COMP 353 CC
Warm Up Project Report

Group Account: zuc353_1

Pablo Arevalo Escobar (40081955)
Sam Mojaverian (40174101)
Benjamin Pizarro (27644914)
Zachary Bruce (40136585)

1. Design

1.1 Summary of the system

The system is composed of six regions (Africa, Americas, Eastern, Mediterranean, Europe, South-East Asia, Western Pacific), which each have their own set of countries. Countries maintain information related to COVID, as well as articles and statistics. Each country has one government agency that uses the system to provide the latest details regarding COVID-19. Researchers from each country can publish articles to the system.

The system has four types of users, who are able to access the system in different ways. Researchers can add, edit, and delete their articles from the system. Administrators can add, edit and delete users of the system. Users with organization delegate privileges can add, edit, and delete their organization's articles from the system. Regular users can view statistics and articles.

1.2 Initial E/R diagram

The figure 1 E/R diagram was built using a tool called gitmind, while the figure 2 E/R diagram was built using MySQL Workbench's reverse engineering feature.

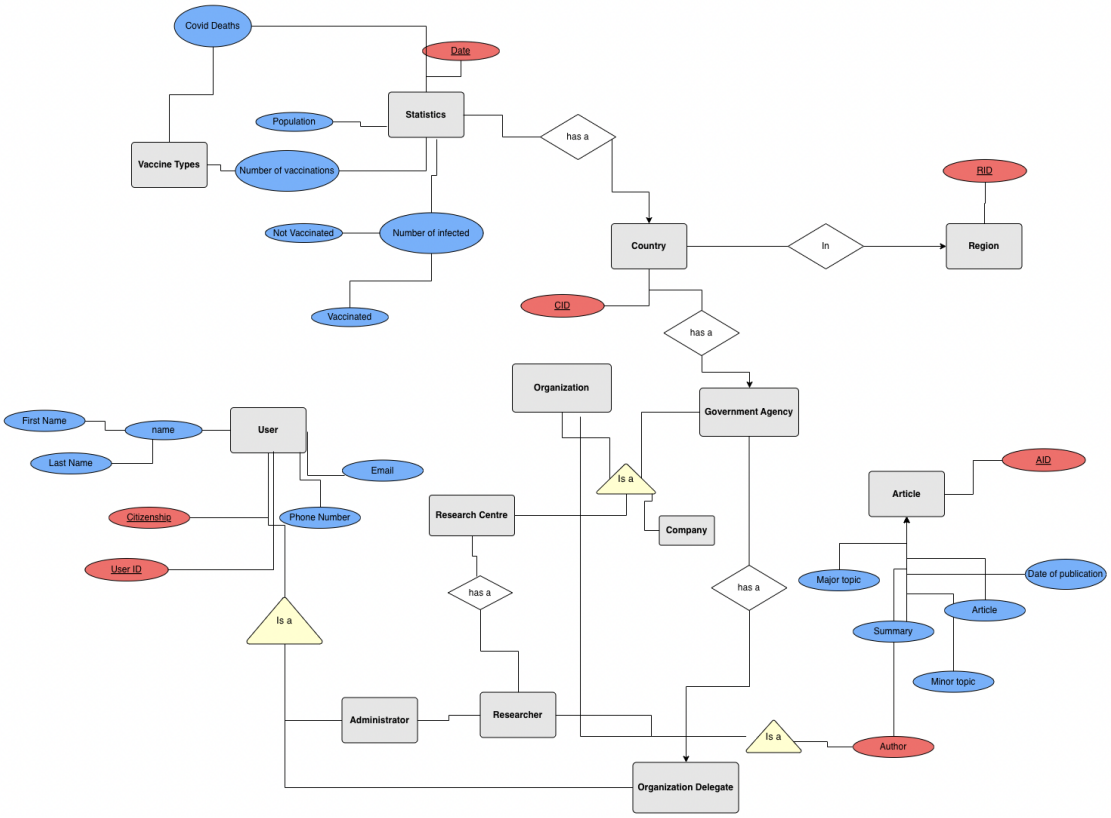


Figure 1. Initial E/R diagram (a full size version is attached in the Appendix)

The initial E/R diagram was converted into the following relations:

- Region(rID, rName)
- Country(rID, cID, cName, Population)
- Organizations(orgID, oName, oType)
- Covid19Stat(statID, cID, rDate, numInfections, numDeaths)
- Vaccines(vID, name)
- VaccineStat(vID, statID, numVaccinations, numVacDeath, numVacInfections)
- Article(aID, Author, majorTopic, minorTopic, summary, content, pubDate)
- Users(uID, cID, fName, lName, phone, email, dob, privilege)

1.3 Design decisions and revisions

In accordance with the minimum requirements, we created the mandatory 6 tables, however, we also included tables for vaccines, general statistics, and vaccine statistics.

Originally, we planned to have a Users table, in which we wanted to include a “Privilege” enum with the different privilege types. However, to stay consistent with the requirements, we opted for separate tables for all privilege types. These include tables for Administrators, Organization Delegates and Regular Users. These tables are similar to the one for Researchers. These tables are linked to the Users table via the foreign key “uID”.

Regions and Countries

The system is composed of six regions:

- Africa
- Americas
- Eastern
- Mediterranean
- Europe
- South-East Asia
- Western Pacific

Each of these regions has a set of countries, which maintain the following information:

- Country ID
- Region ID
- Organization ID
- Name of Country
- Population

Organizations

- Organization ID
- Organization Name
- Organization Delegate ID
- Organization Type

All Users

- User ID
- Country ID
- First Name
- Last Name
- Phone number
- Email
- Date of birth

Researchers

- Researcher ID (user ID)
- Privilege (set by default to Researcher)

Organization Delegates

- Organization Delegate ID (references user ID)
- Privilege (set by default to Organization Delegate)

Administrators

- Administrator ID (references user ID)
- Privilege (set by default to Administrator)

Regular User

- User ID (references user ID)
- Privilege (set by default to Regular)

Article

- Article ID
- Researcher ID
- Author
- Major Topic
- Minor Topic
- Summary
- Content (the actual article itself)
- Publication Date

Statistics Structure:

- Having looked through the queries in part 3, we decided to create a general covid statistics table that we link to countries (via cID) for easier access to any query that required *per country* results. We then linked that table to a vaccine statistics table (via statID) and this way, we can access vaccine statistics per country in an easier manner.

- **COVID19 Statistics**
  - Date of the statistics (used to query latest results)
  - Number of infections
  - Number of deaths

- **Vaccine Statistics**
  - Number of vaccinations
  - Number of vaccinated deaths
  - Number of vaccinated infection

1.4 Revised E/R diagram

After adding the other relations that were necessary to meet the requirements, the following E/R Diagram was generated.

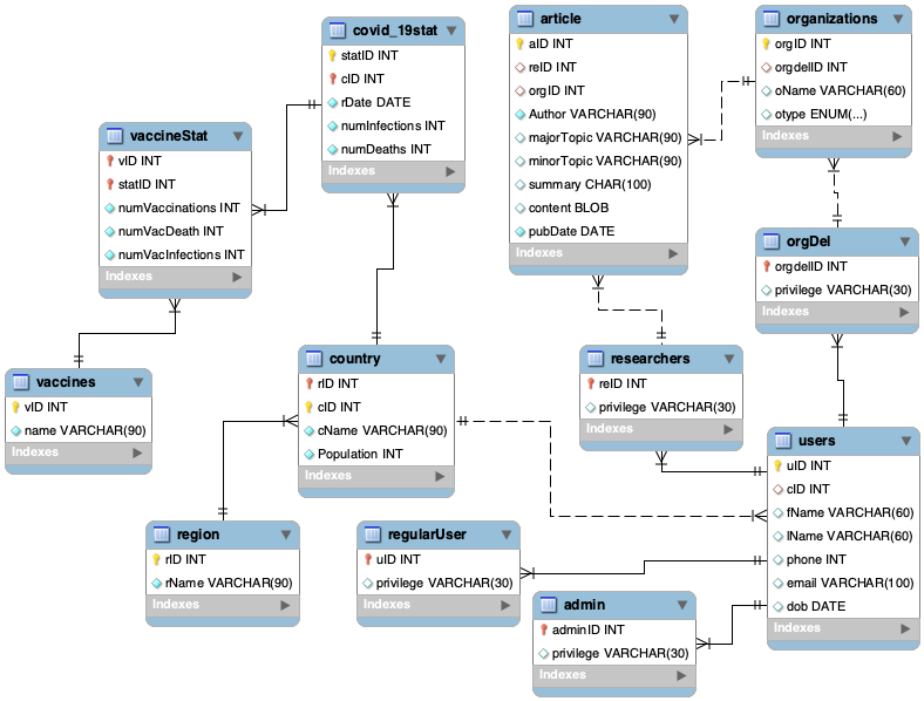


Figure 2. Revised E/R diagram (a full size version is attached in the Appendix)

2. SQL statements used to create the database

```
CREATE TABLE region (
    rID INT UNSIGNED NOT NULL PRIMARY KEY,
    rName VARCHAR(90) NOT NULL
);

CREATE TABLE country (
    rID INT UNSIGNED,
        FOREIGN KEY (rID) REFERENCES region(rID),
    cID INT UNSIGNED AUTO_INCREMENT,
    PRIMARY KEY(cID,rID),
    cName VARCHAR(90) NOT NULL,
    Population INT UNSIGNED NOT NULL
);

CREATE TABLE covid_19stat (
    statID INT UNSIGNED AUTO_INCREMENT,
    cID INT UNSIGNED,
        FOREIGN KEY (cID) REFERENCES country(cID),
    rDate DATE NOT NULL,
    PRIMARY KEY(statID, cID),
    numInfections INT UNSIGNED DEFAULT 0 NOT NULL,
    numDeaths INT UNSIGNED DEFAULT 0 NOT NULL
);

CREATE TABLE vaccines(
    vID INT UNSIGNED AUTO_INCREMENT PRIMARY KEY,
    name VARCHAR (90) NOT NULL
);

CREATE TABLE vaccineStat (
    vID INT UNSIGNED,
        FOREIGN KEY (vID) REFERENCES vaccines (vID),
    statID INT UNSIGNED,
        FOREIGN KEY (statID) REFERENCES covid_19stat (statID),
    PRIMARY KEY(vID,statID),
    numVaccinations INT UNSIGNED DEFAULT 0 NOT NULL,
    numVacDeath INT UNSIGNED DEFAULT 0 NOT NULL,
    numVacInfections INT UNSIGNED DEFAULT 0 NOT NULL
);
```

```
CREATE TABLE users(
    uID INT UNSIGNED AUTO_INCREMENT NOT NULL PRIMARY KEY,
    cID INT UNSIGNED,
        FOREIGN KEY (cID) REFERENCES country (cID),
    fName VARCHAR(60) DEFAULT 'N/A',
    lName VARCHAR(60) DEFAULT 'N/A',
    phone INT UNSIGNED,
    email VARCHAR(100) DEFAULT 'empty',
    dob DATE
);

CREATE TABLE researchers (
    reID INT UNSIGNED PRIMARY KEY,
        FOREIGN KEY (reID) REFERENCES users (uID),
    privilege VARCHAR(30) DEFAULT 'Researcher'

);

CREATE TABLE orgDel (
    orgdelID INT UNSIGNED PRIMARY KEY,
        FOREIGN KEY (orgdelID) REFERENCES users (uID),
    privilege VARCHAR(30) DEFAULT 'Organization Delegate'

);

CREATE TABLE regularUser (
    uID INT UNSIGNED PRIMARY KEY,
        FOREIGN KEY (uID) REFERENCES users (uID),
    privilege VARCHAR(30) DEFAULT 'Regular'
);

CREATE TABLE admin (
    adminID INT UNSIGNED PRIMARY KEY,
        FOREIGN KEY (adminID) REFERENCES users (uID),
    privilege VARCHAR(30) DEFAULT 'Administrator'
);
```

```
CREATE TABLE organizations (
    orgID INT UNSIGNED AUTO_INCREMENT NOT NULL PRIMARY KEY,
    orgdelID INT UNSIGNED,
        FOREIGN KEY (orgdelID) REFERENCES orgDel (orgdelID),
    oName VARCHAR(60),
    otype ENUM('Government','Research Center','Company')
);

CREATE TABLE article (
    aID INT UNSIGNED AUTO_INCREMENT NOT NULL PRIMARY KEY,
    reID INT UNSIGNED,
        FOREIGN KEY (reID) REFERENCES researchers (reID),
    orgdelID INT UNSIGNED,
        FOREIGN KEY (orgdelID) REFERENCES orgDel (orgdelID),
    Author VARCHAR(90) NOT NULL,
    majorTopic VARCHAR(90) DEFAULT 'N/A',
    minorTopic VARCHAR(90) DEFAULT 'N/A',
    summary CHAR(100) DEFAULT 'empty',
    content BLOB,
    pubDate DATE NOT NULL
);
```

3. SQL statements used to express the required queries and transactions

Note: For some queries to function properly, the SQL mode 'ONLY_FULL_GROUP_BY' must be disabled using: `set sql_mode='';`

i)

```
SELECT
    rName, SUM(DISTINCT Population), SUM(numVaccinations), SUM(numDeaths)
FROM
    ((vaccineStat
        INNER JOIN covid_19stat ON vaccineStat.statID = covid_19stat.statID)
        INNER JOIN vaccines ON vaccineStat.vID = vaccines.vID)
            INNER JOIN (country INNER JOIN region ON country.rID = region.rID)
                ON country.cID = covid_19stat.cID
GROUP BY region.rID;
```

ii)

```
SELECT
    rName, cName, Population, SUM(numVaccinations), SUM(numDeaths)
FROM
    ((vaccineStat
    INNER JOIN covid_19stat ON vaccineStat.statID = covid_19stat.statID)
    INNER JOIN vaccines ON vaccineStat.vID = vaccines.vID)
        INNER JOIN (country INNER JOIN region ON country.rID = region.rID)
            ON country.cID = covid_19stat.cID
GROUP BY country.cID
ORDER BY rName,cName ASC;
```

iii)

```
SELECT
    vaccines.name, rName, SUM(numVaccinations), SUM(numVacDeath) AS Deaths
FROM
    ((vaccines
    INNER JOIN vaccineStat ON vaccines.vID = vaccineStat.vID)
    INNER JOIN covid_19stat ON covid_19stat.statID = vaccineStat.statID)
        INNER JOIN (country INNER JOIN region ON country.rID = region.rID)
            ON country.cID = covid_19stat.cID
WHERE rName = 'Americas'
GROUP BY vaccines.name
ORDER BY Deaths ASC;
```

iv)

```
SELECT
    newBag.rName,
    SUM(DISTINCT newBag.Population) AS totalPopulation,
    COUNT(newBag.rID) AS totalResearchers,
    SUM(newBag.totalArticles) AS totalArticles,
    AVG(newBag.totalArticles) averageArticles,
    ((COUNT(newBag.rID) * 100) / Population) * 100 AS percentPopResearcher
FROM (SELECT
         rName,
         region.rID,
         Population,
         COUNT(users.uID) AS totalArticles,
         email
     FROM
         ((region
         INNER JOIN country ON region.rID = country.rID)
         INNER JOIN users ON country.cID = users.cID)
             INNER JOIN researchers ON users.uID = researchers.reID
             LEFT JOIN article ON researchers.reID = article.reID
    GROUP BY users.uID) AS newBag
GROUP BY newBag.rName
ORDER BY percentPopResearcher ASC;
```

v)

```
SELECT cName, rDate, Population, numVaccinations, numInfections, numDeaths
FROM (covid_19stat INNER JOIN country ON covid_19stat.cID = country.cID)
         INNER JOIN vaccineStat ON vaccineStat.statID = covid_19stat.statID
WHERE
    cName =  'Canada'
    ORDER BY rDate DESC;
```

vi)

```
SELECT pubDate, author, majorTopic, minorTopic, summary
FROM article
WHERE author = 'Joe Smith'
ORDER BY pubDate ASC;
```

vii)

```
SELECT MAX(pubDate), author, majorTopic, minorTopic, summary
FROM article
WHERE author = 'Joe Smith';
```

viii)

```
(SELECT admin.privilege, fName, lName, email, phone, dob
FROM (users INNER JOIN admin ON users.uID = admin.adminID))

UNION

(SELECT regularUser.privilege, fName, lName, email, phone, dob
FROM (users INNER JOIN regularUser ON users.uID = regularUser.uID))

UNION

(SELECT researchers.privilege, fName, lName, email, phone, dob
FROM (users INNER JOIN researchers ON users.uID = researchers.reID))

UNION
(SELECT orgDel.privilege, fName, lName, email, phone, dob
FROM (users INNER JOIN orgDel ON users.uID = orgDel.orgdelID));
```

ix)

```
SELECT
    newBag.fName,
    newBag.lName,
    newBag.email,
    newBag.phone,
    newBag.dob,
    newBag.cName,
    newBag.totalArticles,
    MAX(newBag.totalArticles) AS totNumArticles
FROM
    (SELECT region.rID, fName, lName, email, phone, dob, cName,
       COUNT(users.uID) AS totalArticles
    FROM
        ((region
            INNER JOIN country ON region.rID = country.rID)
            INNER JOIN users ON country.cID = users.cID)
            INNER JOIN researchers ON users.uID = researchers.reID
            LEFT JOIN article ON researchers.reID = article.reID
        GROUP BY users.uID) AS newBag
GROUP BY newBag.rID;
```

x)

```
SELECT fName, lName, email, phone, dob, country.cName AS Citizenship
FROM ((users
        INNER JOIN country ON users.cID = country.cID)
        LEFT JOIN researchers ON researchers.reID = users.uID)
            LEFT JOIN article ON researchers.reID = article.reID
WHERE researchers.reID IS NOT NULL AND aID IS NULL;
```

4. Populating the database tables

Other than the Region and Vaccines tables, which only needed to have 6 and 4 tuples, respectively, data was generated using a Python script written to generate appropriate tuples for each table. The Python script is stored at the project repository on GitHub. Due to the large number of tuples generated, only a subset of the tuples generated for each table are shown below. The full set of tuples for each table is visible at the project repository on GitHub.

```
INSERT INTO region VALUES
(1,'Africa'),
(2,'Americas'),
(3,'Eastern Mediterranean'),
(4,'Europe'),
(5,'South-East Asia'),
(6,'Western Pacific');

INSERT INTO vaccines VALUES
(NULL,'Pfizer'),
(NULL,'Moderna'),
(NULL,'AstraZeneca'),
(NULL,'Johnson & Johnson');

INSERT INTO country VALUES
(1,NULL,'Djibouti',2043225),
(1,NULL,'Seychelles',373742413),
(1,NULL,'DR Congo',235635434),
(1,NULL,'Sierra Leone',674894291),
(1,NULL,'South Africa',105064874),
(1,NULL,'Egypt',835110948),
(2,NULL,'United States',798429260),
(2,NULL,'Canada',418634510),
(2,NULL,'Mexico',726308016),
(2,NULL,'Bolivia',587107917),
(2,NULL,'Colombia',580079840),
(2,NULL,'Peru',498147838),
(2,NULL,'Brazil',490733390),
(2,NULL,'Argentina',234462042),
(2,NULL,'Chile',334443743), ... ;
```

```
INSERT INTO covid_19stat VALUES
(NULL,1,'2020-11-11',1024814,662206),
(NULL,2,'2021-06-22',407743,159623),
(NULL,3,'2021-03-14',1462771,223974),
(NULL,4,'2022-03-06',908139,7239),
(NULL,5,'2020-05-05',389949,127488),
(NULL,6,'2020-10-16',1015698,69818),
(NULL,7,'2021-10-10',1101535,323851),
(NULL,8,'2021-02-28',1279198,115909),
(NULL,9,'2020-07-23',1268273,32464),
(NULL,10,'2021-02-03',542273,576917),
(NULL,11,'2022-02-06',1096437,51061),
(NULL,12,'2022-10-03',407093,124320),
(NULL,13,'2022-04-14',496818,66351),
(NULL,14,'2021-05-21',360269,29935),
(NULL,15,'2020-09-16',902054,479199), ... ;

INSERT INTO vaccineStat VALUES
(3,1,490646,80,24234),
(2,2,324650789,393831,88915723),
(2,3,57444268,112178,19321417),
(3,4,439446512,475967,153694530),
(3,5,16749181,14535,4912683),
(3,6,238582480,401516,87920356),
(4,7,592356914,702224,161535842),
(2,8,351816540,352493,61672506),
(3,9,381777589,887681,139122148),
(2,10,224675771,158642,59990774),
(2,11,256076651,206408,49015801),
(4,12,439304304,369979,93717915),
(2,13,265973513,655132,102532121),
(3,14,142229802,122459,30566153), ... ;

INSERT INTO users VALUES
(NULL,18,'Joe','Smith',92829,'joesmith@gmail.com','1970-06-24'),
(NULL,26,'Ruadhan','McCormack',84354,'RuadhanMcCormack@gmail.com','2008-06-12'),
(NULL,9,'Bryden','Tandy',48072,'BrydenTandy@gmail.com','1989-08-24'),
(NULL,11,'Zack','Root',92298,'ZackRoot@gmail.com','1970-11-19'),
(NULL,69,'Kameron','Tandy',22409,'KameronTandy@gmail.com','1986-10-09'),
(NULL,22,'Tegan','Schmitt',80186,'TeganSchmitt@gmail.com','1981-10-28'),
(NULL,14,'Siddharth','Hancock',21118,'SiddharthHancock@gmail.com','1976-03-14'),
(NULL,22,'I','Cast',14347,'ICast@gmail.com','2006-10-25'),
```

```
(NULL,82,'Leyland','Sagar',40954,'LeylandSagar@gmail.com','2000-07-20'),
(NULL,10,'Sayf','Van Zandt',14366,'SayfVan Zandt@gmail.com','1980-04-24'),
(NULL,38,'Sam','Chapman',35983,'SamChapman@gmail.com','2005-08-04'),
(NULL,34,'Moayd','Wakefield',74198,'MoaydWakefield@gmail.com','1986-08-03'),
(NULL,25,'Thierry','Yang',22766,'ThierryYang@gmail.com','1978-02-12'),
(NULL,64,'Pawel','Nelson',91319,'PawelNelson@gmail.com','1999-10-09'),
(NULL,92,'Harikrishna','Yocum',88314,'HarikrishnaYocum@gmail.com','1970-07-20'),
... ;

INSERT INTO researchers VALUES
(1, DEFAULT),
(8, DEFAULT),
(35, DEFAULT),
(65, DEFAULT),
(92, DEFAULT),
(90, DEFAULT),
(6, DEFAULT),
(53, DEFAULT),
(88, DEFAULT),
(66, DEFAULT),
(86, DEFAULT),
(79, DEFAULT),
(23, DEFAULT),
(91, DEFAULT),
(36, DEFAULT), ... ;

INSERT INTO orgdel VALUES
(68, DEFAULT),
(17, DEFAULT),
(49, DEFAULT),
(32, DEFAULT),
(5, DEFAULT),
(89, DEFAULT),
(80, DEFAULT),
(52, DEFAULT),
(18, DEFAULT),
(10, DEFAULT),
(12, DEFAULT),
(48, DEFAULT),
(25, DEFAULT),
(33, DEFAULT), ... ;
```

```
INSERT INTO admin VALUES
(4, DEFAULT),
(51, DEFAULT),
(73, DEFAULT),
(83, DEFAULT),
(71, DEFAULT);

INSERT INTO regularUser VALUES
(2, DEFAULT),
(3, DEFAULT),
(9, DEFAULT),
(11, DEFAULT),
(13, DEFAULT),
(14, DEFAULT),
(15, DEFAULT),
(16, DEFAULT),
(19, DEFAULT),
(20, DEFAULT),
(21, DEFAULT),
(22, DEFAULT),
(24, DEFAULT),
(26, DEFAULT),
(27, DEFAULT), ... ;

INSERT INTO organizations VALUES
(1,68,'Company Incorporated','Company'),
(2,17,'National Government Agency','Government'),
(3,49,'National Government Agency','Government'),
(4,32,'National Government Agency','Government'),
(5,5,'Company Incorporated','Company'),
(6,89,'Company Incorporated','Company'),
(7,80,'Company Incorporated','Company'),
(8,52,'Company Incorporated','Company'),
(9,18,'Company Incorporated','Company'),
(10,10,'Company Incorporated','Company'),
(11,12,'Company Incorporated','Company'),
(12,48,'National Government Agency','Government'),
(13,25,'National Government Agency','Government'),
(14,33,'Company Incorporated','Company'),
(15,98,'National Government Agency','Government'), ... ;
```

```
INSERT INTO article VALUES
(NULL,1,NULL,'Joe Smith','MajorTop','MinorTop','This is a summary','This is the
 content of the article','2021-08-18'),
(NULL,NULL,31,'UAE Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2022-08-07'),
(NULL,7,NULL,'Siddharth Hancock','MajorTop','MinorTop','This is a summary','This
 is the content of the article','2021-05-09'),
(NULL,NULL,89,'Switzerland Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2022-02-06'),
(NULL,NULL,32,'Jordan Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2022-07-20'),
(NULL,NULL,17,'Cuba Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2022-06-28'),
(NULL,35,NULL,'Isaa Hylan','MajorTop','MinorTop','This is a summary','This is the
 content of the article','2022-01-01'),
(NULL,NULL,32,'Jordan Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2022-12-10'),
(NULL,7,NULL,'Siddharth Hancock','MajorTop','MinorTop','This is a summary','This
 is the content of the article','2020-06-13'),
(NULL,NULL,10,'Colombia Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2021-08-22'),
(NULL,65,NULL,'Cadyn Christensen','MajorTop','MinorTop','This is a summary','This
 is the content of the article','2020-03-12'),
(NULL,NULL,33,'South Africa Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2020-09-03'),
(NULL,NULL,12,'Germany Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2021-06-09'),
(NULL,40,NULL,'Philip Miller','MajorTop','MinorTop','This is a summary','This is
 the content of the article','2020-04-26'),
(NULL,NULL,100,'Italy Government Agency','MajorTop','MinorTop','This is a
 summary','This is the content of the article','2020-09-28'), ... ;
```

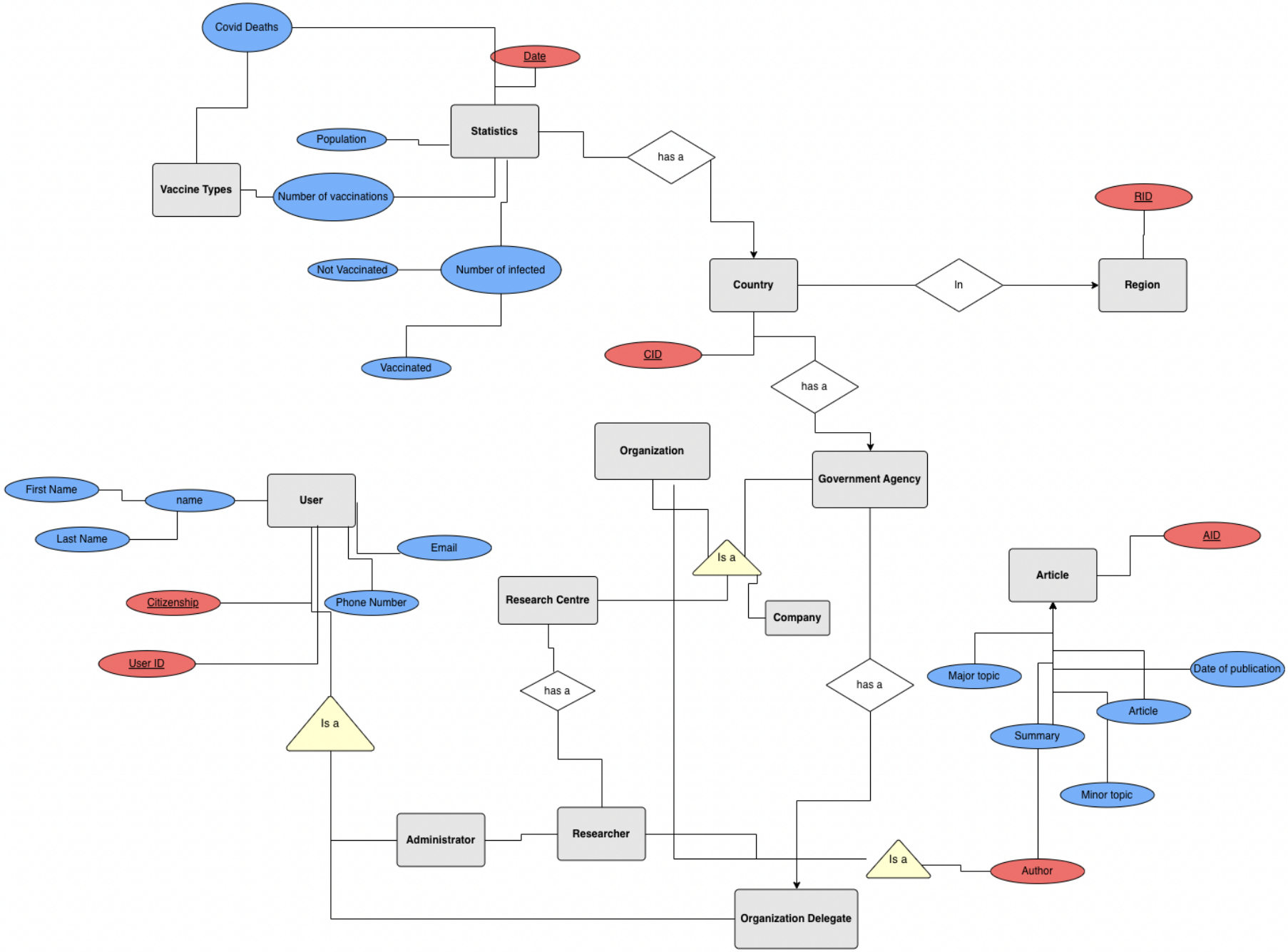
5. SELECT COUNT(*) FROM R

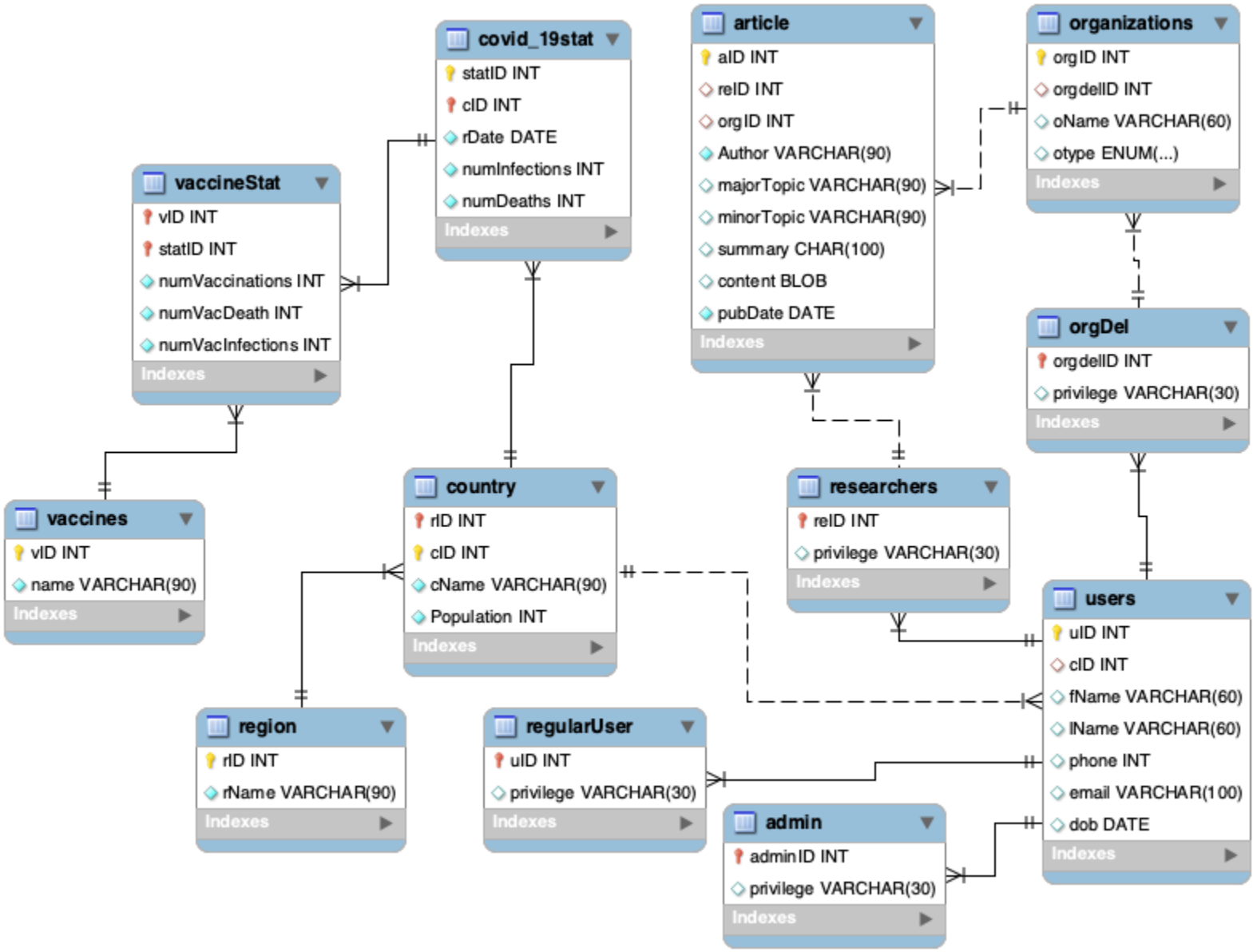
The result of `SELECT COUNT(*)` for each relation in the database is reported below.

```
SELECT COUNT(*)
FROM admin;
                  RESULT:     COUNT()
                                    5

SELECT COUNT(*)
FROM article;
                  RESULT:     COUNT()
                                  100

SELECT COUNT(*)
FROM country;
                  RESULT:     COUNT()
                                   95

SELECT COUNT(*)
FROM covid_19stat;
                  RESULT:     COUNT()
                                  950

SELECT COUNT(*)
FROM region;
                  RESULT:     COUNT()
                                    6

SELECT COUNT(*)
FROM organizations;
                  RESULT:     COUNT()
                                   20

SELECT COUNT(*)
FROM orgDel;
                  RESULT:     COUNT()
                                   20
```

```
SELECT COUNT(*)
FROM regularUser;
                  RESULT:     COUNT()
                                   55

SELECT COUNT(*)
FROM researchers;
                  RESULT:     COUNT()
                                   20

SELECT COUNT(*)
FROM users;
                  RESULT:     COUNT()
                                  100

SELECT COUNT(*)
FROM vaccines;
                  RESULT:     COUNT()
                                    4

SELECT COUNT(*)
FROM vaccineStat;
                  RESULT:     COUNT()
                                  950
```

APPENDIX


Covid Deaths
Date
Statistics
Population
has a
Vaccine Types
Number of vaccinations
RID
Not Vaccinated
Number of infected
Country
In
Region
Vaccinated
CID
has a
Organization
Government Agency
First Name
name
User
Last Name
Email
Is a
AID
Article
Citizenship
Phone Number
Research Centre
Company
User ID
has a
has a
Major topic
Date of publication
Is a
Article
Summary
Minor topic
Administrator
Researcher
Is a
Author
Organization Delegate


covid_19stat
statID INT
cID INT
rDate DATE
numInfections INT
numDeaths INT
Indexes
article
aID INT
reID INT
orgID INT
Author VARCHAR(90)
majorTopic VARCHAR(90)
minorTopic VARCHAR(90)
summary CHAR(100)
content BLOB
pubDate DATE
Indexes
organizations
orgID INT
orgdelID INT
oName VARCHAR(60)
otype ENUM(...)
Indexes
vaccineStat
vID INT
statID INT
numVaccinations INT
numVacDeath INT
numVacInfections INT
Indexes
orgDel
orgdelID INT
privilege VARCHAR(30)
Indexes
vaccines
vID INT
name VARCHAR(90)
Indexes
country
rID INT
cID INT
cName VARCHAR(90)
Population INT
Indexes
researchers
reID INT
privilege VARCHAR(30)
Indexes
users
uID INT
cID INT
fName VARCHAR(60)
lName VARCHAR(60)
phone INT
email VARCHAR(100)
dob DATE
Indexes
region
rID INT
rName VARCHAR(90)
Indexes
regularUser
uID INT
privilege VARCHAR(30)
Indexes
admin
adminID INT
privilege VARCHAR(30)
Indexes