life skills) and presenting needs(counseling, mental health, legal service, etc). Since each incident corresponds to one or more needs of the help seeker, the data sample size for incident_need_list is expanded to 200 to accommodate the varying needs of the help seekers.

**Value format:**

SQL statement: INSERT INTO incident_need_list (incident_id, need_id, need_description) VALUES;

Example value: (1, 1, 'Shelter needed for Alice after domestic violence.'),

**Constraints:**[3][17][18][19]
Based on the table constraint design(foreign keys) itself, the data value should be random but must adhere to the following requirements:

1. For domestic violence (violence 1) and intimate partner violence (violence 2), each incident should have 1-3 needs, averaging two needs:

(1)Basic needs: Shelter, food: >1/3 of help seekers; housing, employment: >1/3 of help seekers.

(2)Counseling: >50% of help seekers; psychological: >40% of help seekers; medical, legal: >15% of help seekers. Other needs should be random and not material.

2. For bullying (violence 3), the top need is psychological service, followed by medical. Other needs should be random, with each incident having 1-2 needs, mostly 1 need.

3. For workplace violence (violence 4), the top needs are psychological and employment, followed by counseling. Each incident should have 1-2 needs.

**14. help_seeker_functional_need:** Captures the functional needs of help seekers.

**Sample size:**

Since not every help seeker might have functional needs, we set the sample size to 50.

**Value format:**

SQL statement: INSERT INTO help_seeker_functional_need
(help_seeker_id, functional_need_id) VALUES;
Example value: (1, 1)

**Constraints:**

Based on the table constraint design(foreign keys) itself, the data value should be randomly generated.

**15. volunteer:** Stores information about volunteers.

**Sample size:**

Considering that the number of volunteers is often relatively smaller in real-world scenarios compared to help seekers, we set the data sample size to 51, around half of that of the help seekers.

**Value format:**

SQL statement: INSERT INTO volunteer (volunteer_name, gender_id, specialty_id, age, email, phone, availability) VALUES;

Example value: ('John Doe', 1, 1, 45, 'john.doe@example.com', '555-1234', 1)

**Constraints:**

Based on the table constraint design(foreign keys) itself, the data value should be random but must adhere to the following requirements:
Set all the availability to 1(the volunteers are available for all resources they have).

**16. volunteer_has_resources_list:** Lists the resources with each volunteer to address needs.

**Sample size:**

Given that each volunteer may be capable of providing multiple types of resources, we expanded the data sample size to 101 entries(around double the size of volunteer). This ensures that the fifty volunteers in our dataset have detailed descriptions of the resources they can offer, thereby enhancing the system's ability to match help seekers with suitable volunteers.

**Value format:**

SQL statement: INSERT INTO volunteer_has_resource_list (volunteer_id, need_id, resource_description) VALUES;

Example value: (1, 1, 'Temporary shelter arrangements available.')

**Constraints:**
Based on the table constraint design(foreign keys) itself, the data value should be randomly generated.

17. **Help_seeker_language,** Records the languages spoken by help seekers.
18. **volunteer_language:** Records the languages spoken by volunteers.

**Sample size:**

Random, based on constraints.

**Value format:**

(1)For help seekers:

SQL statement: INSERT INTO help_seeker_language (help_seeker_id, language_id) VALUES;

Example value: (1, 3), (1, 1).

(2)For volunteers:

SQL statement: INSERT INTO volunteer_language (volunteer_id, language_id) VALUES;

Example value: (1, 1), (1, 3).

**Constraints:**
Based on the table constraint design(foreign keys) itself, the data value should be random but must adhere to the following requirements:

Language coverage:
(1) For help seekers, we aimed to represent the diversity of languages spoken in Vancouver. Each help seeker is associated with one or more languages they speak, ensuring a realistic representation of the linguistic diversity in the population.
(2) For volunteers, each language in our list should be covered by at least one volunteer. This is crucial for providing effective support to help seekers in their preferred languages.

The frequency of volunteers speaking common languages like English, Mandarin, Cantonese, and French is higher, while those speaking less common languages are fewer but still represented.

**19. intervention:** Records interventions made for each incident.

**Sample Size:**
Assume current incidents could be processed 90% due to waiting time and resource allocation. We need around 180 records for the first 180 incident_need_id (200 incidents * 90%).

**Value format:**
SQL statement: INSERT INTO intervention (incident_need_id, volunteer_id, intervention_start_date, intervention_end_date, intervention_description, intervention_status_id) VALUES
Example value: (1, 1, '2023-01-26', NULL, NULL, 2).

**Constraints:**[20][21][22][23][24]
Based on the table constraint design(foreign keys) itself, the data value should be random but must adhere to the following requirements:
1. The intervention start date should be after the incident date and within 1 month. The end date and intervention status depend on the following factors:
(1) Not all end dates should have values; some should be null, especially for rural areas where effectiveness is lower.
(2) Housing needs should last at least 6 months, while shelter and food should take less than 3 months. The overall duration should range from 10 days to 18 months, with most over 6 months (180 days after the start date).

2. For the intervention_status, randomly allocate the 5 status options, with most being effective (3) or closed (5). Only interventions with end dates should be effective or closed. Those with a null end date should have one of the other 3 status options. If the end date is NULL, the intervention status should not be 3 or 5.

3. For help seekers with functional needs, the effectiveness within 12 months should be 40% less than for those without functional needs.

## User Cases – Filtering And Matching

**Stored Procedure**
**to Delete Ineligible Interventions:**

In this application prototype, we use one stored procedure to validate whether the volunteer has the resources to match the incident needs before starting the intervention. We effectively filter out 8 invalid intervention records by calling this stored procedure after mock data insertion. Therefore, the valid sample size of intervention that is utilized in the following data analysis should be 180 - 8 = 172.

**Stored Procedure**
**to Create Intervention and Match Volunteers with Help Seekers:**

** Note:
*This procedure is implemented after mocking data, writing other queries and writing the report. Originally we have implemented only the procedure above to delete invalid intervention mock data entries afterwards, where volunteers' resources are mismatched with needs.*

*However, we found that a prior control makes more sense in the functional logic and is crucial in actual user cases, so we created this alternative stored procedure for reference, which is more comprehensive for our main app feature.*

*For the queries, we are still using our mocked data with already inserted interventions.*

As for using this to create interventions in the actual application, this procedure is intended to be run after the operations of storing volunteer and help_seeker information, creating incident, need, and incident_need_list.

Since there might be several needs an incident involves, the JOIN table incident_need_list would include several matchings of incident and need. Thus, multiple interventions need to be created to address these different pairings of incident_need.

We applied the CURSOR operations learned in class to write this procedure and used a while loop to find one volunteer to match each incident need pairing in order to create a new unique intervention for this need.

*DECLARE CURSOR*
*DECLARE CONTINUE HANDLER*
*FETCH one row to incident_need and need variables*
*WHILE LOOP:*
1. *Find one volunteer to fulfill the need*

*2. Create an intervention*
*FETCH next row*
*END WHILE*

There will be no updating of the volunteer's availability after being chosen for a particular intervention, since one volunteer can intervene with multiple cases in reality depending how much they want to commit. Volunteer's availability should be updated by themselves on the app.

However, in the future, we could implement TRIGGER operations to better the logic of matching volunteers for each incident need.

## User Cases – Data Analysis for Volunteering Work

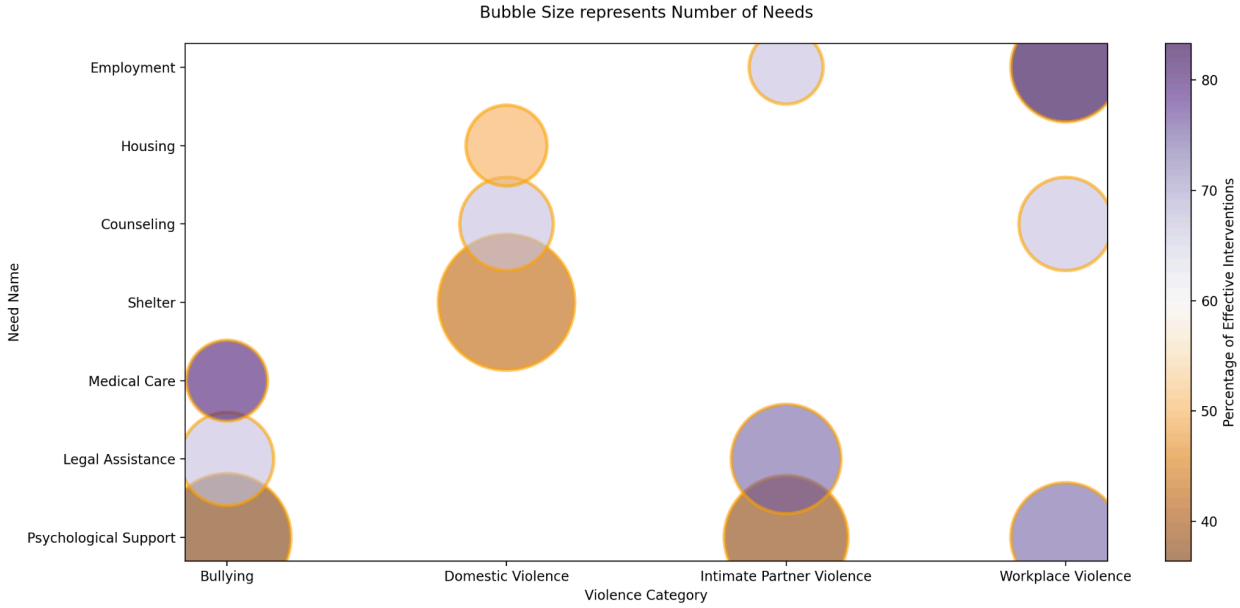
This app is used by both help seekers and volunteers. In this application prototype, we designed 7 queries (5.1 and 5.2 are two queries) from the volunteer user's perspective, aiming to facilitate decision making for volunteers to better allocate resources and satisfy help seeker's needs through data analysis based on the information stored in the database.

**Question 1: On Violence Category-Need Satisfaction**

(Explanation of the schema) As each **intervention** is directly addressing an incident's various needs, it means there will be multiple interventions created to meet the same incident's different needs.

(Question) For each **violence category**'s various needs (category of need), how many needs have been requested? What is the percentage of **intervention** being effective out of all interventions (excluding the ones that are closed and escalated) for each pairing of violence category and need? List the top 3 largest number of needs requested for each violence category.

| violence_category | need_name | num_of_need | per_effective |
|---|---|---|---|
| Bullying | Psychological Support | 15 | 36.4 |
| Bullying | Legal Assistance | 8 | 66.7 |
| Bullying | Medical Care | 6 | 80.0 |
| Domestic Violence | Shelter | 17 | 42.9 |
| Domestic Violence | Counseling | 8 | 66.7 |
| Domestic Violence | Housing | 6 | 50.0 |
| Intimate Partner Violence | Psychological Support | 14 | 37.5 |
| Intimate Partner Violence | Legal Assistance | 11 | 75.0 |
| Intimate Partner Violence | Employment | 5 | 66.7 |
| Workplace Violence | Psychological Support | 11 | 75.0 |
| Workplace Violence | Employment | 11 | 83.3 |
| Workplace Violence | Counseling | 8 | 66.7 |




*Analysis and Insights:*

*By identifying the pairings with the highest number of needs requested as well as if it has a low effectiveness percentage, volunteers and organizations can prioritize resources and attention towards the most critical areas according to each different violence category. This can help ensure that the most prevalent needs are addressed promptly and adequately.*

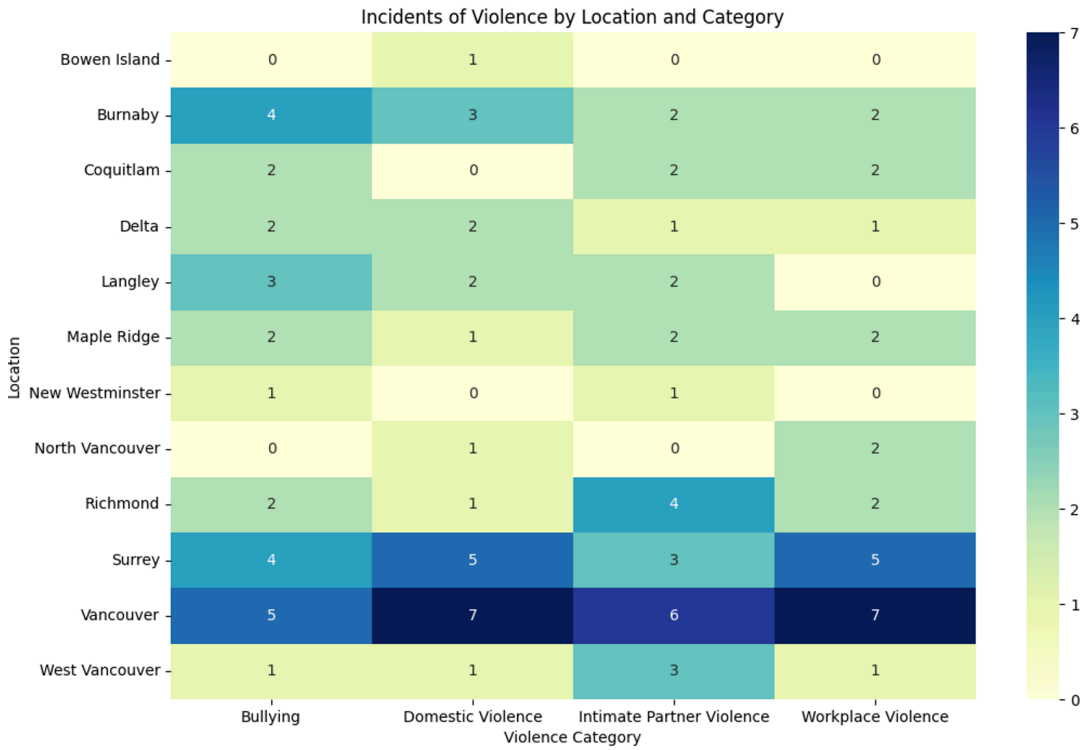
*For example, with our current data input based on research, we can see from the bubble chart that the biggest bubbles (most need) also have darker orange (interventions not effective) are psychological support needed for bullying & intimate partner violence, shelter needs for*

*domestic violence. We found that the National Network to End Domestic Violence from Washington DC[20] has raised the policy recommendation to "(advance) access to safe, affordable housing for survivors through enhanced legal protections".*

**Question 2: On Violence-Location Statistics**

(Question) For each **location,** what **type of violence** happens and how many **incidents**? For each location and violence pairing, list the number of incidents in descending order. Show the top 10 most number of incident pairings for tabular view and all pairings for heatmap visualization.

| location | violence_category | num_incident |
|---|---|---|
| Vancouver | Domestic Violence | 7 |
| Vancouver | Workplace Violence | 7 |
| Vancouver | Intimate Partner Violence | 6 |
| Surrey | Domestic Violence | 5 |
| Surrey | Workplace Violence | 5 |
| Vancouver | Bullying | 5 |
| Surrey | Bullying | 4 |
| Richmond | Intimate Partner Violence | 4 |
| Burnaby | Bullying | 4 |
| Surrey | Intimate Partner Violence | 3 |




*Analysis and Insights:*

*By examining the most frequent location-violence pairings, we can discern patterns, which may indicate underlying social, economic, political, or cultural factors contributing to violence in specific areas.*

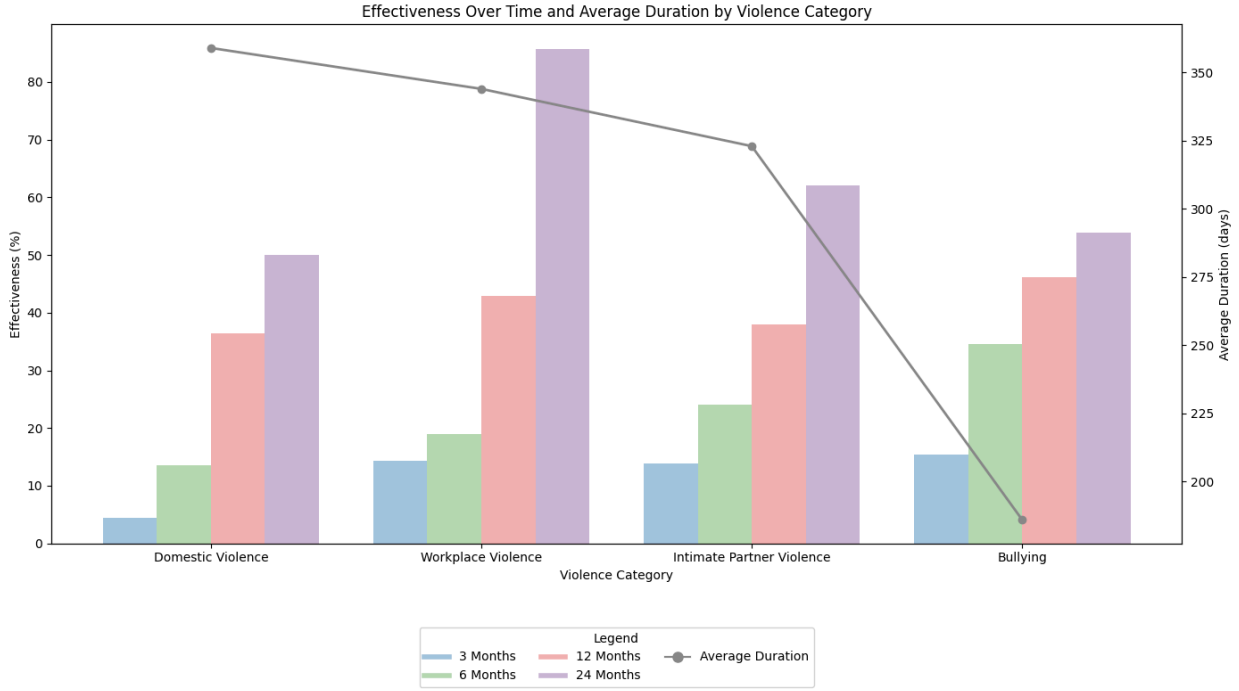
*According to our current limited research, the more people reside in an area, the more crimes there are. We can see from the heatmap that the more saturated blue the colors of a block are, the more violence of a certain category happens. This result can also be affected if the residents of an area are actively seeking help from an incident on this app.*

**Question 3: On Intervention - Effectiveness Over Time**

(Explanation of the schema) Intervention status: 'Pending', 'In Progress', 'Effective, 'Escalated' (escalated to a higher authority), 'Closed' (marked as closed for other reasons)

(Question) For each **violence category,** show the percentage of **interventions** being effective out of all interventions (excluding cases where intervention status = "closed" and = "escalated") after 3 months, 6 months, 12 months and 24 months. In the same query, output the average duration each **violence category**'s **interventions** take to be effective (even if it took more than 24 months), ordered by the longest average duration.

| violence_category | effective_3_months | effective_6_months | effective_12_months | effective_24_months | average_duration |
|---|---|---|---|---|---|
| Domestic Violence | 4.3 | 13.0 | 34.8 | 47.8 | 359 |
| Workplace Violence | 10.0 | 15.0 | 40.0 | 80.0 | 352 |
| Intimate Partner Violence | 13.8 | 24.1 | 37.9 | 62.1 | 323 |
| Bullying | 15.4 | 34.6 | 53.8 | 61.5 | 193 |




*Analysis and Insights:*

*With this analysis we can determine, for each violence category, the effectiveness of the interventions are the duration it requires for interventions to be successful. This insight allows volunteers and organizations to better prepare for future interventions according to each category of violence, in terms of allocating resources, planning strategies, estimating time commitments, and specialty training.*

*As we can see from the line and bar chart combined visualization, domestic violence's incident intervention (line chart) takes the longest time and is relatively less effective (bar chart) which is easy to guess, but workplace violence also has quite a high line for the duration, and are almost equally not effective before 24 months. In our research, we found that Rapp-Paglicci et al's Handbook of Violence[22] stated that if the work is long-term (more than six months), the survivor would suffer from various kinds of trauma that require psychological healing.*

**Question 4: On Language Family - Need Satisfaction**

(Question) For each different language family, which needs (such as shelter, food, medical care, etc.) are the least satisfied? Calculate the satisfaction rate (i.e., the proportion of effective interventions, *with no time limit) for each need within each language family, order from least percentage of effectiveness to most. Show the satisfaction rate that is lower than 60% for tabular view and all pairings for sunburst visualization.

| language_family | need_name | per_effective |
|---|---|---|
| Afro-Asiatic | Housing | 0.0 |
| Algonquian | Shelter | 0.0 |
| Austronesian | Psychological Support | 0.0 |
| Austronesian | Legal Assistance | 0.0 |
| Sino-Tibetan | Shelter | 0.0 |
| Sino-Tibetan | Medical Care | 0.0 |
| Indo-European | Psychological Support | 43.2 |
| Afro-Asiatic | Counseling | 50.0 |
| Afro-Asiatic | Psychological Support | 50.0 |


*Analysis and Insights:*

*This interactive sunburst chart we made plotly can not only show us how effective each language family's corresponding needs are, but also show us how each language family's overall need satisfaction rate.*

*Understanding the satisfaction of needs within different language families can help optimize resource allocation to better meet the needs of specific language groups. It can help in organizing targeted volunteer activities or submitting petitions to the government.*

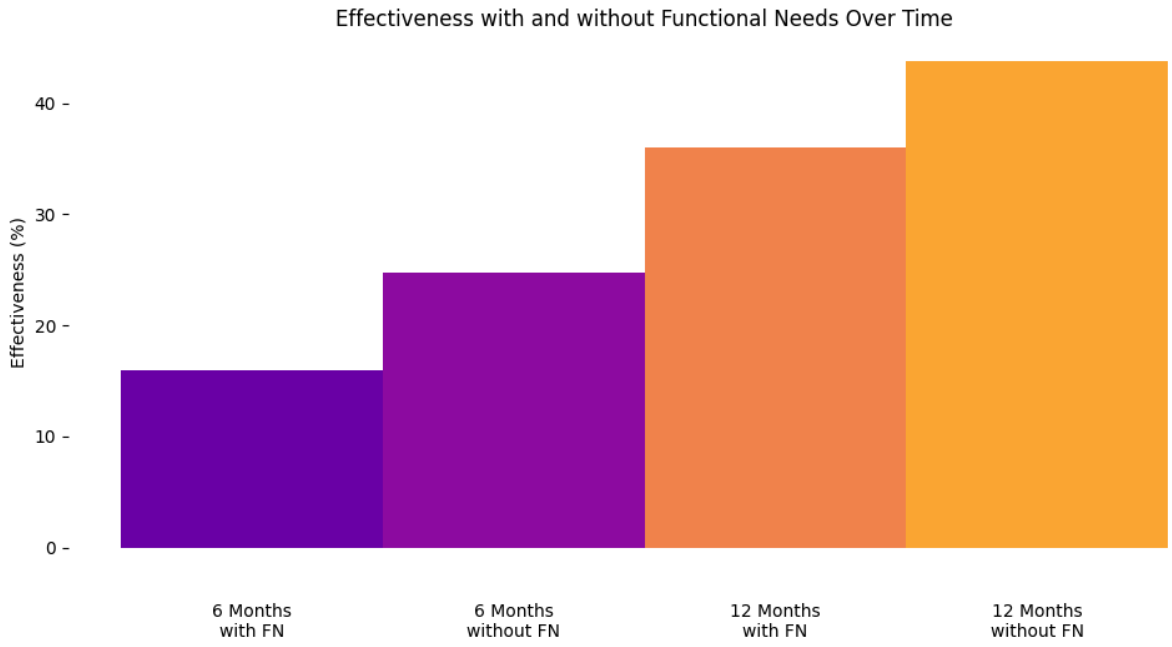
**Question 5.1: On Intervention Effectiveness if with Functional Needs**

(Explanation of the schema) Since functional_need's relationship with incident_need still need other implementations, this query can be currently used as a hypothetical analysis. We currently think of functional need for now as a need that each help_seeker might have but not necessarily

addressed by the volunteers (the volunteers do not have such a resource) to exemplify how these needs in our current environment are unseen.

(Question) Is there any connection between **help seekers** having **functional needs** or not and whether the **interventions** are successfully resolved or not within a long enough period of time (6 months and 12 months)?

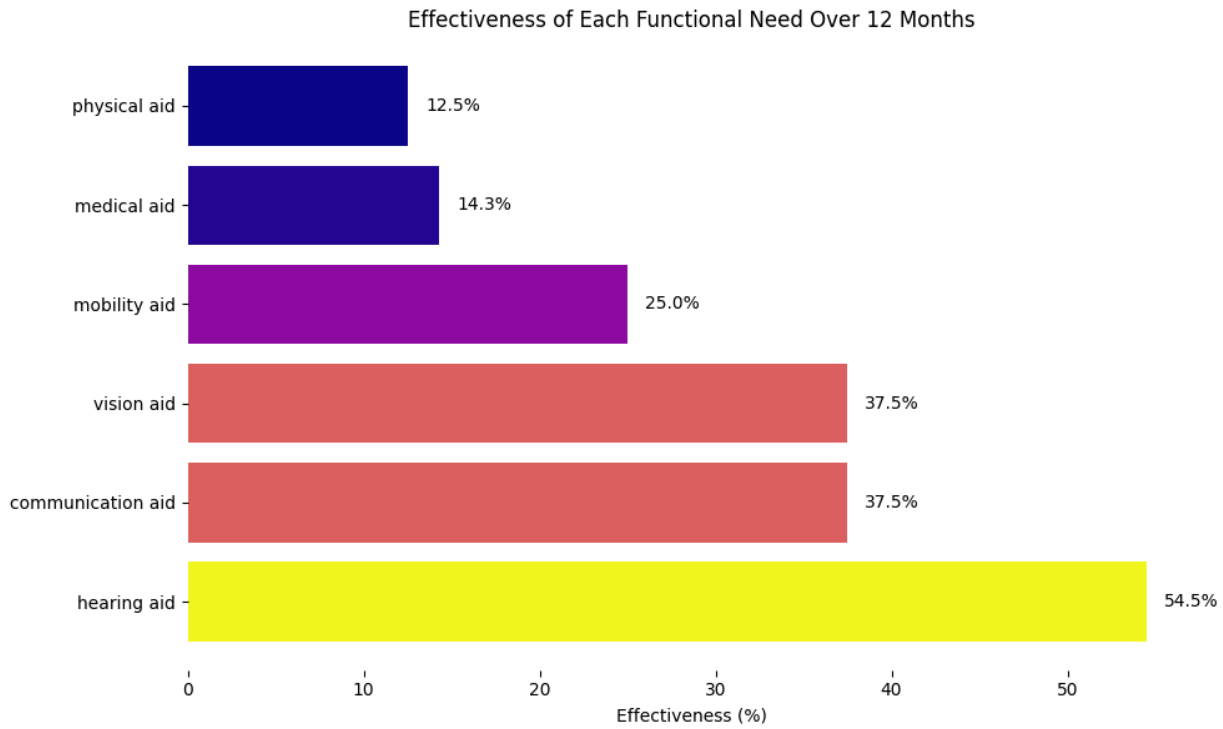
| 6_months_with_fn | 6_months_without_fn | 12_months_with_fn | 12_months_without_fn |
|---|---|---|---|
| 16.0 | 24.7 | 36.0 | 43.8 |




**Question 5.2: On Intervention Effectiveness with Each Functional Need**

(Question) Are there any correlations between what the help seeker's functional needs are and whether the interventions are successfully resolved or not within a long enough period of time (12 months)? For each kind of functional need of help seekers, what are the percentage of the interventions being effective, order from least percentage of effectiveness to most?

| functional_need_name | effective_12_months |
|---|---|
| physical aid | 12.5 |
| medical aid | 14.3 |
| mobility aid | 25.0 |
| communication aid | 37.5 |
| vision aid | 37.5 |
| hearing aid | 54.5 |




*Analysis and Insights:*

*Is it possible that since certain functional needs are less well addressed in our public environment and that both the volunteers and the help_seeker find it more difficult to resolve the incident? This will help volunteers to do organized work on addressing certain accessibility problems and even make campaigns to voice concerns in the lack of resources publicly for these needs.*

*As stated by Lightfoot and Williams in The Intersection of Disability, Diversity, and Domestic Violence: Results of National Focus Groups[23] there are six primary themes emerged relating to helping seekers with disabilities, which include "lack of accessible services, different manifestations of abuse, isolation, negative attitudes toward people with disabilities". Although*

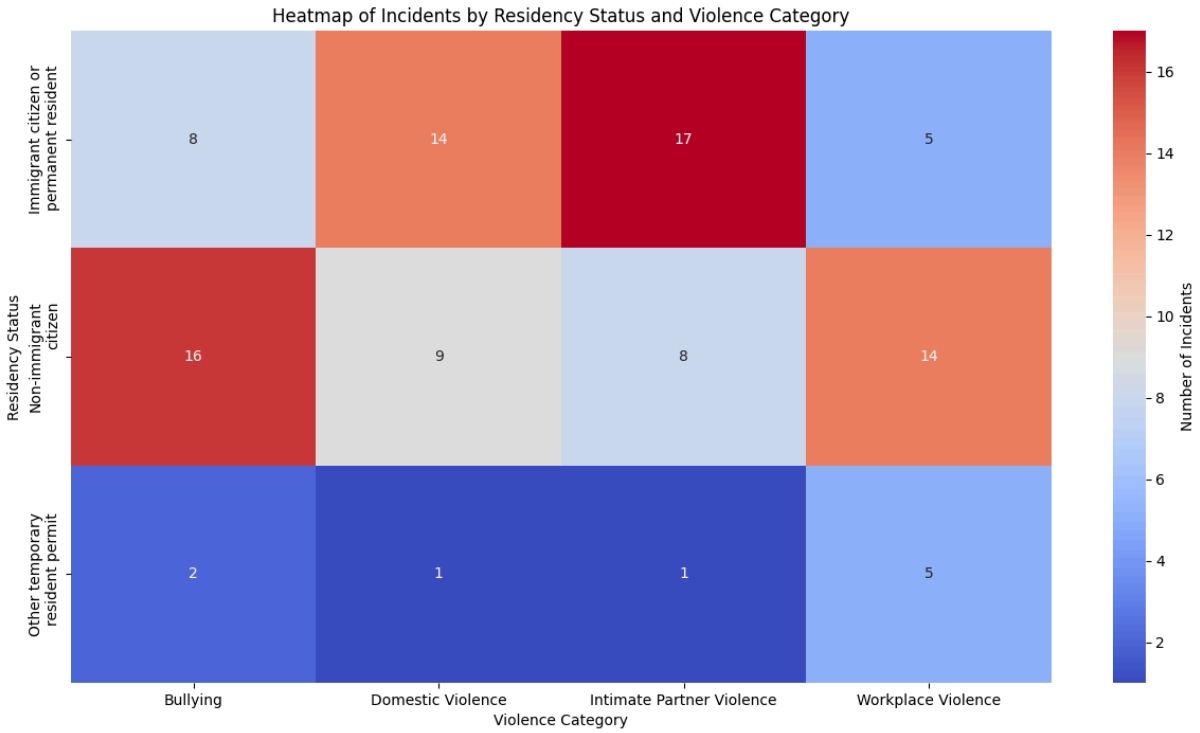
*we have no actual user data inserted, based on previous experience of volunteering, people with different functional needs, no matter if it is for communication aid or mobility aid, meet a lot of barriers in terms of seeking intervention for an incident.*

*Another study by Hague et al.[24] on women with disability found that "support services frequently overlook or respond inappropriately" to women with disability in domestic violence situations in the UK.*

**Question 6: On Help Seeker Statistics - Residency Status**

(Question) For help seekers with different **residency status**, what **kind of violence** do they face the most? Show the number of **incidents** that happen for different pairs of residency status and violence categories. Which pair experiences the most violent incidents? Show the top 10 most number of incident pairings in tabular view and all pairings in heatmap visualization.

| residency_status_category | violence_category_name | num_incident |
|---|---|---|
| Immigrant citizen or permanent resident | Intimate Partner Violence | 17 |
| Non-immigrant citizen | Bullying | 16 |
| Immigrant citizen or permanent resident | Domestic Violence | 14 |
| Non-immigrant citizen | Workplace Violence | 14 |
| Non-immigrant citizen | Domestic Violence | 9 |
| Immigrant citizen or permanent resident | Bullying | 8 |
| Non-immigrant citizen | Intimate Partner Violence | 8 |
| Immigrant citizen or permanent resident | Workplace Violence | 5 |
| Other temporary resident permit | Workplace Violence | 5 |
| Other temporary resident permit | Bullying | 2 |




*Analysis and Insights:*

*The findings can inform volunteers to conduct more community outreach and education programs that benefit specific groups of people who are more vulnerable to certain types of violence.*

*With our data input based on research, the heatmap shows with the most saturated red that immigrants or permanent residents experience the most cases of intimate partner violence. For example, based on Maher and Segrave's research[15] on migration status and family violence, women who are immigrants face extra barriers in intimate partner violence because of the immigration policy.*

## Conclusions

In conclusion, the SafeConnect database project has developed a scalable management system that bridges the gap between help seekers and volunteers. The project involved designing and implementing a database with 19 entities, each capturing essential aspects of help-seekers, volunteers, incidents, and the intervention that reflected the interactions between them. This design process included defining the structure for various entities, covering a comprehensive

dimension of help seekers, their needs, incident details, and volunteer resources, ensuring a holistic representation of the support ecosystem.

We incorporated an AI tool to generate key user case entity mock data based on the predefined basic entity information. To ensure AI generates data as realistically as possible, we quantified assumptions and prompts informed by reputable sources. These assumptions guided the creation of mock data for help seekers, incidents, needs, volunteers, and interventions.

The data analysis mainly focused on addressing specific, insightful questions, such as identifying the most common needs for each type of violence, analyzing the distribution of violence incidents across different locations, and evaluating the effectiveness of interventions. By developing detailed SQL queries, we extracted insights that highlight resource allocation, satisfaction levels, and intervention outcomes.

The project demonstrated the importance of a well-structured database in enhancing community safety and support networks. Through comprehensive documentation, we outlined the database design, data sources, assumptions, and methods used in data generation, ensuring transparency and reproducibility.

**Key accomplishments**

1. **Database Development:** Planned, designed, and implemented a robust database with 19 entities, holding essential information about help-seekers, volunteers, incidents, and interventions.
2. **Data Simulation:** Used data from reputable sources as benchmarks to generate realistic mock data with AI, simulating real-world scenarios for analysis.
3. **Comprehensive Documentation:** Documented the database design and implementation step by step, including the ER diagram, data sources, assumptions, and methods used in data generation.
4. **Query Development:** Developed SQL queries to address complex questions, providing valuable insights into resource satisfaction, violence statistics, and intervention effectiveness.
5. **Store Procedure Developments:** Developed two stored procedures to use in different stages of the app operations, one to filter out invalid interventions, and another to create matched volunteers with multiple incident needs of a help seeker.

**Limitations**

1. With limited time and a lack of available datasets on violence-related research, we used AI-generated mock data, which required extensive manual cleaning using reputable

related data as benchmarks. This may also cause our insights and analysis to differ from actual research.

2. Ensuring data validation requires constraints on the correlation of user cases. In our prototype, we implemented a sample stored procedure for data input validation. However, a real-world application would need more complex validation mechanisms.

3. The current version of the application prototype offers limited features and a general-level concept. It does not fully address correlations such as matching the languages of volunteers and help-seekers, aligning volunteer specialties with specific needs, considering functional needs affecting general needs, and accounting for volunteer capacity in resource allocation.

**Future improvements**

1. Devote more time for data collection and cleaning, and seek partnerships with organizations that can provide real-world datasets. Validate findings against research studies and continuously refine the data generation process for higher accuracy.
2. Develop and implement additional features to ensure data input validation and improve the overall functionality. This may include validating correlations between entities, such as ensuring volunteer resources match help seeker's needs before intervention, data integrity checks, and even automated error handling, thus providing a better user experience.
3. Enable more detailed analysis of correlations to offer deeper insights and improve resource allocation and intervention strategies.

## Author Contributions

Annan Fu: Project planning, background research, database design and ER diagram, data source and assumptions, data validation and cleaning, stored procedure, and report writing.

Danni Hu: Project idea and planning, background research, database design, data source assumption research (q.3 &5), stored procedure for matching volunteers, queries, data analysis and visualization, corresponding parts report writing.

Jiani Guo: The problem study, background research, project planning, report structure and writing, presentation slides structure and writing.

Xing Gao: Project planning, background research, database design and setup, data assumptions, data analysis and visualization, and report writing.

## References


[1] Public Health Agency of Canada. (2023). *Types of violence and abuse.* Government of Canada. https://www.canada.ca/en/public-health/topics/types-violence-abuse.html

[2] Destination Vancouver. (n.d.). Municipalities. Destination Vancouver. https://www.destinationvancouver.com/vancouver/municipalities/

[3] Illinois Criminal Justice Information Authority. (2023). *Victim need report: Service providers' perspectives on the needs of crime victims and service gaps.* Illinois Criminal Justice Information Authority. https://icjia.illinois.gov/researchhub/articles/victim-need-report-service-providers-perspectives-on-the-needs-of-crime-victims-and-service-gaps/

[4] Anti-Violence Project. (n.d.). *Resources. Anti-Violence Project.* https://avp.org/resources/

[5] California State University San Marcos. (n.d.). *Inclusive language guidelines.* https://www.csusm.edu/ipa/surveys/inclusive-language-guidelines.html

[6] WelcomeBC. (n.d.). Language in B.C. Government of British Columbia. https://www.welcomebc.ca/choose-b-c/explore-british-columbia/language-in-b-c

[7] Language family. (2023, June 20). In *Wikipedia*. https://en.wikipedia.org/wiki/Language_family

[8] Statistics Canada. (2016, June 10). *Immigrant status of person.* https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DEC&Id=103339

[9] Stanford University. (n.d.). *Disability language guide*. Stanford University. https://disability.stanford.edu/sites/g/files/sbiybj26391/files/media/file/disability-language-guide-stanford_1.pdf

[10] Statistics Canada. (2023, November 21). *The Daily — Trends in police-reported family violence and intimate partner violence in Canada, 2022*. Www150.Statcan.gc.ca. https://www150.statcan.gc.ca/n1/daily-quotidien/231121/dq231121b-eng.htm

[11] Statistics Canada. (2024, February 12). Proportion of employees who reported experiencing workplace harassment or sexual assault in the workplace, by gender and immigrant status. https://www150.statcan.gc.ca/n1/daily-quotidien/240212/cg-a008-eng.htm

[12] Statistics Canada. (2018a, May 30). Victims of police-reported violent crime and traffic violations causing bodily harm or death, by type of violation, sex of victim and age of victim. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510005101

[13] Saunders, N. R., Guan, J., Macpherson, A., Lu, H., & Guttmann, A. (2020). Association of Immigrant and Refugee Status With Risk Factors for Exposure to Violent Assault Among Youths and Young Adults in Canada. *JAMA network open*, *3*(3), e200375. https://doi.org/10.1001/jamanetworkopen.2020.0375

[14] Fisher, B. S., & Regan, S. L. (2006). The extent and frequency of abuse in the lives of older women and their relationship with health outcomes. *Gerontologist, 46*(2), 200-209. https://doi.org/10.1093/geront/46.2.200

[15] Maher, J., & Segrave, M. (2018). Family violence risk, migration status and 'vulnerability': hearing the voices of immigrant women. Journal of Gender-Based Violence, 2(3), 503-518. Retrieved Jun 22, 2024, from https://doi.org/10.1332/239868018X15375304047178

[16] City Population. (n.d.). Greater Vancouver (Metropolitan area, Metropolitan areas in Canada, Canada) - Population Statistics, Charts, Map and Location. https://www.citypopulation.de/en/canada/britishcolumbia/admin/5915__greater_vancouver/

[17] Wadsworth, P. , Kothari, C. , Lubwama, G. , Brown, C. & Frank Benton, J. (2018). Health and Health Care From the Perspective of Intimate Partner Violence Adult Female Victims in Shelters. *Family & Community Health, 41* (2), 123-133. doi: 10.1097/FCH.0000000000000186.

[18] Anti-Bullying Alliance. (n.d.). Mental health and bullying module. Anti-Bullying Alliance. https://anti-bullyingalliance.org.uk/sites/default/files/uploads/attachments/Mental-health-and-bullying-module-FINAL_0.pdf

[19] Canadian Labour Congress. (2023). National survey reports widespread harassment and violence in workplaces. https://canadianlabour.ca/national-survey-reports-widespread-harassment-and-violence-in-workplaces/

[20] National Network to End Domestic Violence. (2023). *17th annual domestic violence counts report: Full report*. Retrieved from https://nnedv.org/wp-content/uploads/2023/03/17th-Annual-Domestic-Violence-Counts-Report-Full-Report-March-2023.pdf

[21] Canadian Women's Foundation. (2023). The abuse support gap. https://canadianwomen.org/blog/abuse-support-gap/

[22] Rapp-Paglicci, L. A., Roberts, A. R., & Wodarski, J. S. (2002). *Handbook of violence*. John Wiley & Sons. Retrieved from https://books.google.ca/books?id=YtOvUULaNZEC

[23] Lightfoot, E., & Williams, O. (2009). The intersection of disability, diversity, and domestic violence: Results of national focus groups. *Journal of Aggression, Maltreatment & Trauma, 18*(2), 133–152. https://doi.org/10.1080/10926770802675551

[24] Hague, G., Thiara, R., & Mullender, A. (2011). Disabled women, domestic violence and social care: The risk of isolation, vulnerability and neglect. *The British Journal of Social Work, 41*(1), 148–165. https://doi.org/10.1093/bjsw/bcq057