# Practicum-1 Design and Implementation of Relational Database - Contact Tracer

Varshitha Uppula Saisrihitha Yadlapalli (Group 3)

### 1.Introduction

Contact Tracing, a core disease control measure employed by local and state health department personnel for decades, is a key strategy for preventing further spread of COVID-19. A relational database has been implemented to store information about a mobile application user who is tracked based on GPS which would aid to find the users who have come in contact with the infected patients. Surveys are conducted for users who can record their data about temperature, symptoms, etc. Additionally, information about labs and tests conducted in particular cities is maintained.

# 2. Assumptions

The following are the key assumptions that we made with respect to developing the contact tracing database/models:

- **1.** We have modeled our application such that it can be used for contact tracing for the Covid-19 pandemic, but with minor modifications it can be used for generic contact tracing and for analytical use cases as well.
- **2.** An app user is anyone who can use the contact tracing application and can be divided into two categories: a **naïve user** and a **health monitor admin**. An app user can have only a **single phone number**.
- **3.** A naïve user is one who can use the application and is regularly surveyed for symptoms and tracked to check if they have come in contact with an infected user. A naïve user is either **infected**, **not-infected** or **suspected** based on the **infection status** which is given to him/her based on the results of tests conducted by a particular lab that are verified by the health monitor admin.
- **4.** A health monitor admin can view information regarding naïve users and will be notified if a naïve user in a region that the admin is responsible for monitoring has a high risk of having contracted Covid-19.
- **5.** A naïve user will also be notified if he/she is at **risk** of contracting Covid-19. The risk (can be calculated using triggers) of a naïve user having contracted Covid-19 depends upon various factors such as:
  - user symptoms and other information that get recorded as a part of regular surveys
  - medical condition of the user that compromises their immunity
  - travel history of user within a given time period
  - user coming in contact with zero or more infected user based on GPS tracking.
- **6.** A naïve user is tracked based on GPS coordinates and we have considered "gpsLoc" as an attribute that uniquely identifies the coordinates the user on a particular date as we have considered the virus to be active/present where the infected person has been to for at least one day.

# 3.Conceptual model

Conceptual model showing the various entities at a high level is built using Visual Paradigm.

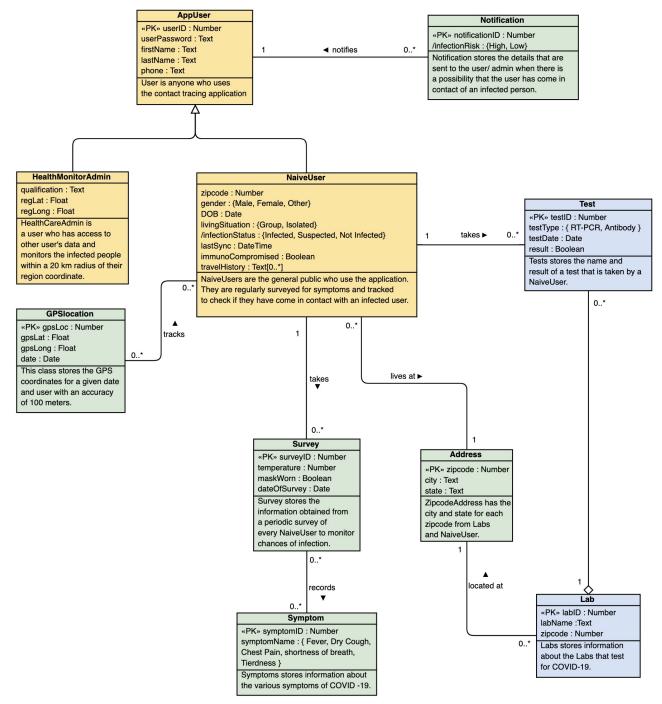


Fig 3.1: Conceptual Model for Contact Tracing in Visual Paradigm

# 4.Logical Model

Logical diagram with Crow's Foot Notation was built using Lucid Chart.

Link to Logical Model (https://app.lucidchart.com/invitations/accept/55d3e1a0-dc6e-4a04-8d6c-6d30657dec3e)

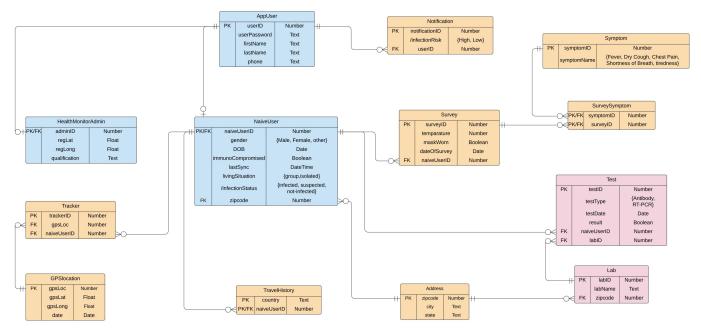


Fig 4.1: Logical Model for Contact Tracing in Lucid Chart

# 5. Relational Schema Definitions

The logical schema shown above is resolved into the below relational schema. A relation R with attributes A1,A2,A3..An is shown as R(A1,A2,A3..An), where the primary key is <u>underlined</u> and foreign keys are shown in *italics*.

AppUser (<u>userID</u>, userPassword, firstName, lastName, phone)

HealthMonitorAdmin (<u>adminID</u>, regLat, regLong, qualification)

NaiveUser (*naiveUserID*, gender, immunoCompromised, DOB, lastSync, livingSituation, infectionStatus, *zipcode*)

Tracker (<u>trackerID</u>, *gpsLoc*, *naiveUserID*)

GPSlocation (gpsLoc, gpsLat, gpsLong, date)

TravelHistory (country, naiveUserID)

Survey (<u>surveyID</u>, temperature, maskWorn, dateOfSurvey, *naiveUserID*)

Symptom (<u>symptomID</u>, symptomName)

SurveySymptom (symptomID, surveyID)

Test (<u>testID</u>, testType, testDate, result, *naiveUserID*, *labID*)

Lab (<u>labID</u>, labName, *zipcode*)

Notification (notificationID, infectionRisk, userID)

# 6.Proof using functional dependencies to show that schema is in BCNF:

A relational schema is considered to be in BCNF if for every one of its dependencies  $X \to Y$  one of the following conditions holds true:

 $X \rightarrow Y$  is a trivial functional dependency (i.e Y is a subset of X)

X is a superkey for the schema

#### **AppUser**

```
userID \rightarrow { userPassword, firstName, lastName, phone } phone \rightarrow { userID, userPassword, firstName, lastName }
```

(Here, there are two candidate keys userID and phone which uniquely identify the non-key attributes and hence satisfies BCNF)

#### **HealthMonitorAdmin**

adminID → { regLat, regLong, qualification }

#### **NaiveUser**

naiveUserID → { gender, immunoCompromised, DOB, lastSync, livingSituation, infectionStatus, zipcode }

#### **GPSlocation**

gpsLoc → { gpsLat, gpsLong, date }

#### **Tracker**

trackerID → { gpsLoc, naiveUserID }

#### **TravelHistory**

country, naiveUserID form a composite primary key (trivial FD hence in BCNF)

#### Survey

surveyID → { temperature, maskWorn, dateOfSurvey, naiveUserID }

#### **Symptom**

symptomID → symptomName

#### SurveySymptom

surveyID, symptomID form a composite primary key (trivial FD hence in BCNF)

#### Test

testID → { testType, testDate, result, naiveUserID, labID }

#### Lab

labID → { labName, zipcode }

#### **Address**

Zipcode → { city, state }

notificationID → { infectionRisk, userID }

# 7. Schema creation and Table definition in MySQL

Tables created in MySQL database using CREATE statements

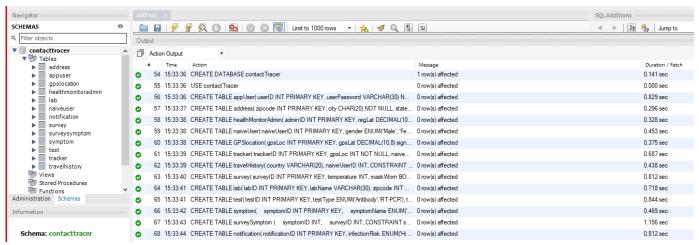


Fig 7.1: Schema creation in MySQL

#### **CREATE Statements**

CREATE DATABASE contactTracer; USE contactTracer;

```
CREATE TABLE appUser(
userID INT PRIMARY KEY,
userPassword VARCHAR(30) NOT NULL,
firstName CHAR(30) NOT NULL,
lastName CHAR(30),
phone VARCHAR(13) NOT NULL UNIQUE
);
```

CREATE TABLE address( zipcode INT PRIMARY KEY, city CHAR(20) NOT NULL, state CHAR(20) NOT NULL );

```
CREATE TABLE healthMonitorAdmin(
adminID INT PRIMARY KEY,
regLat DECIMAL(10,8) signed NOT NULL,
regLong DECIMAL(11,8) signed NOT NULL,
qualification VARCHAR(30),
CONSTRAINT hm_admin FOREIGN KEY(adminID) REFERENCES appUser(userID)
);
```

```
CREATE TABLE naiveUser(
naiveUserID INT PRIMARY KEY,
gender ENUM('Male', 'Female', 'Other'),
dob DATE NOT NULL,
immunoCompromised BOOLEAN NOT NULL,
lastSync DATETIME,
livingSituation ENUM('Group', 'isolated'),
infectionStatus ENUM('infected', 'suspected', 'not-infected') NOT NULL,
zipcode INT NOT NULL,
CONSTRAINT naivU FOREIGN KEY(naiveUserID) REFERENCES appUser(userID),
CONSTRAINT zip FOREIGN KEY(zipcode) REFERENCES address(zipcode)
);
CREATE TABLE GPSlocation(
gpsLoc INT PRIMARY KEY,
gpsLat DECIMAL(10,8) signed NOT NULL,
gpsLong DECIMAL(11,8) signed NOT NULL,
date DATE NOT NULL
);
CREATE TABLE tracker(
trackerID INT PRIMARY KEY,
gpsLoc INT NOT NULL,
naiveUserID INT NOT NULL,
CONSTRAINT gps_track FOREIGN KEY(gpsLoc) REFERENCES GPSlocation(gpsLoc),
CONSTRAINT track naive FOREIGN KEY(naiveUserID) REFERENCES naiveUser(naiveUserID)
);
CREATE TABLE travelHistory(
country VARCHAR(20),
naiveUserID INT,
CONSTRAINT on naive PRIMARY KEY (country, naiveUserID),
CONSTRAINT travel naive FOREIGN KEY(naiveUserID) REFERENCES naiveUser(naiveUserID)
);
CREATE TABLE survey(
surveyID INT PRIMARY KEY,
temperature INT,
maskWorn BOOLEAN NOT NULL,
dateOfSurvey DATE NOT NULL,
naiveUserID INT NOT NULL,
CONSTRAINT survey user FOREIGN KEY(naiveUserID) REFERENCES naiveUser(naiveUserID)
);
```

```
CREATE TABLE lab(
labID INT PRIMARY KEY,
labName VARCHAR(30),
zipcode INT NOT NULL,
CONSTRAINT zip lab FOREIGN KEY(zipcode) REFERENCES address(zipcode)
);
CREATE TABLE test(
testID INT PRIMARY KEY,
testType ENUM('Antibody', 'RT-PCR'),
testDate DATE NOT NULL,
result BOOLEAN NOT NULL,
naiveUserID INT NOT NULL,
labID INT NOT NULL.
CONSTRAINT test lab FOREIGN KEY(labID) REFERENCES lab(labID),
CONSTRAINT test_user FOREIGN KEY(naiveUserID) REFERENCES naiveUser(naiveUserID)
);
CREATE TABLE symptom(
symptomID INT PRIMARY KEY,
symptomName ENUM('Fever', 'Dry Cough', 'Chest Pain', 'Shortness of Breath', 'tiredness')
);
CREATE TABLE surveySymptom (
symptomID INT,
surveyID INT,
CONSTRAINT sy_sr PRIMARY KEY(symptomID,surveyID),
CONSTRAINT sym_sur FOREIGN KEY(symptomID) REFERENCES symptom(symptomID),
CONSTRAINT sur sym FOREIGN KEY(surveyID) REFERENCES survey(surveyID)
);
CREATE TABLE notification(
notificationID INT PRIMARY KEY,
infectionRisk ENUM('High', 'loW'),
userID INT NOT NULL,
CONSTRAINT notif_user FOREIGN KEY(NotificationID) REFERENCES appUser(userID)
);
```

### 8. Constraint Checks

Check to ensure that none of the constraints are violated.

```
INSERT INTO appUser(userID, userPassword, firstName, lastName, phone)
            (57."A0Y71IUZ3F0"."Charde"."Morse"."16460329 5595"):
            INSERT INTO appUser(userID, userPassword, firstName, lastName, phone)
            /* 3) constraint check for ENUM - should fail as "Medium" is not part of the list of permitted values */
            INSERT INTO notification (notificationID, infectionRisk, userID) VALUES (875, "Medium", 43);
   13
               * 4) constraint check for referencing non-existant foreign key */
            INSERT INTO lab (labID, labName, zipCode)
   16
            VALUES (51,"In","29778847");
             /* 5) NOT NULL constraint check */
   19
            INSERT INTO address (zipcode, city, state)
            VALUES ("16503", NULL, "MD");
            /* 6) data type violation */
            INSERT INTO address (zipcode,city,state)
            VALUES ("Hello", "Gaithersburgg", "MDD");
 100% 🗘 40:24
1 2 INSERT INTO appUser(userID,userPassword,firstName,lastName,pho... Error Code: 1062. Duplicate entry '16460329 5595' for key 'appuser.phone' 2 2 INSERT INTO appUser(userID,userPassword,firstName,lastName,pho... Error Code: 1062. Duplicate entry '2' for key 'appuser.PRIMARY'
       2: INSERT INTO notification (notificationID, infectionRisk, userID) VALUE... Error Code: 1265. Data truncated for column 'infectionRisk' at row 1
4 2: INSERT INTO lab (labil) labName_zipCode) VALUES (51;"n";"297788... Error Code: 1452. Cannot add or update a child row: a foreign key constraint fails ('contacttracer'.'lab', CONSTRAINT 'zip_lab' FOREIGN KEY ('zipcode') REFERENCES 'address' ('zipcode'))
5 2: INSERT INTO address (zipcode, city, state) VALUES ("16503", NULL,"... Error Code: 1048. Column 'city' cannot be null
6 21 INSERT INTO address (zipcode,city,state) VALUES ("Hello","Gaithers... Error Code: 1366. Incorrect integer value: 'Hello' for column 'zipcode' at row 1
```

Fig 8.1: Proof of constraint checks

# 9.Inserting Data into database

Relevant dummy data inserted into the database using www.generatedata.com

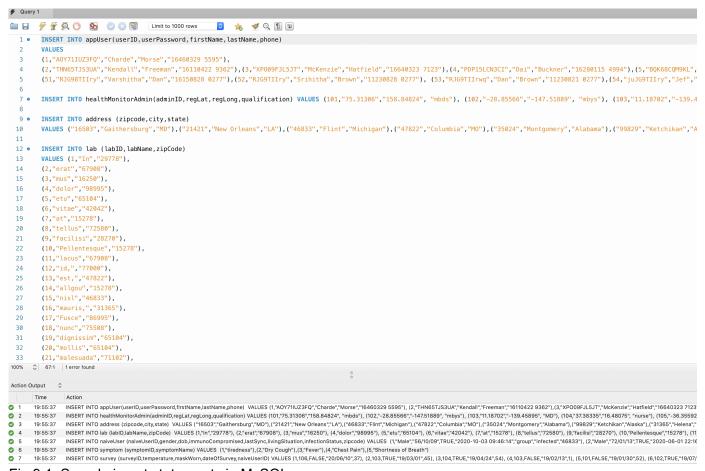


Fig 9.1: Sample insert statements in MySQL

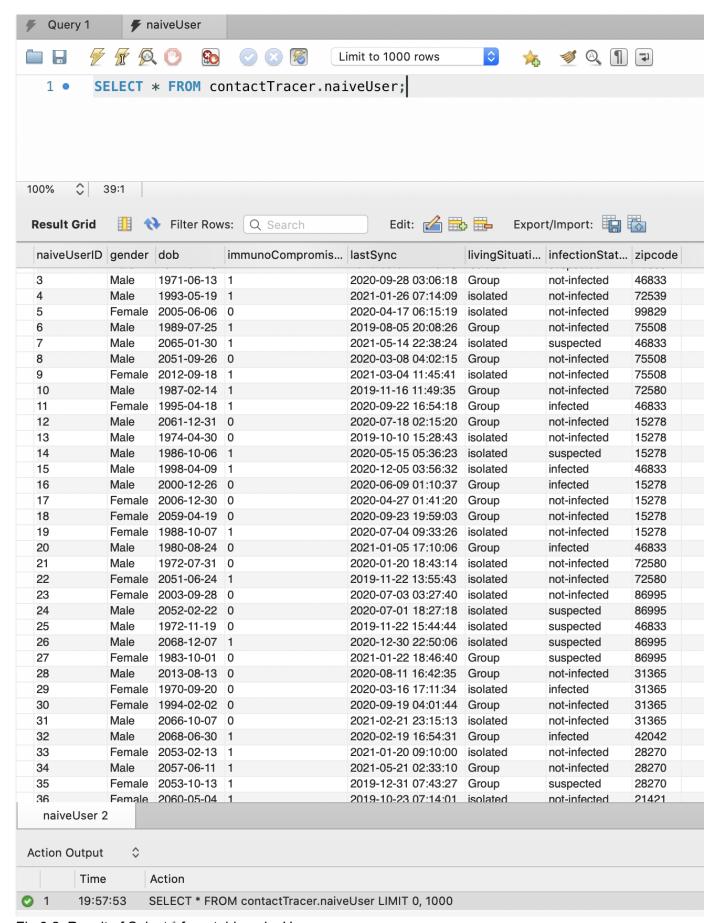


Fig 9.2: Result of Select \* from table naiveUser

# 10.Load Data into R

```
## [1] "address" "appuser" "gpslocation"
## [4] "healthmonitoradmin" "lab" "naiveuser"
## [7] "notification" "survey" "surveysymptom"
## [10] "symptom" "test" "tracker"
## [13] "travelhistory"
```

# 11. Execute Queries in R to show ContactTrace database

1.(Join of 3 tables) Names of users from cities Grand Island, Harrisburg and Fayetteville who have a travel history to any country

```
query1<- "SELECT appUser.firstName , appUser.lastName, address.city, travelHistory.country
   'Country travelled to'
FROM appUser,NaiveUser,travelHistory, address
WHERE appUser.userID = naiveUser.naiveUserID AND
naiveUser.naiveUserID= travelHistory.naiveUserID AND
address.zipcode=naiveUser.zipcode AND
address.city IN ('Grand Island','Fayetteville','Harrisburg');"</pre>
```

```
##
     firstName lastName
                                 city Country travelled to
## 1 Allistair
                  Conway Grand Island
                                                    Japan
## 2
         Maile
                    Lott Grand Island
                                                    Japan
## 3
           Ira
                    Leon Grand Island
                                                   Uganda
## 4
         Alden
                Carson Grand Island
                                                   Sweden
## 5
          Aiko Barrett Grand Island
                                                   Kuwait
                    Duke Grand Island
                                                   Sweden
## 6
         Allen
## 7
         Upton
                 Bradley Grand Island
                                                 Botswana
                Jackson Fayetteville
## 8
        Rashad
                                                   Uganda
## 9
      Chastity
                   Bruce Fayetteville
                                                  Andorra
## 10
                    Pena Fayetteville
                                                   Sweden
        Kessie
## 11 Christen
                   Bowen Harrisburg
                                            Guinea-Bissau
## 12
         Wyatt Lancaster Harrisburg
                                                   Sweden
         Judah
## 13
                    Sims Harrisburg
                                                    Nauru
## 14
         Keith
                    Hess
                           Harrisburg
                                                Australia
```

# 2.(Subquery) Find non-infected users and their living situation who have come in contact with a infected person (*CONTACT TRACING*) based on GPS location and date

```
query2<-"SELECT distinct naiveUser.naiveUserID, naiveUser.livingSituation
FROM naiveUser, tracker, GPSlocation
WHERE naiveUser.naiveUserID = tracker.naiveUserID
AND tracker.gpsLoc = GPSlocation.gpsLoc
AND naiveUser.infectionStatus != 'infected'
AND tracker.gpsLoc IN ( SELECT tracker.gpsLoc FROM naiveUser, tracker
WHERE naiveUser.naiveUserID = tracker.naiveUserID
AND naiveUser.infectionStatus = 'infected');"</pre>
```

```
naiveUserID livingSituation
##
## 1
                3
                             Group
## 2
                4
                          isolated
## 3
                5
                          isolated
## 4
                7
                          isolated
## 5
               10
                             Group
## 6
               14
                          isolated
## 7
               19
                          isolated
## 8
               23
                          isolated
## 9
               24
                          isolated
## 10
                          isolated
               36
               40
## 11
                             Group
## 12
               46
                          isolated
               52
## 13
                             Group
```

#### 3.(Use having and Group By) List of all labs which have conducted more than 2 tests

```
query3<-"SELECT Count(test.testID) AS 'Test Count', lab.labName 'Lab Name'
FROM lab,test
WHERE test.labID = lab.labID
GROUP BY lab.labID
HAVING COUNT(test.testID)>1;"
```

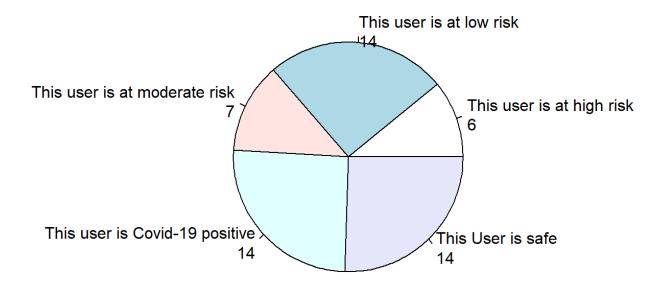
```
##
     Test Count Lab Name
## 1
              2
                      mus
## 2
              2
                    lacus
              2
## 3
                     nisl
## 4
              2
                   mollis
## 5
              2
                       at
## 6
              2
                   cursus
## 7
              2
                    etpoy
              2 malesuada
## 8
```

#### 4.(Complex query using SELECT CASE/WHEN)To determine risk status of a person

```
query4<-"SELECT naiveUserID, infectionStatus, immunoCompromised,
CASE WHEN infectionStatus = 'suspected' AND immunoCompromised = TRUE THEN 'This user is at
high risk'
WHEN infectionStatus = 'suspected' AND immunoCompromised = FALSE THEN 'This user is at mode
rate risk'
WHEN infectionStatus = 'infected' THEN 'This user is Covid-19 positive'
WHEN infectionStatus = 'not-infected' AND immunoCompromised = TRUE THEN 'This user is at lo
w risk'
ELSE 'This User is safe'
END AS Alert
FROM naiveUser;"</pre>
```

##	naivelleerID	infectionStatus	immunoCompromised	Alert
## 1	1	infected	•	This user is Covid-19 positive
## 2	2	suspected	1	This user is at high risk
## 3	3	not-infected	1	This user is at low risk
## 4	4	not-infected	1	This user is at low risk
## 5	5	not-infected	0	This User is safe
## 6	6	not-infected	1	This user is at low risk
## 7	7	suspected	1	This user is at high risk
## 8	8	not-infected	9	This User is safe
## 9	9	not-infected	1	This user is at low risk
## 10	10	not-infected	1	This user is at low risk
## 11	11	infected		This user is Covid-19 positive
## 12	12	not-infected	0	This User is safe
## 13	13	not-infected	0	This User is safe
## 14	14	suspected	1	This user is at high risk
## 15	15	infected	1	This user is Covid-19 positive
## 16	16	infected	0	This user is Covid-19 positive
## 17	17	not-infected	0	This User is safe
## 18	18	not-infected	0	This User is safe
## 19	19	not-infected	1	This user is at low risk
## 20	20	infected	0	This user is Covid-19 positive
## 21	21	not-infected	0	This User is safe
## 22	22	not-infected	1	This user is at low risk
## 23	23	not-infected	0	This User is safe
## 24	24	suspected	0	This user is at moderate risk
## 25	25	suspected	0	This user is at moderate risk
## 26	26	suspected	1	This user is at high risk
## 27	27	suspected	0	This user is at moderate risk
## 28	28	not-infected	0	This User is safe
## 29	29	infected	0	This user is Covid-19 positive
## 30	30	not-infected	0	This User is safe
## 31	31	not-infected	0	This User is safe
## 32	32	infected		This user is Covid-19 positive
## 33	33	not-infected	1	This user is at low risk
## 34		not-infected	1	This user is at low risk
## 35	35	suspected	1	This user is at high risk
## 36	36	not-infected	1	This user is at low risk
## 37	37	suspected	0	This user is at moderate risk
## 38	38	not-infected	0	This User is safe
## 39	39	not-infected	0	This User is safe
## 40	40	<pre>not-infected not-infected</pre>	1 0	This user is at low risk This User is safe
## 41	41 42	infected		This user is Covid-19 positive
## 42 ## 43	42	infected		This user is Covid-19 positive
## 44	44	not-infected	1	This user is at low risk
## 45	44	infected		This user is Covid-19 positive
## 46	46	suspected	1	This user is at high risk
## 40	46	infected	_	This user is at high risk  This user is Covid-19 positive
## 48	48	infected		This user is Covid-19 positive
## 49	49	suspected	0	This user is at moderate risk
## 50	50	infected	_	This user is Covid-19 positive
## 51		infected		This user is Covid-19 positive
## 52	52	suspected	0	This user is at moderate risk
## 53	53	not-infected	1	This user is at low risk
## 54		suspected	0	This user is at moderate risk
## 55	55	not-infected	1	This user is at low risk
55	23	2 2	-	220 20 40 1011 1131

#### Pie-chart depicting user's risk of being infected by Covid-19



#### 5.(Query of choice) GPS coordinates of infected users for Plotting- dummy data

```
query5<-"SELECT gpslocation.gpsLat AS Lattitude, gpslocation.gpsLong AS Longitude
FROM naiveUser, tracker, gpslocation
WHERE naiveUser.naiveUserID = tracker.naiveUserID AND
tracker.gpsLoc = gpslocation.gpsLoc
AND naiveUser.infectionStatus = 'infected';"</pre>
```

```
Lattitude Longitude
##
## 1
     47.40403 131.83584
      22.38534
## 2
               35.53054
     52.10685 142.92803
## 3
## 4 86.66287 22.99641
## 5
     47.40403 131.83584
## 6
       1.88978 -21.19278
## 7 -86.95076 -131.54891
## 8 -56.85797 94.08396
## 9
      36.54954 -157.57636
## 10 39.21384 -66.08290
## 11 -26.39023 139.71786
## 12 -85.96642 -127.08869
## 13 -87.20021
               36.18430
## 14 64.46163 104.56520
## 15 -27.58789 68.46440
## 16 -31.69499 24.58768
## 17 -27.58789 68.46440
```



#### 6.(optional) Select of number of infected cases per city

```
query7<-"SELECT count(naiveUser.naiveUserID) AS 'infected case count', address.city
FROM naiveUser, address
WHERE naiveUser.zipcode= address.zipcode AND
naiveUser.infectionStatus = 'infected'
GROUP BY(address.city);
"</pre>
```

```
infected case count
##
                                    city
                                   Flint
## 1
                       7
## 2
                       1
                            Grand Island
## 3
                       2
                                  Helena
## 4
                       1
                                 Kapolei
## 5
                       1
                             New Orleans
## 6
                             San Antonio
## 7
                       1 South Portland
```