Pandemic Information System Model

SYSTEMS AND METHODS FOR BIG AND UNSTRUCTURED DATA

PROF. MARCO BRAMBILLA

FIRST DELIVERY NEO4J PROJECT

November 2021

Avci Oguzhan - xxxxxxxx Gentile Nicole - 10594355 Rigamonti Davide - 10629791 Singh Raul - xxxxxxxx Tagliaferri Mattia - xxxxxxxx add personal codes



Contents

1	Intr	oduction 2	2
	1.1	Problem Specification	2
	1.2	Hypoteses	2
2	Data	abase	3
	2.1	ER Diagram	3
	2.2	Dataset description	3
	2.3	Queries	3
	2.4	Commands	1
3	App	lication	5
	3.1	Description	5
	3.2	User Guide	5
4	Refe	erences and sources	5
T	odo	list	
ad	d pers	onal codes	
	_	population size	2
		accesses from the same person can be generated at the same time	2
		ill true?	2
		ill true?	
		1 add more	

1 Introduction

1.1 Problem Specification

We represented a population of XXX people in the USA.

The database receives data coming from tracing applications that use sensor in smartphones, wearable objects or other devices to understand whether two people had a contact; data includes date and time of the contact.

Some places, like restaurants, theaters and hospitals, collect date and name of visiting people.

The idea is to use this database so that, if a person gets positive, we can understand who are all people who had a contact with him/her. Contacts are of different types: in family (since people from the same family are always in contact), in a location (if someone is positive, we alert people who were in the same location on the same day) or are given by an application using sensors. Data are recorded from 02/2020 and can be used for analytical purposes.

1.2 Hypoteses

Vaccine date, date of the last contagion, date of the last negative test and healing date are optional fields.

A person is considered either *infected* if they have a contagion date, *healed* if they have an healing date and a contagion date or neither if they have none yet.

We assumed that people can do tests without being infected (or after healing) and that people can decide not to get the vaccine but vaccinated people can still get infected.

We also assumed that people can get covid at most once (realistic if we consider perfect antibodies), in this way it's easier to store and retrieve data in order to build statistics.

Members of a family are assumed to be all the people who live together, relatives who see each other very often or roommates. Obviously, all members of a family live in the same city and are considered "always in contact".

When adding data about people going places, we do not consider distances between the city where they live in and the location they go to, so they can exist in different places at the same time.

Regarding the places we have: people who got covid after 18/10 went to places from 18/09/2021 to 17/10/2021, since we assume that later they are in quarantine; all the others went to visit locations from 18/09/2021 to 17/11/2021. If a person becomes positive, we alert all people who were in the same location in a 4 hour span of time centered on the time of arrival of the infected person.

For simplicity, for people who are in hospital date of hospitalization and contagion coincide; we also assume they leave the hospital on the healing date.

In order to populate the database, we imposed that every person visits from 6 to 10 locations.

fill with population size

check if accesses from the same person can be generated at the same time

is this still true?

is this still true?

we could add more

2 Database

2.1 ER Diagram

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.2 Dataset description

We used three types of nodes: Person, Location and City. People are characterized by their name and surname, birthdate, city, email, social security number, vaccine date, date of the last contagion, date of the last negative test and healing date. People can be linked by family relationships. Locations have a name and a type (restaurant, theater, hospital); type is used so that the dataset can be easily expanded with new types of location. Possible relationships are:

- WENT_TO to track people who visited a location at a certain time;
- IN_FAMILY for family relationships;
- IS_IN between locations and cities;
- LIVES_IN to link people to the city they live in;
- HAS_MET to indicate contacts between people using tracing app or devices;
- IS_HOSPITALIZED_IN to indicate people who are/were hospitalized for covid reasons.

2.3 Queries

2.4 Commands

3 Application

3.1 Description

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3.2 User Guide

4 References and sources