UNIVERSIDAD PANAMERICANA

Subject: Knowledge Management and Analysis(COM145)

Professor: Sarahí Aguilar González

Delivery date: 05/24/22

Cycle: 1222

Project name: Quality of life rating in Mexico

Team Members		
ID	**Name**	**Career**
0217201	Carlos Cabrera Castrejon	TI
0211956	Víctor Manuel Goya Constantino	TI

Rubrics				
ID	2-social		7-knowledge	
	D	C	A	JI

Víctor Manuel Goya Constantino ID: 0211956, Carlos Cabrera Castrejón ID: 0217201

Universidad Panamericana

Final Report - Knowledge Management and Analysis

Abstract.- During the realization of this project we were asked to carry out a project where the problem and the question was posed: Is it possible to know the ratio of the municipalities/municipalities through different indicators?

Keywords - municipal development index, data, python, datasets, rating.

INTRODUCTION

Is it possible to know the "qualification" of the municipalities through different indicators? This was the question that we have decided to ask, taking into consideration different points, such as several types of variables that are calculated in the municipal development index, taking into account four main features of the municipality: institutional, social, economic and environmental.

These 4 factors are evaluated with different variables that will be described in more detail later in the project.
This rubric will give us the guideline to be able to make the rating of the mayor's office/municipality.

PROBLEM TO BE SOLVED

There is no doubt that in the country there are many places that people do not know, this is due to the size of it, as the years go by technology grows with us, and this helps us to have access to information about these places, all this is done in order to make a measurement to "qualify" the quality of life/development that could exist in the municipalities/municipalities, the reason for this index is that we can not ignore that there are so many places in the country, we realize that there are places that are not safe to visit or we will find ourselves in trouble in case of an emergency, so we seek that with this rating we give information to all people about what to expect from this place through the rubric that is considered in the same rating.

PURPOSE OF REPORT

The purpose of this report is to show how we put together the rubric, as well as to explain the process involved in the development of this research, so that the existence of these ratings can be reliably supported, as well as to logically show the decision for which we chose these items for the rating.

SCOPE

During the project we tried to cover most of the topics that were seen throughout the semester in the subject of knowledge management and analysis, all this so that the document provides the necessary information for the correct understanding of the project, as well as that anyone who uses the document can use it as a reference in case they need to make use of any of the tools that were used in the realization of the project.

PROJECT REPORT

I.   APPROACH

At the beginning of the project we looked for something like a housing qualification, where taking aspects that we can find through some data, we can make a calculation that serves as a qualification for the municipalities. Our point of reference was the UN-Habitat CPI, which is a methodology for understanding, analyzing, planning, taking action and observing the effects of public policies on citizen well-being. It translates well-being into a measurable metric, and measures the efficiency of the city and the effect of public policies over time.

Our first goal in starting this project was to define the options on how we would do a similar version of this same methodology, but with different rubrics, so we decided to take into account four municipal development indexes.

II.   INVESTIGATION

The first step to have the right approach to our way of answering the question, we then decided to base our project on four municipal development indices (MDI).

- Institutional Municipal Development Index (MDI-I).
- Municipal Social Development Index (MDI-S).
- Municipal Economic Development Index (MDI-E).
- Municipal Environmental Development Index (MDI-MA).

Institutional Municipal Development Index (MDI-I). seeks to measure performance in the governmental area of the city/municipality through tax effort, transparency, electoral participation, number of municipal employees per capita, and safety. We from this category have decided to use:

- Employees per Capita.

$$EpC = \frac{ElAP}{PM} \times 1000$$

- EpC = Employees per Capita
- ElAP = Employees working in public administrations
- PM = Municipal population

- Percentage of Own Income.

$$PoInPr = \frac{IpM}{IbM} \times 100$$

- PoInPr = Percentage of own income
- IpM = City/municipality's own revenues
- IbM = City/municipality's own income

- Poverty

$$P = \frac{PPM}{PM} \times 1000$$

- P = Poverty
- PPM = Number of people in poverty
- PM = Municipal population

Municipal Social Development Index (MDI-S) seeks to measure the socioeconomic conditions of the population through eleven variables that fall into three basic dimensions: education, health and goods, assets and basic services in the home.

- Alphabetization

$$A = \frac{PqA}{Pq} \times 100$$

- A = Alphabetization
- PqA = Population 15 years of age and over literate
- Pq = Population 15 years and older

- School attendance

$$AE = \frac{PcAE}{Pc} \times 100$$

- AE = School attendance
- PcAE = Population from 6 to 14 years old that exists at school
- Pc = Population from 6 to 14 years old

- Basic and Post Basic Education

$$EBP = \frac{PqEBC}{Pq} \times 100$$

- EBP = Basic and Post Basic Education
- PqEBC = Population 15 years of age and older with completed basic and post-basic education
- Pq = Population 15 years and older

- Eligibility

$$Dh = \frac{PDh}{PM} \times 100$$

- Dh = Eligibility
- PDh = Population with entitlement
- PM = Municipal population

- Different Floor To Ground

$$PDT = \frac{VpDT}{TVH} \times 100$$

- PDT = Different Floor To Ground
- VpDT = Dwellings with floors of material other than earth
- TVH = Total number of inhabited dwellings

- ToiSan

$$ES = \frac{VES}{TVH} \times 100$$

- ES = Toilet or Sanitary
- VES = Dwellings that have a toilet or sanitary facilities
- TVH = Total number of inhabited dwellings

- PipedWater

$$AE = \frac{VAE}{TVH} x100$$

- AE = Piped water

- VAE = Dwellings that have piped water from the public water mains
- TVH= Total number of inhabited dwellings

- Drainage

$$D = \frac{VD}{TVH} \times 100$$

- D = Drainage
- VD = Dwellings with drainage
- TVH = Total number of inhabited dwellings

- ElecPow

$$EE = \frac{VEE}{TVH} \times 100$$

- EE = Electric power
- VEE = Dwellings with electric power
- TVH = Total number of inhabited dwellings

- Washing machine

$$L = \frac{VL}{TVH} \times 100$$

- L = Washing machine
- VL = Dwellings with washing machine
- TVH = Total number of inhabited dwellings

- Refrigerator

$$R = \frac{VR}{TVH} \times 100$$

- R = Refrigerator
- VR = Dwellings with refrigerator
- TVH = Total number of inhabited dwellings

Municipal Economic Development Index (MDI-E) seeks to measure the economic development of the city/municipalities through four elements that help in this measurement, which are the gross census value added; the value of agricultural, livestock and forestry production; the employed population; and the permanent and temporary workers added to the Mexican Social Security Institute (IMSS):

- Percentage of population employed

$$PPO = \frac{PO}{Pd} \times 100$$

- PPO = Percentage of population employed
- PO = Employed Population
- Pd = Population 12 years old and over

- Insured workers

$$TA = \frac{TAPE}{Pd} \times 1000$$

- TA = Insured workers
- TAPE = Permanent and temporary insured workers
- Pd = Population 12 years old and over

Municipal environmental development index (MDI-MA) is organized according to the pressure-state-response model and is calculated using twelve thematic indicators, we decided to make use of other indicators, which are:

- Water supply coverage

$$CAA = \frac{VAE}{TVH} \times 100$$

- CAA = Water supply coverage
- VAE = Dwellings with piped water from the public water mains
- TVH = Total number of inhabited dwellings

- Drainage coverage

$$CD = \frac{VD}{TVH} \times 100$$

- CD = Drainage coverage
- VD = Dwellings with drainage
- TVH = Total number of inhabited dwellings

III. PROCESS

fter the exhaustive analysis of the data and options, it was agreed to make the process plan, which involves the development of the project, then the first step was to get the CSV containing the data that would help us in the development of the project. The data were obtained mostly through INEGI, they are a conglomerate of six datasets that we used, these were exported to SQL, where we used this language to make a union of these data, because there were some details with some municipalities that had not been censused.

Once the data was cleaned, we opted to make all this data in a single CSV file in which all the data we used from the six datasets used are enclosed. We proceed to do now in Python the general calculation of the IDM. To do the calculation of this we calculated the average of each MDI (MDI-I, MDI-S, MDI-E, MDI-MA).

We continue in such a way that once we get the MDI we will make a CSV with this data, which will leave us with an MDI for each municipality in the country, this would already be the calculation of the qualification that we originally wanted to achieve with our research question.

Finally, we made this model in Python, where the results of this project are explained graphically, which had a high degree of complexity because we did not expect to have to work with so many datasets.

In the model, once we got this CSV of the MDI, they were handled as dataframes, likewise these had to be normalized by means of the average, this so that everything can be handled in the same way and there are no discrepancies between the types of data. Once the normalization was achieved, we proceeded to divide the data into two datasets, one for training and the other for testing, of which 80% was for training and the remaining 20% for testing.

To continue with the multiple linear regression, in which we obtained the necessary coefficients to be able to make the prediction of the model, once this prediction is made, we continue to evaluate it, so we decided to get the error and the percentage error, so that in this case we can make a final test, where we add a new data frame where we test it with the model and thus have the result of the model, where we could have a range calculated with the error.

CONCLUSIONS

As for the result of the project, the expected result was achieved because we were able to do this supervised problem of multiple linear regression, the research question was also answered, which is the main objective of the project, and we also fulfilled our own expectations about the complexity of the project, because when we saw everything we used we realized that it is a project that can be very robust, Even if we had been able to obtain more data we could have done it in a more precise way, because there are municipalities where the census is not completed or not taken directly, this could be attributed to the fact that in many parts of the states there is no security for the people to take these censuses and that there are states in a level of extreme poverty where the censuses are not taken or they are simply considered isolated from the country in general and are isolated from their jurisdiction. Similarly, to conclude, we were able to make models that showed that it is feasible to answer: is it possible to know the ratio of the

municipalities/municipalities through different indicators, so we consider this project satisfactory.

REFERENCES

[1]
METODOLOGÍA DEL ÍNDICE DE DESARROLLO MUNICIPAL (IDM)
https://iieg.gob.mx/ns/wp-content/uploads/2020/08/Metodologia_IDM.pdf

[2]
PRESENTACIÓN DE RESULTADOS
https://www.inegi.org.mx/contenidos/programas/ccpv/2020/doc/cpv2020_pres_res_ags.pdf

[3]
PRINCIPALES RESULTADOS POR LOCALIDAD (ITER)
https://www.inegi.org.mx/contenidos/programas/ccpv/2020/doc/fd_iter_cpv2020.pdf

[4]
METODOLOGÍA DE CÁLCULO
ÍNDICE DE DESARROLLO MUNICIPAL BÁSICO (IDMb)
http://www.inafed.gob.mx/work/siha_2015/4/MetodologiaIDMb-Feb07.pdf

[5]
LIFE EXPECTANCY ACROSS THE UNITED STATES
https://github.com/taylorsyde/USA-Life-Expectancy

[6]
COUNTRY PROFILING USING PCA AND CLUSTERING
https://github.com/leo-you/Country-PCA-and-Custering

[7]
LINEAR REGRESSION IN PYTHON WITH SCIKIT-LEARN
https://stackabuse.com/linear-regression-in-python-with-scikit-learn/

[8]
DESCRIPCIÓN DE LA BASE DE DATOS DEL INEGI
Principales resultados por localidad (ITER) (inegi.org.mx)