Data mining for the study of the Epidemic (SARS-CoV-2) COVID-19: Algorithm for the identification of patients speaking the native language in the Totonacapan area – Mexico

Medel-Ramírez, Carlos and Medel-López, Hilario

Universidad Veracruzana / Instituto de Investigaciones y Estudios Superiores Económicos y Sociales, Universidad Veracruzana / Instituto de Antropología

24 July 2020

Data Article

Data mining for the study of the Epidemic (SARS-CoV-2) COVID-19: Algorithm for the identification of patients speaking the native language in the Totonacapan area – Mexico

Authors

Carlos Medel-Ramírez [1], Hilario Medel-López [2]

Affiliations

1. Universidad Veracruzana / Instituto de Investigaciones y Estudios Superiores Económicos y Sociales
2. Universidad Veracruzana / Instituto de Antropología

Corresponding author

Carlos Medel-Ramírez (cmedel@uv.mx)

Abstract

The importance of the working document is that it allows analyzing the information and the status of the cases associated with (SARS-CoV-2) COVID-19 as data open to the municipal government and especially in the Totonacapan Zone in Mexico, from the registry patient diary, according to age, sex, comorbidities and condition of (SARS-CoV-2) COVID-19, according to the following characteristics: a) Positive, b) Negative, c) Suspect. Likewise, it presents information on the identification of an outpatient and / or hospitalized patient, attending to their medical development, identifying: a) Recovered, b) Deaths and c) Assets. Data analysis is carried out by applying a data mining algorithm, which provides the information, fast and timely, necessary for the estimation of the healthcare scenarios of (SARS-CoV-2) COVID-19.

Specifications Table

Subject	Infectious Diseases
Specific subject area	Information from the Viral Respiratory Diseases Epidemiological Surveillance System for (SARS-CoV-2) COVID-19 in the Totonacapan Zone - Mexico
Type of data	Table Figure
How data were acquired	Government of Mexico. Health Secretary. Databases Covid-19 México https://datos.gob.mx/busca/dataset/informacion-referente-a-casos-covid-19-en-mexico/resource/e8c7079c-dc2a-4b6e-8035-08042ed37165 Instruments: Software Orange Data Mining version 3.26.0 https://orange.biolab.si Make and model and of the instruments used: Algorithm for the identification of patients according to following characteristics: a) Positive, b) Negatives, c) Suspects. Likewise, it presents

	information regarding the identification of an outpatient and / or hospitalized patient, attending to their medical development, identifying: a) Recovered, b) Deaths and c) Assets
Data format	The information is presented in raw in CVS format, the Ministry of Health of Mexico since April 14, 2020 published the cases associated with (SARS-CoV-2) COVID-19 as open data. The data processing corresponds to the records on the epidemic (SARS-CoV-2) COVID-19 at 15 July 2020. The treatment of the information is carried out through the application software for data mining Orange version 3.26.0, in which the algorithm for the analysis of information is filtered to present the current scenario in Mexico of the SARS-CoV-2 (COVID 19).
Parameters for data collection	The study area corresponds to the variant identified as Totonaco de la Costa and that identifies the Totonacapan Zone from the political and administrative integration of the municipalities of Cazones de Herrera, Coahuitlán, Coatzintla, Coxquihui, Coyutla, Chumatlán, Espinal, Filomeno Mata, Gutiérrez Zamora, Mecatlán, Papantla, Poza Rica, Tecolutla, Tihuatlán and Zozocolco de Hidalgo. The information is presented at the municipal level, and particularly in the municipalities that make up the Totonacapan Zone from a daily registry of patients, according to age, sex, comorbidities, for the condition for positive results of (SARS-CoV-2) COVID -19, presenting the following characteristics: a) Positive, b) Negatives, c) Suspects. It also presents information on the identification of an outpatient and / or hospitalized patient, attending to their medical development, identifying: a) Recovered, b) Deaths and c) Assets.
Description of data collection	This information is filtered to present the current scenario in the Totonacapan Zone in Mexico of the SARS-CoV-2 (COVID 19) in a fast and timely manner, to support public decision-making in health matters.
Data source location	Institution: Universidad Veracruzana / Instituto de Investigaciones y Estudios Superiores Económicos y Sociales. Country: México
Data accessibility	Raw data can be retrieved from the Github repository https://github.com/CMedelR

Value of the Data

- The Algorithm for the identification of patients (SARS-CoV-2) COVID 19 in the Totonacapan Zone in Mexico allows to analyze at the municipal level the registry of patients, according to age, sex, comorbidities, indigenous speech condition and condition of (SARS-CoV-2) COVID-19 according to the following characteristics: a) Positive, b) Negative, c) Suspicious, as well as presenting information on the identification of an outpatient and / or hospitalized patient, attending to their medical development in Phase 3 and Phase 4, in a fast and timely manner, to support public decision-making in health matters.
- Taking into account their strategic roles in public health and researchers can use the data from this study to identify the action scenario for decision-making in the combat of (SARS-CoV-2) COVID 19 in Phase 3 and Phase 4.
- The importance of data analysis is that it allows identifying the cases (SARS-CoV-2) COVID-19 in Mexico is concentrated on a daily for patients of (SARS-CoV-2) COVID-19 and and allows preparing action scenarios for making public health policy decisions to combat SARS-CoV-2) COVID-19 in in the Totonacapan Zone in Mexico.

Data Description

The source of information on the number of registered cases of (SARS-CoV-2) COVID-19 at 15 July 2020 for Mexico comes from the website https://datos.gob.mx/busca/dataset/informacion-referente-a-casos-covid-19-en-mexico/resource/e8c7079c-dc2a-4b6e-8035-08042ed37165 by the Ministry of Health, with the participation of the National Council for Science and Technology (CONACYT), the Center for Research in Geospatial Information Sciences (CENTROGEO), the National Laboratory for Geo-Intelligence (GEOINT), the Data Laboratory of the National Laboratory for Geointelligence (DataLab), where the registry of COVID-19 cases (SARS-CoV-2) COVID-19 is concentrated, and is the official means of communication and information on the epidemic in Mexico.

The information of the cases (SARS-CoV-2) COVID-19 in Mexico is concentrated on a daily basis since April 19, 2020, communication and official information on the epidemic in Mexico, the data are presented at the municipal, state and national levels, with a daily registry of patients, according to age, sex, comorbidities, for the condition of (SARS-CoV-2) COVID-19 according to the following characteristics: a) Positive, b) Negatives, c) Suspects. Likewise, it presents information regarding the identification of an outpatient and / or hospitalized patient, attending to their medical development. The data processing corresponds to the records on the epidemic (SARS-CoV-2) COVID-19 at July 15, 2020. The treatment of the information is carried out through the application software for data mining and visual programming Orange Data Mining version 3.26.0. Orange Data Mining is a machine learning and data mining suite for data analysis through Python scripting and visual programming. [1]

According to (WHO, 2020) the (SARS-CoV-2) COVID-19 disease pattern presents 4 scenarios identified from the confirmation of Laboratory Diagnosis: a) Not Infected or b) Infected, in this finally, the following categories are observed, taking into account age and specific comorbidities in each case: a) Mild Infection, a) Moderate Infection, c) Severe Infection and d) Critical Infection.

Depending on the category observed in Patients who have a Confirmation of Infected, as in the case of a) or b) it can assume the character of Outpatient, so the strategy is isolation or "quarantine" at home, where the result It is hoped that he will recover. Regarding the Patients who have a Confirmation of Infected, in categories c) and d) they assume the character of Hospitalized Patient, with a probability of requiring care in Intensive Care Units and requiring Intubation, and where it is hoped to save as many patients as possible.

The importance of the research is that it allows identifying the action scenario for making public health policy decisions to combat CO(SARS-CoV-2) COVID-19, since they consider the following states of process in medical treatment, in order to carry out the Estimate of Scenarios for Medical Care of the (SARS-CoV-2) COVID-19 in the Totonacapan Zone in Mexico under the following premises of hospital care:
1. A patient with a positive (SARS-CoV-2) COVID-19 laboratory diagnosis can be considered: a) Outpatient, or b) Hospitalized.
2. If the (SARS-CoV-2) COVID-19 Positive patient is Hospitalized, the following should be considered: a) Enter the Intensive Care Unit or b) Do not enter the Intensive Care Unit.
3. If the (SARS-CoV-2) COVID-19 Positive patient is Hospitalized and Entered into the Intensive Care Unit, the following should be considered: a) The patient requires intubation or b) The patient does NOT require intubation.

Methods

The information is presented in raw in CVS format, the Ministry of Health of Mexico. The data processing corresponds to the records on the epidemic (SARS-CoV-2) COVID-19 at 15 July 2020. The treatment of the information is carried out through the application software for data mining Orange version 3.26.0, in which the algorithm for the analysis of information are developed and it is filtered to present the current scenario in the Totonacapan Zone in Mexico of the SARS-CoV-2 (COVID 19). In this way, the algorithm that is presented allows us to project the requirements for the use of installed infrastructure in the face of the growing requirement for patient care Positive (SARS-CoV-2) COVID-19, allowing the identification of scenarios at the national, state and municipal levels. The construction of the algorithm is based on the following definitions. (See Figure 1 at the end of the section).

Definition 1: Total Patients to consider in Model (SARS-CoV-2) COVID-19.- It is the number of total patients according to the confirmatory laboratory result or not of (SARS-CoV-2) COVID-19).

Be:
TP SARS-CoV-2 i j = Total patients according to (SARS-CoV-2) COVID-19 confirmatory laboratory result

Which consists of:
TP SARS-CoV-2 i j = (P + SARS-CoV-2 i j) + (P- SARS-CoV-2 i j) + (Px SARS-CoV-2 i j), where: i = State, j = Totonacapan Zone

Of which:
P+ SARS-CoV-2 i j = Patient with a positive (SARS-CoV-2) COVID-19 result in the State, Totonacapan Zone
P- SARS-CoV-2 i j = Patient with negative (SARS-CoV-2) COVID-19 result in the State, Totonacapan Zone
Px SARS-CoV-2 i j = Patient with pending confirmation (SARS-CoV-2) COVID-19 in the State, Totonacapan Zone

Definition 2: Identification of a suspected (SARS-CoV-2) COVID-19 case.- This is the patient who undergoes an initial qualification according to the initial diagnostic characteristics indicated in the case definitions for surveillance by the World Health Organization for primary care of (SARS-CoV-2) COVID-19 cases.

Be:
CsCOVID 19 (SARS-CoV-2) = Patient with initial classification as a suspected case of (SARS-CoV-2) COVID 19

Where:
Cs (SARS-CoV-2) COVID-19 = Cs (SARS-CoV-2) COVID-19 Type 1 + Cs (SARS-CoV-2) COVID-19 Type 2 + Cs (SARS-CoV-2) COVID-19 Type 3

Of which:
According to the World Health Organization, there are 3 categories (identified as Type 1, Type 2 and Type 3) to identify suspected cases of (SARS-CoV-2) COVID-19, defined below:

1. Cs (SARS-CoV-2) COVID-19 Type 1.- Is a patient with acute respiratory disease (fever and at least one sign / symptom of respiratory disease, with no other aetiology that fully explains the clinical presentation and a history of travel or residence in a country / area or territory that reports local transmission of COVID-19 disease during the 14 days prior to the on set of symptoms.
2. Cs (SARS-CoV-2) COVID-19 Type 2.- He is a patient with an acute respiratory disease, who has been in contact with a confirmed or probable COVID-19 case in the last 14 days before the onset of symptoms.

4

3. Cs (SARS-CoV-2) COVID-19 Type 3.- Is a patient with severe acute respiratory infection (fever and at least one sign / symptom of respiratory illness (e.g. cough, shortness of breath) and requiring hospitalization and without another etiology that fully explains the clinical presentation.

Definition 3: Total Patients to consider in the (SARS-CoV-2) COVID-19 Model .- It is the number of total patients according to the confirmatory laboratory result or not of (SARS-CoV-2) COVID-19).

Be:
TP SARS-CoV-2 i j = Total patients according to confirmatory laboratory result or not of (SARS-CoV-2) COVID-19

Which consists of:
TP SARS-CoV-2 i j = (P + SARS-CoV-2 i j) + (P- ARS-CoV-2 i j) + (Px ARS-CoV-2 i j) , where: i = State, j = Totonacapan Zone
Of which:
P + SARS-CoV-2 i j = Patient with a positive (SARS-CoV-2) COVID-19 result in the State, Totonacapan Zone
P- ARS-CoV-2 i j = Patient with negative (SARS-CoV-2) COVID-19 result in the State, Totonacapan Zone
Px ARS-CoV-2 i j = Patient with pending confirmation (SARS-CoV-2) COVID-19 in the State, Totonacapan Zone

Definition 4: Positive Patients for (SARS-CoV-2) COVID-19 i j.- It is the number of patients with laboratory results with positive confirmation for (SARS-CoV-2) COVID-19 i j .

It has:
P + SARS-CoV-2 i j = Patient with a positive (SARS-CoV-2) COVID-19 result in the State, Totonacapan Zone

Definition 5.- Medical Treatment Strategy for a patient with positive laboratory confirmation for (SARS-CoV-2) COVID-19 i j .- It is the Action Plan in Medical Treatment for a patient with positive laboratory confirmation for SARS-CoV-2 in attention to your degree of infection and comorbidities present that is channeled to determine the Physician.

According to the Strategy of Medical Care required for Patients with a Positive SARS-CoV-2 Result, according to their degree of identified infection, they have the following.

Be: ET P + SARS-CoV-2 i j = Medical Treatment Strategy P + SARS-CoV-2 i j

The medical treatment for a patient with a positive laboratory result for (SARS-CoV-2) COVID-19, based on the Medical Treatment Strategy (ETM P + SARS-CoV-2 ij), based on his degree of infection and present comorbidities, poses two action scenarios : i) Outpatient (SARS-CoV-2) COVID-19 patient or ii) Hospitalized (SARS-CoV-2) COVID-19 patient.
Be:
i) Outpatient COVID19 patient.
P + SARS-CoV-2 i j Outpatient = Positive (SARS-CoV-2) COVID-19 with Outpatient mode in the State, Totonacapan Zone
ii) COVID19 Patient Hospitalized.
P + SARS-CoV-2 i j Hospitalized = Positive (SARS-CoV-2) COVID-19 with modality Hospitalized in the State, Municipality
where:
Depending on the degree of infection (I1, I2 or I3), the Hospitalized (SARS-CoV-2) COVID-19 Patient may require: i) Access to the Intensive Care Area without Intubation or ii) Access to the Intensive Care Area with Intubation.

Definition 6.- Patients with a Positive (SARS-CoV-2) COVID-19 Result Hospitalized with Access to the Intensive Care area.- It is the number of Patients with a Positive SARS-CoV-2 Result Hospitalized with Access to the Intensive Care area, according to its degree of infection.

Be:
P + SARS-CoV-2 i j Hospital Intensive Care = Positive (SARS-CoV-2) COVID-19 with modality Hospitalized in the State, Totonacapan Zone.

Definition 7.- Patients with a positive (SARS-CoV-2) COVID-19 result Hospitalized with access to the Intensive Care Area with Intubation.- It is the number of Patients with a Positive (SARS-CoV-2) COVID-19 Result Hospitalized with Access to the Intensive Care area with Intubation.
Be:
P + SARS-CoV-2 i j Hospital Intensive Care with Intubation = Positive (SARS-CoV-2) COVID-19 with Hospitalized modality and intubation in the State, Totonacapan Zone.

Definition 8.- P + SARS-CoV-2 i j Deaths.- Deaths of Patients with a positive result for SARS-CoV-2. Deaths are all those positive to (SARS-CoV-2) COVID-19 where one is indicated in the data record (DATE_DEF other than the value "99-99-9999").

Definition 9.- (SARS-CoV-2) COVID-19 case fatality rate.- It is the proportion of people who die from (SARS-CoV-2) COVID-19 among the Patients with a positive (SARS-CoV-2) COVID-19 result in a given period and area.
Be:

TL SARS-CoV-2 i j = (SARS-CoV-2) COVID-19 case fatality rate

Where:
(SARS-CoV-2) COVID-19 case fatality rate = [(Deaths of Patients with a Positive (SARS-CoV-2) COVID-19 Result in the State or Municipality) / (Total of Patients with a Positive (SARS-CoV-2) COVID-19 result in the State or Totonacapan Zone)] x 100

Of which:
DP+ SARS-CoV-2 i j = Deaths of Patients with a positive (SARS-CoV-2) COVID-19 result in the State / Municipality
And:
P+ SARS-CoV-2 i j = Total Patients with a positive (SARS-CoV-2) COVID-19 result in the State / Municipality
So:
TL SARS-CoV-2 i j = [D P + SARS-CoV-2 i j / P + SARS-CoV-2 i j] x 100

The data processing corresponds to the records on the epidemic (SARS-CoV-2) COVID-19 at 16 July 2020. The treatment of the information is carried out through the application software for data mining Orange version 3.26.0, in which the algorithm for the information analysis are developed. (See Table 1 and Figure 1, below). According to information from the Ministry of Health, the following records are available at the national level:

1. The number of patients with a positive (SARS-CoV-2) COVID-19 result is 324,041 of which: a) 230,219 are Care Outpatients and b) 93,822 are Hospitalized patients.
2. Of the 93,822 Hospitalized (SARS-CoV-2) COVID-19 Positive patients: a) 7,801 patients enter the Intensive Care Unit; while b) 86,021 patients do not enter the Intensive Care Unit.
3. Only 70 Hospitalized (SARS-CoV-2) COVID-19 Positive patients admitted to the Intensive Care Unit required intubation are speakers of indigenous language, at the national level, of which: a) 26 are women and) 44 are men.
4. Likewise, to date 56 deaths from positive (SARS-CoV-2) COVID-19 patients speakers of indigenous language, at the national level have been registered nationwide.

According to information from the Ministry of Health, in the Totonacapan Zone the following records are available at the national level:

1. The number of patients with a positive (SARS-CoV-2) COVID-19 result is 1,225 of which: a) 518 are women and b) 707 are men. (See Table 2).
2. The total number of women with a positive (SARS-CoV-2) COVID-19, only 6 are women who speak the indigenous language and 512 are women who do not speak the indigenous language. On the other hand, the total number of men with a positive (SARS-CoV-2) COVID-19, 17 are men who speak the indigenous language and 690 are men who do not speak the indigenous language. (See Table 3).
3. 1,225 patients with (SARS-CoV-2) positive COVID-19, according to their primary hospital care, there are: a) 732 are patients who did not require hospitalization and b) 493 are patients who required hospitalization. (See Table 4).
4. 493 patients with (SARS-CoV-2) COVID-19 positive who were hospitalized: a) 15 correspond to speakers of the indigenous language and b) 478 are patients who do not speak the indigenous language. (See Table 4).
5. From the 732 patients with (SARS-CoV-2) COVID-19 positive who were not hospitalized, only 8 were patients speaking the indigenous language and 724 were patients who did not speak the indigenous language. (See Table 4).
6. From the 1,225 patients with positive (SARS-CoV-2) COVID-19, 23 are patients who speak the indigenous language and 1,202 correspond to patients who do not speak the indigenous language. (See Table 3).
7. From the 23 patient patients who speak the indigenous language, these come from: Coatzintla 1 patient, Coxquihui 1, Coyutla 2, Chumatlán 1 and Papantla 12. (See Table 5)
8. From the 493 Hospitalized (SARS-CoV-2) COVID-19 Positive patients: a) 4 patients enter the Intensive Care Unit; while b) 157 patients do not enter the Intensive Care Unit. (See Table 6).
9. Only 4 Hospitalized (SARS-CoV-2) COVID-19 Positive patients admitted to the Intensive Care Unit required intubation.
10. As of July 16, 2020, there have been 3 deaths of hospitalized patients (SARS-CoV-2) COVID-19 whose results were positive and were indigenous language speakers in the Totonacapan zone. (See Figure 3, at the end of the section, where the associated comorbidities of the indigenous language-speaking patients who died in the Totonacapan).

Table 1. Total number of cases in Mexico as of July 16, 2020, According to Sex and Result at (SARS-CoV-2) COVID-1

Result	Sex		Total
	Women	Men	
Positive (SARS-CoV-2) COVID-1	149,637	174,404	324,041
No positive (SARS-CoV-2) COVID-1	198,824	176,631	375,455
Pending result	40,948	41,619	82,567
Total	389,409	392,654	782,063

Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Table 2. Total number of cases in Totonacapan Zone in Mexico as of July 16, 2020, According to Sex and Result at (SARS-CoV-2) COVID-1

Result	Sex		Total
	Women	Men	
Positive (SARS-CoV-2) COVID-1	518	707	1,225
No positive (SARS-CoV-2) COVID-1	251	247	498
Pending result	70	75	145
Total	839	1,029	1,868

Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Table 3. Total number of cases in Totonacapan Zone in Mexico as of July 16, 2020, According to sex, indigenous language speaker status and result positive at (SARS-CoV-2) COVID-19

Result	Women	Men	Total
Indigenous language speaker	6	17	23
Non-indigenous language speaker	507	683	1,190
Not specified	5	7	12
Total	518	707	1,225

Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Table 4. Total number of cases in Totonacapan Zone in Mexico as of July 16, 2020, According to primary hospital care condition, indigenous language speaker status and result positive at (SARS-CoV-2) COVID-19

| Result	Non Hospitalized	Hospitalized	Total
Indigenous language speaker	8	15	23
Non-indigenous language speaker	720	470	1,190
Not specified	4	8	12
Total	732	493	1,225

Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Table 5. Total number of cases in Totonacapan Zone in Mexico as of July 16, 2020, According to primary indigenous language speaker status and result positive at (SARS-CoV-2) COVID-19

Municipality	Speak indigenous language	Non Speak indigenous language	Not specified	Total
Cazones de Herrera	0	14	0	14
Coatzintla	1	102	1	104
Coxquihui	1	2	0	3
Coyutla	2	4	0	6
Chumatlán	1	0	0	1
Espinal	0	8	0	8
Filomeno Mata	0	1	0	1
Gutierrez Zamora	0	20	0	20
Mecatlan	0	1	0	1
Papantla	12	168	2	182
Poza Rica	6	775	9	790
Tecolutla	0	8	0	8
Tihuatlán	0	87	0	87
Total	23	1,190	12	1,225

Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Table 6. Total number of cases in Totonacapan Zone in Mexico as of July 16, 2020, According to primary indigenous language speaker status, result positive at (SARS-CoV-2) COVID-19 and status according to primary hospital care condition in Intensive care unit

| Municipality	Intensive care unit	Non Intensive care unit	Not specified	Total
Cazones de Herrera	0	4	0	4
Coatzintla	1	8	1	10
Coxquihui	0	1	0	1
Coyutla	0	2	0	2
Espinal	0	1	0	1
Gutierrez Zamora	0	4	0	4
Papantla	3	13	0	16
Poza Rica	0	100	2	102
Tihuatlán	0	21	0	21
Total	4	154	3	161

Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Figure 1. Algorithm for the identification of patients
(SARS-CoV-2) COVID 19 in the Totonacapan Zone in Mexico
Orange Data Mining version 3.26.0
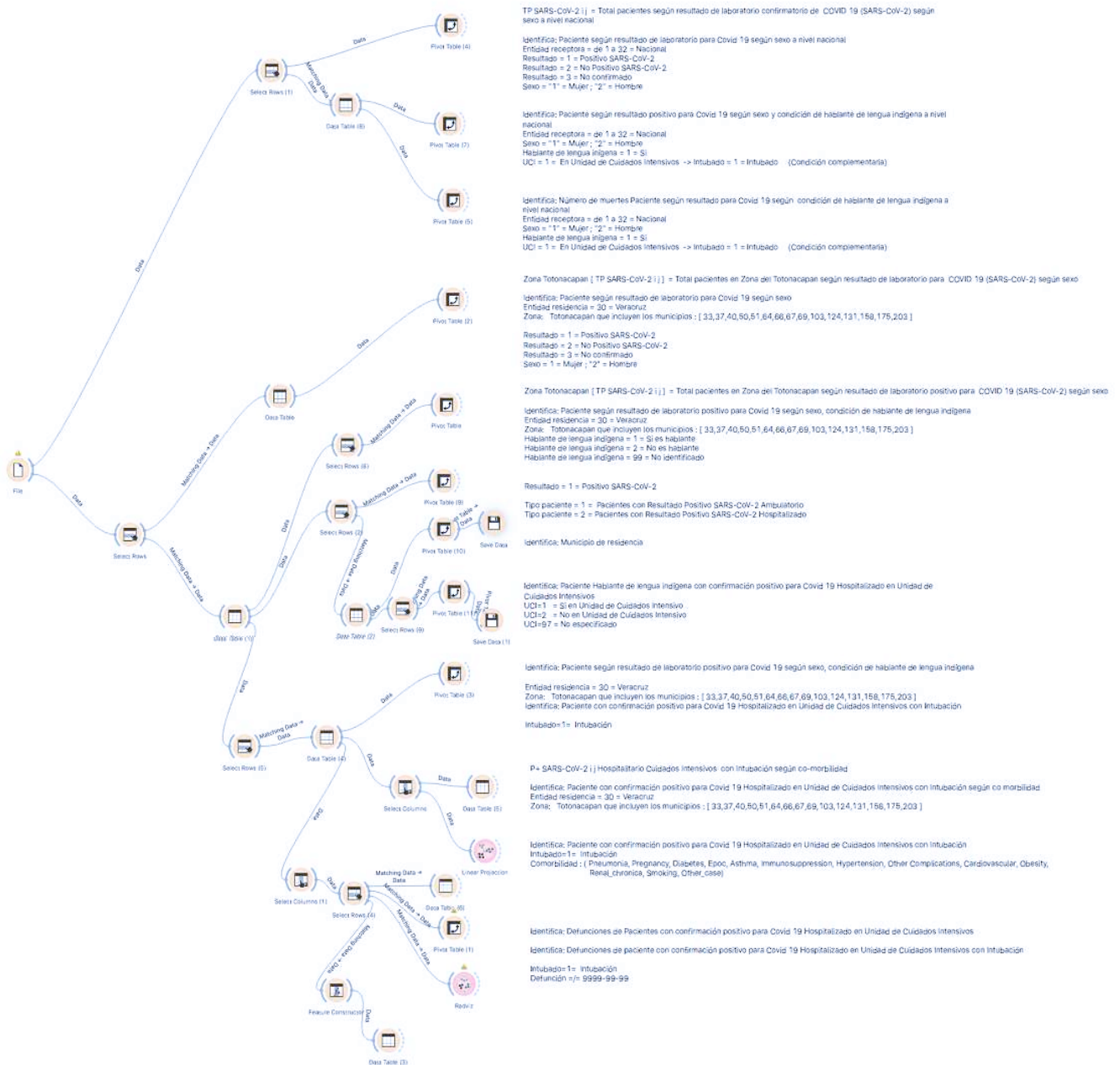
Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020 and Software Orange Data Mining version 3.26.0.

Figure 2. Comorbidities in patients with a positive result in (SARS-CoV-2) COVID-19, speakers of the indigenous language in the Totonapacapan area, according to sex and age groups

Figure 3. Comorbidities in deceased patients who presented with a positive result for COVID-19 (SARS-CoV-2), speakers of the indigenous language in the Totonapacapan area, according to sex and age groups
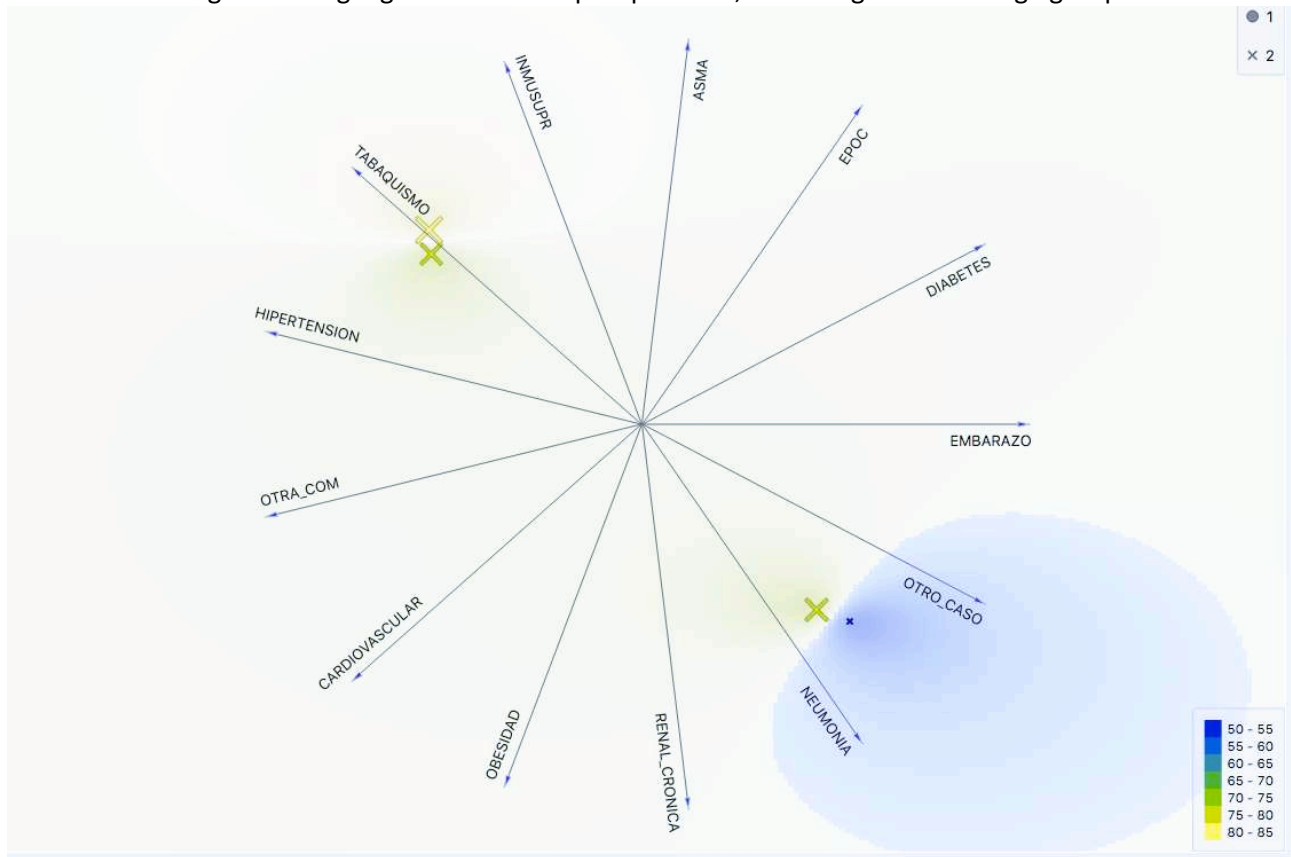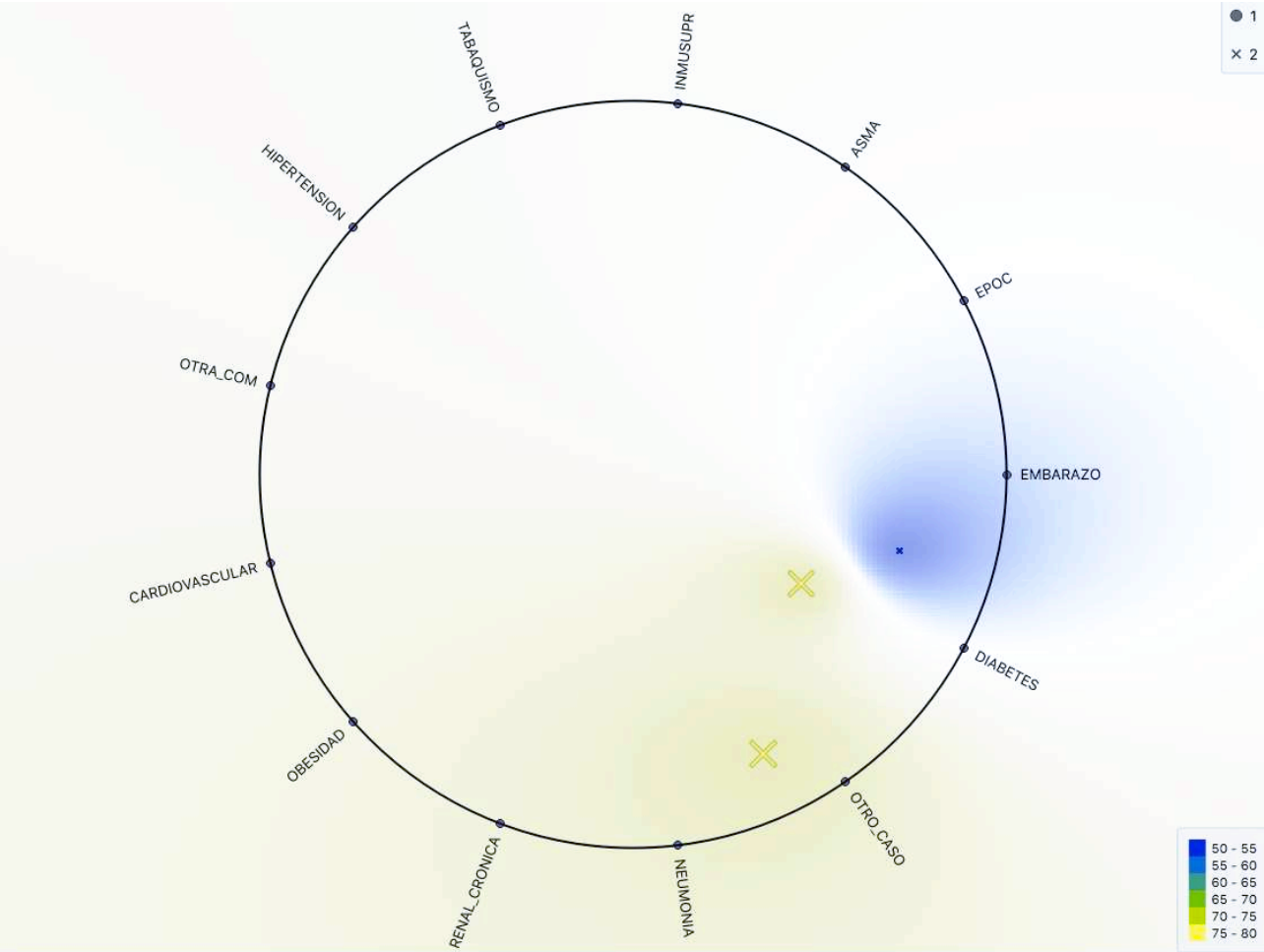
Technical note: 1 = Woman (o), 2 = Man (x)
Source: Own elaboration with Government of Information from the Mexico. Health Secretary. Epidemiological Surveillance System for Viral Respiratory Diseases as of July 16, 2020

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

References

[1] Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python, Journal of Machine Learning Research 14(Aug): 2349–2353. https://dl.acm.org/doi/pdf/10.5555/2567709.2567736

[2] Government of Mexico. Health Secretary. Databases Covid-19 México. https://datos.gob.mx/busca/dataset/informacion-referente-a-casos-covid-19-en-mexico/resource/e8c7079c-dc2a-4b6e-8035-08042ed37165

[3] Medel-Ramírez, Carlos and Medel-López, Hilario, (2020), Data Article. Data mining for the study of the Epidemic (SARS- CoV-2) COVID-19: Algorithm for the identification of patients (SARS-CoV-2) COVID 19 in Mexico, MPRA Paper, University Library of Munich, Germany, https://EconPapers.repec.org/RePEc:pra:mprapa:100888

[4] Software Orange Data Mining version 3.25.1 https://orange.biolab.si

[5] World Health Organization. (2020). Laboratory testing for coronavirus disease (COVID-19) in suspected human cases. Interim guidance. 19 March 2020. Recuperado de: https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117

Reference to a dataset:
[6] Raw data can be retrieved from the Github repository https://github.com/CMedelR